

---

# STRUCTURAL BIOINFORMATICS

---

**Edited By**

Philip E. Bourne

San Diego Supercomputer Center  
Department of Pharmacology  
University of California San Diego  
La Jolla, CA

Helge Weissig

Structural Bioinformatics  
San Diego SuperComputer Center  
University of California San Diego  
La Jolla, CA

 **WILEY-LISS**

A JOHN WILEY & SONS PUBLICATION

E4102 V0144



This book is printed on acid-free paper.

Copyright © 2003 by John Wiley & Sons, Inc. All rights reserved.

Published by Wiley-Liss, 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](http://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, e-mail: [permreq@wiley.com](mailto:permreq@wiley.com).

**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 please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. 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, however, may not be available in electronic format.

Cover Design: Jennifer Matthews

***Library of Congress Cataloging-in-Publication Data:***

Structural bioinformatics / edited by Philip E. Bourne, Helge Weissig.

p. ; cm. – (Methods of biochemical analysis; v. 44)

Includes bibliographical references and index.

ISBN 0-471-20200-2 (cloth : alk. paper)

1. Macromolecules–Structure–Mathematical models. 2. Macromolecules–Structure–Computer simulation. 3. Bioinformatics.

I. Bourne, Philip E. II. Weissig, Helge. III. Series.

[DNLM: 1. Computational Biology. QH 506 S9266 2003]

QP517.M3S776 2003

572.8'733–dc21

2002011156

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

---

# CONTENTS

---

Foreword . . . . .	ix
Preface . . . . .	xiii
Acknowledgments . . . . .	xv
Contributors . . . . .	xvii

## Section I INTRODUCTION 1

### 1 DEFINING BIOINFORMATICS AND STRUCTURAL BIOINFORMATICS 3

*Russ B. Altman and Jonathan M. Dugan*

### 2 FUNDAMENTALS OF PROTEIN STRUCTURE 15

*Eric D. Scheeff and J. Lynn Fink*

### 3 FUNDAMENTALS OF DNA AND RNA STRUCTURE 41

*Stephen Neidle, Bohdan Schneider, and Helen M. Berman*

### 4 COMPUTATIONAL ASPECTS OF HIGH-THROUGHPUT CRYSTALLOGRAPHIC MACROMOLECULAR STRUCTURE DETERMINATION 75

*Paul D. Adams, Ralf W. Grosse-Kunstleve, and Axel T. Brunger*

### 5 MACROMOLECULAR STRUCTURE DETERMINATION BY NMR SPECTROSCOPY 89

*John L. Markley, Eldon L. Ulrich, William M. Westler, and Brian F. Volkman*

### 6 ELECTRON MICROSCOPY 115

*Niels Volkman and Dorit Hanein*

<b>7</b>	<b>MOLECULAR VISUALIZATION</b>	<b>135</b>
	<i>John Tate</i>	
<b>Section II DATA REPRESENTATION AND DATABASES</b>		<b>159</b>
<b>8</b>	<b>THE PDB FORMAT, mmCIF FORMATS, AND OTHER DATA FORMATS</b>	<b>161</b>
	<i>John D. Westbrook and P. M. D. Fitzgerald</i>	
<b>9</b>	<b>THE PROTEIN DATA BANK</b>	<b>181</b>
	<i>The PDB Team</i>	
<b>10</b>	<b>THE NUCLEIC ACID DATABASE</b>	<b>199</b>
	<i>Helen M. Berman, John Westbrook, Zukang Feng, Lisa Iype, Bohdan Schneider, and Christine Zardecki</i>	
<b>11</b>	<b>OTHER STRUCTURE-BASED DATABASES</b>	<b>217</b>
	<i>Helge Weissig and Philip E. Bourne</i>	
<b>Section III COMPARATIVE FEATURES</b>		<b>237</b>
<b>12</b>	<b>PROTEIN STRUCTURE EVOLUTION AND THE SCOP DATABASE</b>	<b>239</b>
	<i>Boojala V. B. Reddy and Philip E. Bourne</i>	
<b>13</b>	<b>THE CATH DOMAIN STRUCTURE DATABASE</b>	<b>249</b>
	<i>C. A. Orengo, F. M. G. Pearl, and J. M. Thornton</i>	
<b>14</b>	<b>STRUCTURAL QUALITY ASSURANCE</b>	<b>273</b>
	<i>Roman A. Laskowski</i>	
<b>15</b>	<b>ALL-ATOM CONTACTS: A NEW APPROACH TO STRUCTURE VALIDATION</b>	<b>305</b>
	<i>Jane S. Richardson</i>	
<b>16</b>	<b>STRUCTURE COMPARISON AND ALIGNMENT</b>	<b>321</b>
	<i>Philip E. Bourne and Ilya N. Shindyalov</i>	



<b>Section IV</b>	<b>STRUCTURE AND FUNCTIONAL ASSIGNMENT</b>	<b>339</b>
<b>17</b>	<b>SECONDARY STRUCTURE ASSIGNMENT</b> <i>Claus A. F. Andersen and Burkhard Rost</i>	<b>341</b>
<b>18</b>	<b>IDENTIFYING STRUCTURAL DOMAINS IN PROTEINS</b> <i>Lorenz Wernisch and Shoshana J. Wodak</i>	<b>365</b>
<b>19</b>	<b>INFERRING PROTEIN FUNCTION FROM STRUCTURE</b> <i>Gail J. Bartlett, Annabel E. Todd, and Janet M. Thornton</i>	<b>387</b>
<b>Section V</b>	<b>PROTEIN INTERACTIONS</b>	<b>409</b>
<b>20</b>	<b>PREDICTION OF PROTEIN–PROTEIN INTERACTIONS FROM EVOLUTIONARY INFORMATION</b> <i>Alfonso Valencia and Florencio Pazos</i>	<b>411</b>
<b>21</b>	<b>ELECTROSTATIC INTERACTIONS</b> <i>Nathan A. Baker and J. Andrew McCammon</i>	<b>427</b>
<b>Section VI</b>	<b>PROTEINS AS DRUG TARGETS</b>	<b>441</b>
<b>22</b>	<b>PRINCIPLES AND METHODS OF DOCKING AND LIGAND DESIGN</b> <i>J. Krumrine, F. Raubacher, N. Brooijmans, and I. Kuntz</i>	<b>443</b>
<b>23</b>	<b>STRUCTURAL BIOINFORMATICS IN DRUG DISCOVERY</b> <i>Eric B. Fauman, Andrew L. Hopkins, and Colin R. Groom</i>	<b>477</b>
<b>Section VII</b>	<b>STRUCTURE PREDICTION</b>	<b>499</b>
<b>24</b>	<b>CASP AND CAFASP EXPERIMENTS AND THEIR FINDINGS</b> <i>Philip E. Bourne</i>	<b>501</b>
<b>25</b>	<b>HOMOLOGY MODELING</b> <i>Elmar Krieger, Sander B. Nabuurs, and Gert Vriend</i>	<b>509</b>

<b>26</b>	<b>FOLD RECOGNITION METHODS</b>	<b>525</b>
	<i>Adam Godzik</i>	
<b>27</b>	<b>AB INITIO METHODS</b>	<b>547</b>
	<i>Dylan Chivian, Timothy Robertson, Richard Bonneau, and David Baker</i>	
<b>28</b>	<b>PREDICTION IN 1D: SECONDARY STRUCTURE, MEMBRANE HELICES, AND ACCESSIBILITY</b>	<b>559</b>
	<i>Burkhard Rost</i>	
	<b>Section VIII THE FUTURE</b>	<b>589</b>
<b>29</b>	<b>STRUCTURAL GENOMICS</b>	<b>591</b>
	<i>Stephen K. Burley and Jeffrey B. Bonanno</i>	
	<b>INDEX</b>	<b>613</b>