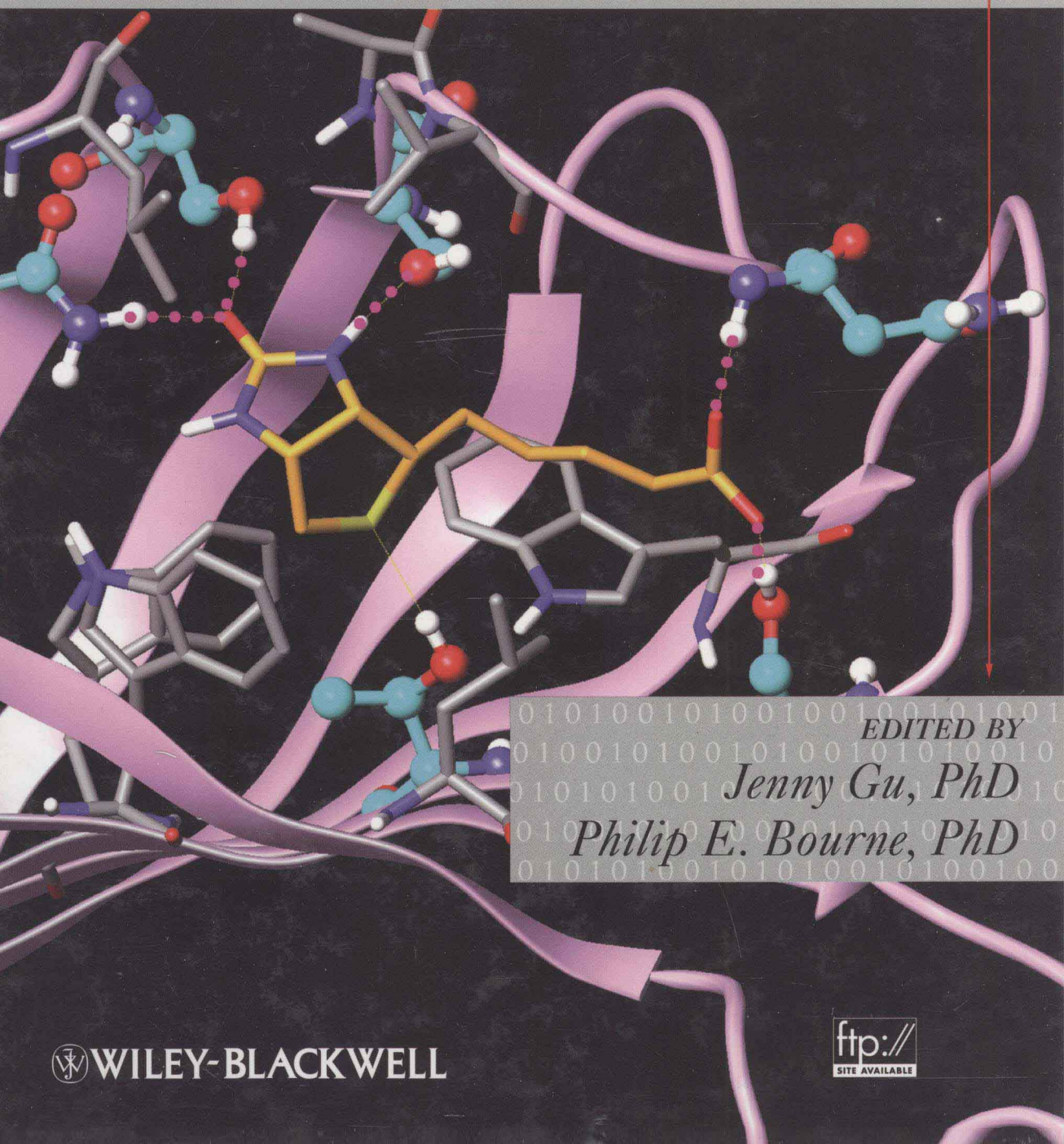


Second Edition

STRUCTURAL BIOINFORMATICS



EDITED BY

Jenny Gu, PhD

Philip E. Bourne, PhD

 WILEY-BLACKWELL

ftp://
SITE AVAILABLE

STRUCTURAL BIOINFORMATICS

Second Edition

Edited By

Jenny Gu, Ph.D.
Philip E. Bourne, Ph.D.

 **WILEY-BLACKWELL**

A JOHN WILEY & SONS, INC., PUBLICATION



Copyright © 2009 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

ISBN 978-0-470-18105-8

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

STRUCTURAL BIOINFORMATICS

Second Edition

FOREWORD

The quality of the coverage and the timeliness of the first edition of *Structural Bioinformatics* led to its wide usage. In turn, the collection has been adopted by the community as a defining articulation for the utility of bioinformatics and structural biology applied to address a range of functional studies in 21st Century Biology. I personally found the book to be read by students, postdoctoral fellows and more senior researchers around the world. The success reflected the excitement in the growing impact of both domains and also helped to accelerate that impact. As pointed out in the Introduction, bioinformatics is now a mainstream activity within the biological sciences; similarly, the implementation of the structural genomics initiative worldwide as a logical and necessary follow up to the Human Genome Project represents the consensus that structure is indeed important for understanding the mechanisms of molecular function. The increased impact and rapid progress during the 6 years, since the first edition was published, demonstrate the extraordinary significance of the interface between structural biology and informatics. Such significant advances indicate the clear need for an update, which has now been provided through the editorial efforts of Gu and Bourne and the writing team of leading researchers who bring the reader to the edge of the frontier.

The second edition, simultaneously as a textbook and an expert monograph, contains a balanced set of contributions, which include updates and advances over the past 6 years and new innovative domains made possible by the sustained application of structural bioinformatics, which were undoubtedly catalyzed by the first edition. Just as was notable about the first edition, this new comprehensive collection fully captures the spirit of excitement at the “bleeding edge” of biodiscovery. The frontier between computing and biology itself reflects the decades of extraordinary progress and revolutionary advances in both domains. During the past 6 years, the data provided by the completion of the full human and numerous model genomes has accelerated the importance of computing for biology, and brought new funding opportunities and research training needs. At the same time, the deeper insight from complete genome sequencing has been the need to look beyond individual genes to systems and to look across multiple scales of biology, ultimately to establish an integrative view of biology based on experimentation and computation. With funding from within mainstream science programs and the increased recognition by the experimental community, the combined use of information technology and quantitative approaches are central to building an integrative view of biology. The superb collection of articles in this second edition speaks directly and powerfully for the role of structural bioinformatics in this effort for both the basic and the applied life sciences.

Additional academic training programs that include structural bioinformatics have been introduced consistently over the past 6 years. Yet, faculty in top flight research institutions who do research and teach in first class bioinformatics programs remain overwhelmed by the demands; indeed, the challenges of bioinformatics education has been a common theme at

the major professional meetings in the field. Thus, textbooks of the highest quality and clarity are essential today, as we all struggle to determine the right curriculum and the right content to train what will be the first generation of students who are truly bioinformaticians. Today, every young aspiring biologist wants to learn about bioinformatics, as do those training in computer science and other quantitative sciences. Understanding the underlying assumptions and intricacies of any bioinformatics algorithm is necessary for proper usage and interpretation of the results obtained with the tools. To teach the large and growing numbers of young scientists who wish to utilize or even contribute to bioinformatics requires authoritative treatments that provide the basis to pull in a new generation of scientists. Such works must also use the best treatments possible to reach a much larger audience, including mature scientists who wish to retrain themselves, and need to set a standard for training everywhere in the world. This collection admirably meets those goals and is a must read for those entering the field and for all of us committed to understanding the interplay of structure and function. By assembling the best thinkers to address systematically all of the challenges at the next stage of the genome effort, Gu and Bourne have created a book that will serve to educate the next generation who will be the future young investigators who will create the tools required to interpret the ever advancing frontier of biology.

At the same time, while the rapid pace of research progress in structural bioinformatics has driven the need for a second edition, the prehistory of modern structural bioinformatics, as is well described in this book, has been retained to remind the readers of some fundamental challenges that should not be neglected and forgotten. This update is not an extensive replicate with minor tweaks of the first edition; instead, the role of historical context and origins of structural bioinformatics as they contribute to present advances are discussed, such as the biannual Critical Assessment of Structure Prediction (CASP) and the initial “exponentiation” of progress in biology resulting from the application of advanced computing technology to structural biology. For more historical content, I refer any reader to the forward for the first edition of *Structural Bioinformatics* or any of my status reports (over the past decade) on computational biology, readily accessible via the internet (or PubMed) in today’s world of “e-knowledge.”

What is most notable and important for the potential readership of this edition is that the fields of computational biology and bioinformatics have gone from uncertainty and neglect to the buzz words on everyone’s lips in less than 10 years. While newly created, contemporary bioinformatics and computational biology training programs are in the process of being incorporated into every biological science domain, and those working at the interface will have increased options even within core disciplines. The reason is obvious: we are already living in the future as biologists; we feel fully the impact of having completely sequenced genomes; we are a part of the transition to high throughput and high information content biological research, and to asking global or systemic questions as the norm, rather than using individual macromolecule-specific probes. Early in this century, the extraordinary, early, and even unanticipated successes of the genome project enabled computer search and modeling techniques to open up new vistas in biology.

The continuing availability of complete genomes coupled with high throughput experimental biological methods, structure determination, and annotated databases will definitely advance our current understanding of protein structures as they relate to biological function, processes, and evolution—a basic research curiosity that has captured central stage in the community. The challenge is to decode the rich information content implicit in genomes and apply the resulting knowledge in service to society. Obtaining a better understanding of biological processes will be achieved through the integration of a variety

of methods including the use of structural information, and should be extended in the future to provide improved health care delivery. At the same time, we are only learning how best to exploit computer and information technology to understand biological mechanisms; the educational content of *Structural Bioinformatics* provides a perspective on our progress and educational content to enable the full potential of systematic computational analysis.

Those of us concerned with macromolecular structure, and protein science in particular, have long spoken the mantra: form follows function, a given function requires a specific structure, or, conversely, structure in turn can be seen to determine function. That is, if we know the structure, we can infer many aspects of biochemical and sometimes, even cellular function which can subsequently be experimentally tested. Indeed, given that we now “know” many gene sequences gained implicitly from sequenced genomes, such resources provide the basis for more refined algorithms that either leverage structural information or improve our understanding of protein structures to model them explicitly and more accurately. Subsequently, these improvements facilitate our ability to predict functionality and greatly reduce the search space for experimental efforts by providing a guided focus to test only the most likely functions. Furthermore, computational modeling of these molecules, for both static and dynamic processes, can provide a detailed description of biological processes at the atomic level, an alternative to traditional biological cartoons which have been the descriptive ways in which we biologists think.

The field of structural bioinformatics, to connect the abstract to the practical, continues to push boundaries beyond what was previously thought impossible. The basis for all structural bioinformatics, the central community database for structural biology, is the Protein Data Bank (PDB) and the macromolecule structures it contains. The growth of this resource has been accelerated by structural genomics initiatives, which continue to sustain and enhance the discovery of novel structures and folds paving the way for new opportunities of discovery and insight through structural bioinformatics. Some examples include a more accurate genome annotation and the basis for clarifying evolutionary related questions. The book addresses key points for cutting edge research such as these, beginning with definitions, and conveys the current scope of research and knowledge of protein structures.

Central to this wonderful collection and insightful articles are advances that have been made in building the infrastructure for an integrative approach to understanding biosystems through the power of understanding protein structures and the implications for the mechanics of function. A survey of current resources is provided to highlight the foundation where future developments are needed to integrate experimental data better and provide the basis for abstraction and generalization. Overall, the individual chapters outline the suite of major basic life science questions such as the status of efforts to predict protein structure and how proteins carry out cellular functions, and also the applied life science questions such as how structural bioinformatics can improve health care through accelerating drug discovery. Dictated by the process of uncovering the mechanisms through which macromolecules act, this journey of discovery into the regulation of life's processes will keep biologists entertained for centuries to come. The second edition book is a great guidebook, even more informative than the earlier collection, and represents the basis for this journey. I highly recommend it to all members of our community.

John C. Wooley
Associate Vice Chancellor, Research
University of California, San Diego
La Jolla, CA

PREFACE

Six years have elapsed since the first edition of this book was published. The field of structural bioinformatics has sustained a high level of excitement in that time, leading to innovative developments and considerable progress throughout the topics covered in the first edition and in the extension to many new domains. Through the efforts of the authors of this new edition, we have tried to capture these developments and to provide an accurate, detailed view of the current field. One way of picturing the advances or defining the “structural” change is relatively straightforward; namely, the number of experimental macromolecular structures has doubled since the first edition of this book was published. The Protein Structure Initiative has also led to an increase in the number of novel structures and folds. Overall, the continued growth in experimental structures has created an even richer data source for much of the work described herein. But numbers do not tell the whole story. The complexity of structures, the methods used, the ways structure is represented, our ability to model structures, our understanding of proteomes and their structural coverage, and so on, have also changed.

Describing the advances in “bioinformatics” per se is more difficult. Change in this case reflects both scientific advances and an increase in recognition within the biological sciences for the importance of computational methods. Due in part to the explosion in high-throughput experimental methods, bioinformatics is certainly more mainstream than it was 6 years ago and most experimental (i.e., non-computational) life scientists would acknowledge the role bioinformatics now plays in furthering our understanding of living systems. Some years from now, whether or not bioinformatics will exist as a separate entity, rather than as a core effort in every biological science department, is a subject for debate. What is important here is that there is an active effort to apply computational methods to a rapidly growing corpus of macromolecular structure data. Our primary goal is to provide a comprehensive description of what this field has accomplished to date and to make the reader aware of what we have gained and could gain in the future toward our understanding of living systems through the study of macromolecular structure and the continued, rigorous application of bioinformatics. As such, this edition should provide a fully current, useful reference to those already in the field, and a suitable text for those educating others. The first edition already encouraged new scholars to enter the field and we believe the case for engaging in structural bioinformatics is stronger than ever.

To meet this goal, the second edition includes not only updated chapters, but also new chapters covering mass spectrometry, genome annotation, immunology, protein dynamics and disorder, membrane proteins, protein design capabilities, and evolutionary biology as they relate to macromolecular structure.

Macromolecular structure is often underappreciated and bypassed in practice during the current era of high-throughput biology, especially since researchers can jump directly from

genomic sequence to phenotype and conduct biochemical studies on large-scale protein–protein interactions that are often involved in pathways associated with disease states. While a great deal can be learned from such studies, ultimately the devil is in the details which do not arise from traditional functional studies; the field of molecular biophysics, now termed structural biology, came into existence to obtain and use structures to provide those details. We believe structure will play an ever increasingly important role as genome studies seek to explore deeper into the mechanisms of life. As such, computational approaches that analyze structure are essential and we hope this book will be there to guide you.

We begin by describing the scope of this book and the history of the field (Chapter 1). The remainder of the introductory Section I is devoted to the understanding of the data itself, namely protein, DNA and RNA structure, respectively (Chapters 2 and 3). Understanding the nuances (scope, accuracy, completeness, etc.) of structural data is prerequisite to any effective use of that data. Effective data use in turn requires an understanding of the experiments or experimental method that produce the data. The most popular methods for deriving macromolecular structure data are, in order, X-ray crystallography (Chapter 4), NMR spectroscopy (Chapter 5), and electron microscopy (Chapter 6). Constructing structural models of molecules can also be guided with hydrogen–deuterium and cross-linking experiments coupled with mass spectrometry (Chapter 7). The raw data from these methods are most often a set of Cartesian coordinates representing the positions of the atoms in these structures, which are well suited for analysis by computer, but alternative representations of this information rich content are sometimes needed to conduct wide-scale bioinformatics analysis (Chapter 8). That is, structural biology and structural bioinformatics are inherently visual sciences—the tabular output of atomic coordinates can be useful as input for computation, but not for human insight. The visualization of structure has evolved along with the science and many useful tools, mostly free, are available (Chapter 9).

In the early days of structural biology (up to the late 1970s), those in the field could name all the structures that had been solved, some of which had Nobel prizes attached to them. As the field grew this was no longer possible, and databases of structure data began to appear. Consistent use of structural data contained within these databases (and indeed the construction of the databases themselves) requires consistent data representation and Section II is devoted to this topic. Chapter 10 introduces the common data representations used by today's software. The field is very fortunate to have scientists who recognize the importance of having a single source of primary data, the worldwide PDB (wwPDB—Chapter 11), from which a variety of secondary resources are derived. Examples of such resources are provided in Chapters 12 and 13.

As the number of structures has increased, much can be learnt from comparative analysis (Section III), where similarities and differences provide new insights. Chapters 14 and 15 describe structure validation, which is important in understanding the accuracy of the data you are dealing with before a 3D comparison and alignment of structures can be made (Chapter 16). When structure comparisons are made and similarities found, reductionism can be applied to make sense of the vast amount of data. Such reductionism leads to classification in various ways, such as by fold, domain, family, and superfamily (Chapters 17 and 18).

The more we know from comparing structures the more we can learn about structure and functional assignment (Section IV). Secondary structure assignment can now be made consistently and reliably for the majority of structures (Chapter 19). Proteins exist as one or more domains or compact structural and functional units. Hence, automated assignment of domains is important (Chapter 20). Through the structural genomics projects and the NIH

Protein Structure Initiative, structure determination is moving from a functional to a genomic initiative. That is, structures were traditionally determined in an effort to elucidate further details about a known function, and these structural efforts were established based on very extensive prior biological, biochemical and often genetic research, and were done in parallel with continuing biological research on functional properties. In contrast, high-resolution structures with no elucidated or known function are being determined at an accelerated pace, thus making functional assignment critical (Chapter 21). The use of structural information to identify distantly related proteins also serves in annotating genomes (Chapter 22) and clarifying evolutionary relationships (Chapter 23).

Proteins do not act in isolation, that is, most proteins do not function by themselves but act as the result of complex protein–protein, protein–ligand and protein–solvent interactions and are often part of larger macromolecular assemblies. Section V describes these interactions beginning with an introduction to electrostatic forces that have a fundamental impact on recognition between molecules (Chapter 24). The majority of these interactions are not captured in the experimental structure of a complex, but as an apo form of the structure with a signature that can be teased out to predict that interaction. Understanding these signatures when found in protein–DNA and protein–RNA interactions (Chapter 25) and in protein–protein interactions (Chapter 26) aids, for example, in the identification of new transcription sites and reconstruction of protein signaling networks. After the sites of interactions are identified, docking of the molecules is simplified, which is important in drug design (Chapter 27).

While the number of structures is increasing rapidly, the number of protein sequences is increasing much more rapidly; thus, the idea of predicting protein structure from its sequence remains an “obsessive” goal (Section VI). Spurred by an unusual biannual competition, referred to as CASP—the Critical Assessment of Protein Structure Prediction (Chapter 28), progress is being made in subcategories of structure prediction efforts within CASP and the field in general. Structure prediction categories include homology modeling (Chapter 30), fold recognition (Chapter 31) and *ab initio* structure prediction (Chapter 32). Other forms of prediction include secondary structure and membrane components for proteins (Chapter 29). Advances in understanding and predicting RNA structures have also been made and are discussed in Chapter 33.

Structural bioinformatics is playing an increasingly important role in the development of new pharmaceuticals (Section VII). The identification of drug targets, understanding the action of drug binding, and the design of promising leads all involve structural bioinformatics (Chapter 34). In addition to the development of small molecule and peptide-based drugs, contributions are also being made in identifying antigen recognition sites that aid in antibody-based therapeutics (Chapter 35).

Finally, Section VIII identifies challenges at the frontiers of structural bioinformatics. Membrane associated proteins, whose structures are difficult to characterize *in vitro* and thus are underrepresented experimentally, are one example (Chapter 36). Proteins are not static under physiological conditions, yet understanding the dynamics (Chapter 37) and the impact of disorder and conformational variants (Chapter 38), while important to protein function, are all still poorly understood. As our understanding of protein structure improves, so do our design rules and capacity for engineering new proteins for their functions that improve upon nature or provide the potential for novel processes (Chapter 39). The best way to push back these frontiers is with more structures and enhanced generalizations about their roles; structure genomics is doing just that (Chapter 40) and is thus a fitting place to end our tour of structural bioinformatics.

The words that follow are written by many of the leaders in the field and we thank them for their time and energy in sharing what motivates them to unravel the mysteries of nature, which are so beautifully displayed before us in an ever increasing number of macromolecular structures.

Jenny Gu
Philip E. Bourne

Color files of all figures from this book are available for download from the following web address: ftp://ftp.wiley.com/public/sci_tech_med/structural_bioinformatics

ACKNOWLEDGMENTS

“Science does not know its debt to imagination.”
Ralph Waldo Emerson

We are grateful to the University of Texas Medical Branch, the University of Münster, and the University of California at San Diego for institutional support in our research and educational endeavors, including this book. Likewise, we thank the Jeanne Kempner Foundation and the US funding agencies, the National Science Foundation, and the National Institutes of Health for their continual support in the advancement of science. This book could not have been completed without the help of all the contributors who have dedicated their expertise, time, and efforts to this extensive update and indispensable resource for the community—we are greatly indebted to them. Additionally, we would like to thank Wolfgang Bluhm, Kristine Briedis, Naomi Cotton, Lynn Fink, Apostol Gramada, Maria Kontoyianni, Kannan Natarajan, Julia Ponomarenko, Peter Rose, Wayne Townsend-Merino, Ruben Valas, Stella Veretnik, Lei Xie, Song Yang, and Zhanyang Zhu for their assistance and support in reviewing the materials for this book. The successful publication of this edition could not have been achieved without the help and patience of Andrea Baier, Tiffany Williams, and Thomas Moore at Wiley-Blackwell.

On a more personal note, JG would like to thank her family, friends, colleagues and mentors for their continued support in taking on new challenging endeavors. PEB would like to sincerely thank his family, Roma, Melanie, and Scott, for their continued understanding of the role science plays in his life.

CONTRIBUTORS

Paul D. Adams, Physical Biosciences Division, Lawrence Berkeley Laboratory, Berkeley, CA and Department of Bioengineering, University of California, Berkeley, CA

Steven C. Almo, Albert Einstein College of Medicine, Bronx, NY

Russ B. Altman, Departments of Bioengineering & Genetics, Stanford University, Stanford, CA

Claus A. Andersen, Siena Biotech Spa., Siena, Italy

Arash Bahrami, Graduate Program in Biophysics and National Magnetic Resonance Facility at Madison, Biochemistry Department, University of Wisconsin-Madison, Madison, WI

Nathan A. Baker, Department of Biochemistry and Molecular Biophysics, Center for Computational Biology, Washington University, St. Louis, MO

Gail J. Bartlett, School of Chemistry, University of Bristol, Cantock's Close, Clifton, Bristol, UK

Helen M. Berman, RCSB Protein Data Bank, Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ

Jeffrey B. Bonanno, Albert Einstein College of Medicine, Bronx, NY

Richard Bonneau, Department of Biology, New York University, New York, NY; and Department of Computer Science, Courant Institute, New York University, New York, NY

Steven Bottomley, School of Biomedical Sciences, Curtin University of Technology, Perth, Western Australia, Australia

Philip E. Bourne, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA

Natasja Brooijmans, Wyeth, Madison, NJ

Axel T. Brunger, The Howard Hughes Medical Institute and Departments of Molecular and Cellular Physiology, Neurology and Neurological Sciences, Structural Biology, and Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford, CA

Stephen K. Burley, Eli Lilly and Company, San Diego, CA

Emidio Capriotti, Structural Genomics Unit, Bioinformatics and Genomics Department, Centro de Investigación Príncipe Felipe, Valencia, Spain

Mark R. Chance, Case Western Reserve University, Cleveland, OH

- Li Chen**, Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ
- Dylan Chivian**, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA
- Alison Cuff**, Institute of Structural and Molecular Biology, University College London, London, UK
- Joanna de la Cruz**, Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ
- Nikolay V. Dokholyan**, Department of Biochemistry and Biophysics, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC
- Kevin Drew**, Department of Biology, New York University, New York, NY; and Department of Computer Science, Courant Institute, New York University, New York, NY
- Jonathan M. Dugan**, Stanford University, Stanford, CA
- Shuchismita Dutta**, Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ
- Hamid R. Eghbalnia**, Department of Molecular and Cellular Physiology, University of Cincinnati, Cincinnati, OH
- Spencer Emtage**, Eli Lilly and Company, San Diego, CA
- Zukang Feng**, Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ
- J. Lynn Fink**, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA
- Andras Fiser**, Albert Einstein College of Medicine, Bronx, NY
- Paula M. D. Fitzgerald**, Merck Research Laboratories, Boston, MA
- Adam Godzik**, Burnham Institute for Medical Research, La Jolla, CA
- Ralf W. Grosse-Kunstleve**, Physical Biosciences Division, Lawrence Berkeley Laboratory, Berkeley, CA
- Colin R. Groom**, Biochemistry Department, University of Cambridge, Cambridge, UK
- Jenny Gu**, Institute for Evolution and Biodiversity, University of Münster, Münster, Germany
- Dorit Hanein**, Burnham Institute for Medical Research, La Jolla, CA
- Erik Helmerhorst**, School of Biomedical Sciences, Curtin University of Technology, Perth, Western Australia, Australia
- Kim Henrick**, Macromolecular Structural Database, European Bioinformatics Institute, Cambridge, UK
- Alícia Pérez Higuero**, Biochemistry Department, University of Cambridge, Cambridge, UK

Vincent J. Hilser, Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, TX

Thomas Huber, The University of Queensland, School of Molecular and Microbial Sciences, Brisbane, Queensland, Australia

Magdalena A. Jonikas, Department of Bioengineering, Stanford University, Stanford, CA

Gordon King, The University of Queensland, Institute for Molecular Bioscience and ARC Special Research Centre for Functional and Applied Genomics, Brisbane, Queensland, Australia

Michael L. Klein, Center for Molecular Modeling, Department of Chemistry, University of Pennsylvania, Philadelphia, PA

Bostjan Kobe, The University of Queensland, School of Molecular and Microbial Sciences and Institute for Molecular Bioscience and ARC Special Research Centre for Functional and Applied Genomics, Brisbane, Queensland, Australia

Elmar Krieger, CMBI, NCMLS, Radboud University Medical Centre, Nijmegen, The Netherlands

Alain Laederach, Biomedical Sciences Program, School of Public Health, SUNY, Albany, NY

Roman A. Laskowski, European Bioinformatics Institute, Hinxton, Cambridge, UK

Jonathan Lees, Department of Biochemistry & Molecular Biology, University College London, London, UK

Sheng Li, Department of Medicine and Biomedical Sciences Graduate Program, University of California, San Diego, La Jolla, CA

Tong Liu, Department of Medicine and Biomedical Sciences Graduate Program, University of California, San Diego, La Jolla, CA

John L. Markley, Biochemistry Department, University of Wisconsin-Madison, Madison, WI

Marc A. Marti-Renom, Structural Genomics Unit, Bioinformatics and Genomics Department, Centro de Investigación Príncipe Felipe, Valencia, Spain

J. Andrew McCammon, Departments of Chemistry & Biochemistry and Pharmacology, Howard Hughes Medical Institute, University of California, San Diego, La Jolla, CA

Raghu P. R. Metpally, Bioinformatics and In Silico Drug Design Lab., Computer Science Department, Queens College of City University of New York, Flushing, NY

Dmitri Mouradov, The University of Queensland, School of Molecular and Microbial Sciences, Brisbane, Queensland, Australia

Haruki Nakamura, PDBj, Institute for Protein Research, Osaka University, Osaka, Japan

Stephen Neidle, School of Pharmacy, University of London, London, UK

- Christine A. Orengo**, Department of Biochemistry & Molecular Biology, University College London, London, UK
- Florencio Pazos**, Computational Systems Biology Group, National Center for Biotechnology (CNB-CSIC), C/Darwin, Madrid, Spain
- Frances M. G. Pearl**, EMBL/EBI, The Wellcome Trust Genome Campus, Cambridge, UK
- Marialuisa Pellegrini-Calace**, EMBL/EBI, The Wellcome Trust Genome Campus, Cambridge, UK
- Irina Periskova**, Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ
- Francis C. Peterson**, Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI
- William R. Pitt**, UCB Celltech, Slough, UK
- Julia V. Ponomarenko**, Skaggs School of Pharmacy & Pharmaceutical Science, San Diego Supercomputer Center, University of California, San Diego, CA
- Boojala V. B. Reddy**, Bioinformatics and In Silico Drug Design Lab., Computer Science Department, Queens College of City University of New York, Flushing, NY
- Adam J. Reid**, Department of Biochemistry & Molecular Biology, University College London, London, UK
- David C. Richardson**, Department of Biochemistry, Duke University, Durham, NC
- Jane S. Richardson**, Department of Biochemistry, Duke University, Durham, NC
- Timothy Robertson**, Department of Biochemistry, University of Washington, Seattle, WA
- Ian Ross**, The University of Queensland, Institute for Molecular Bioscience and ARC Special Research Centre for Functional and Applied Genomics, Brisbane, Queensland, Australia
- Burkhard Rost**, CUBIC, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY; Columbia University Center for Computational Biology and Bioinformatics (C2B2), New York, NY; Northeast Structural Genomics Consortium (NESG), Columbia University, New York, NY; and Department of Pharmacology, Columbia University, New York, NY
- Andrej Sali**, University of California, San Francisco, CA
- Ilan Samish**, Department of Biochemistry & Biophysics, and Department of Chemistry, University of Pennsylvania, Philadelphia, PA
- J. Michael Sauder**, Eli Lilly and Company, San Diego, CA
- Eric D. Scheeff**, Razavi Newman Center for Bioinformatics, The Salk Institute for Biological Studies, La Jolla, CA
- Bohdan Schneider**, Institute of Biotechnology of the Academy of Sciences of the Czech Republic, Prague, Czech Republic

- Ilya N. Shindyalov**, San Diego Supercomputer Center, San Diego, La Jolla, CA
- Subramanyam Swaminathan**, Brookhaven National Laboratory, Upton, NY
- Janet M. Thornton**, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK
- Robert C. Tyler**, Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI
- Eldon L. Ulrich**, BMRB, Biochemistry Department, University of Wisconsin-Madison, Madison, WI
- Ruben Valas**, Bioinformatics Graduate Program, University of California, San Diego, La Jolla, CA
- Alfonso Valencia**, Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain
- Marc H. V. van Regenmortel**, Biotechnology School of the University of Strasbourg, CNRS, Illkirch, France
- Gabriele Varani**, Departments of Biochemistry and Chemistry, University of Washington, Seattle, WA
- Hanka Venselaar**, CMBI, NCMLS, Radboud University Medical Centre, Nijmegen, The Netherlands
- Stella Veretnik**, San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA
- Brian F. Volkman**, Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI
- Niels Volkmann**, Burnham Institute for Medical Research, La Jolla, CA
- Gert Vriend**, CMBI, NCMLS, Radboud University Medical Centre, Nijmegen, The Netherlands
- James D. Watson**, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK
- Helge Weissig**, Activx Biosciences, San Diego, CA
- John D. Westbrook**, Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ
- William M. Westler**, National Magnetic Resonance Facility at Madison, Biochemistry Department, University of Wisconsin-Madison, Madison, WI
- Shoshana Wodak**, The Hospital for Sick Children, Toronto, Ontario, Canada
- Virgil L. Woods Jr.**, Department of Medicine and Biomedical Sciences Graduate Program, University of California, San Diego, La Jolla, CA
- Huanwang Yang**, Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ