STRUCTURAL BIOINFORMATICS

Edited By

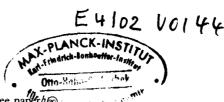
Philip E. Bourne San Diego Supercomputer Center Department of Pharmacology

University of California San Diego La Jolla, CA

Helge Weissig

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





This book is printed on acid-free paper book

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. 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.

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)

Macromolecules-Structure-Mathematical models.
 Macromolecules-Structure-Computer simulation.
 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.

10987654321



CONTENTS

Forev	word	ix
Prefa	nce	xiii
Ackn	owledgments	xv
Cont	ributors	xvii
Sec	etion I Introduction	1
1	DEFINING BIOINFORMATICS AND STRUCTURAL BIOINFORMATICS	3
	Russ B. Altman and Jonathan M. Dugan	
2	FUNDAMENTALS OF PROTEIN STRUCTURE Eric D. Scheeff and J. Lynn Fink	15
3	FUNDAMENTALS OF DNA AND RNA STRUCTURE Stephen Neidle, Bohdan Schneider, and Helen M. Berman	41
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 Volkmann and Dorit Hanein	

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



CONTENTS

Se	$\operatorname{ction} \operatorname{IV}$ structure and functional assignment	339
17	SECONDARY STRUCTURE ASSIGNMENT Claus A. F. Andersen and Burkhard Rost	341
18	IDENTIFYING STRUCTURAL DOMAINS IN PROTEINS Lorenz Wernisch and Shoshana J. Wodak	365
19	INFERRING PROTEIN FUNCTION FROM STRUCTURE Gail J. Bartlett, Annabel E. Todd, and Janet M. Thornton	387
Se	ction V protein interactions	409
20	PREDICTION OF PROTEIN-PROTEIN INTERACTIONS FROM EVOLUTIONARY INFORMATION Alfonso Valencia and Florencio Pazos	411
21	ELECTROSTATIC INTERACTIONS Nathan A. Baker and J. Andrew McCammon	427
Se	${ m ction~VI}$ proteins as drug targets	441
22	PRINCIPLES AND METHODS OF DOCKING AND LIGAND DESIGN J. Krumrine, F. Raubacher, N. Brooijmans, and I. Kuntz	443
23	STRUCTURAL BIOINFORMATICS IN DRUG DISCOVERY Eric B. Fauman, Andrew L. Hopkins, and Colin R. Groom	477
Se	ction VII STRUCTURE PREDICTION	499
24	CASP AND CAFASP EXPERIMENTS AND THEIR FINDINGS Philip E. Bourne	501
25	HOMOLOGY MODELING Elmar Krieger, Sander B. Nabuurs, and Gert Vriend	509

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