

STRUCTURAL PROTEOMICS

and its impact on the
LIFE SCIENCES

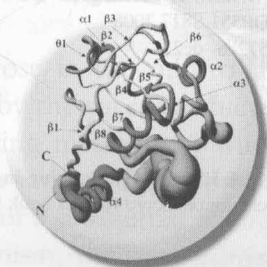
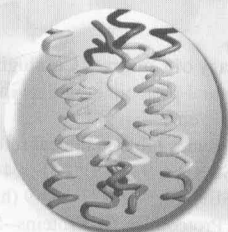
Joel L. Sussman
Israel Silman
editors



STRUCTURAL PROTEOMICS

and its impact on the

LIFE SCIENCES



Joel L. Sussman, Israel Silman

Weizmann Institute of Science, Israel

editors

 **World Scientific**

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data

Structural proteomics and its impact on the life sciences / editors, Joel L. Sussman, Israel Silman.

p. ; cm.

Includes bibliographical references and index.

ISBN-13: 978-981-277-204-6 (hardcover : alk. paper)

ISBN-10: 981-277-204-9 (hardcover : alk. paper)

1. Proteomics. 2. Proteins--Structure. I. Sussman, Joel. II. Silman, Israel.

[DNLM: 1. Proteomics. QU 58.5 S932 2008]

QP551.S817 2008

572'.6--dc22

2008009897

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Copyright © 2008 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

Typeset by Stallion Press

Email: enquiries@stallionpress.com

Printed by FuIsland Offset Printing (S) Pte Ltd, Singapore

Joel L. אהרון, יז'ואל סילמן
Weizmann Institute of Science

STRUCTURAL PROTEOMICS

and its impact on the
LIFE SCIENCES



Joel L. אהרון, יז'ואל סילמן
Weizmann Institute of Science, Israel
editors

 World Scientific

HONG KONG • LONDON • NEW DELHI • BEIJING • SINGAPORE • BOSTON • TAIPEI • CHENNAI

Preface

Joel L. Sussman* and Israel Silman†

The concept of structural genomics (SG) arose in the mid-to-late 1990s in both the USA and Japan, triggered by the success achieved in applying high-throughput (HTP) sequencing methods to whole genomes. It was envisaged that application of a similar HTP approach to obtaining the three-dimensional structures of a substantial fraction of the entire set of proteins of a given organism, the “proteome,” would be an efficient way of filling in the gaps in observed “fold-space.” The decision to adopt such an approach resulted in the investment of large sums of money, i.e. hundreds of millions of dollars, in large-scale structural genomics projects in both countries. Thus, in Japan, the Protein Research Group was established at the RIKEN Genomic Sciences Center in 1998, and in the USA, the Protein Structure Initiative (PSI), funded by the NIH/NIGMS, commenced at nine major centers in 2000. These projects were characterized by concentration of resources in a small number of large centers, by development of novel, automated technologies which would allow a HTP pipeline approach to structure determination, and a focus on novel folds as a major target criteria.

Europe was slower in implementing HTP approaches to structural biology. Various national efforts such as the Protein Structure Factory

Departments of *Structural Biology and †Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel

(PSF) in Berlin, the Oxford Protein Production Facility (OPPF), and the Genopoles in France, led the way. But it was only towards the end of 2002 that the first Europe-wide project began, an EC-funded Integrated Project entitled Structural Proteomics in Europe (SPINE). While benefiting from the technological achievements of the US and Japanese programs, and itself also concerned with developing cutting edge technologies as a means of achieving its objectives, SPINE from the outset focused these technologies on biomedically relevant targets. Indeed, as implied by its name, it aimed to establish a pan-European biomedically oriented structural **proteomics** program, placing significant emphasis on **functional aspects** of the target proteins studied. SPINE was followed by a series of EC-funded programs of various scopes and funding, some placing an emphasis on technological development, and others on attacking various classes of targets.

A similar emphasis on the use of the emerging HTP technologies to solve structures of biomedical relevance was adopted by the Structural Genomics Consortium (SGC), established in 2003 with the support of Canadian and British sponsors from both the public and private sectors, with laboratories in Oxford, Toronto and, subsequently, in Stockholm.

From the outset, the scale of funding of the PSI met with considerable criticism, especially in the USA, which has been going through a period during which funding for research by individual PIs has been hard to come by. Many critics have argued that the \$270M spent on funding the Pilot Phase of the PSI, over a period of five years, could have been more effectively utilized to fund hypothesis-driven research directed towards targets of fundamental or applied interest. Nevertheless, the achievements of the various SG/SP consortia, viewed in aggregate, and achieved in less than a decade, are impressive. In aggregate, the US PSI centers (September 2000 to June 2005) have determined over 1100 structures (<http://www.nigms.nih.gov/Initiatives/PSI/Background/PilotFacts.htm>). During PSI-2, which is still ongoing, ~1200 structures have already been solved, of which the vast majority share less than 30% sequence identity with any structure already deposited in the PDB, and many represent novel folds. As a result of the efforts of all the consortia (US, Japanese and

European), 5968 protein structures had been deposited in the Protein Data Bank as of 11-Dec-2007 (http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics). Although some of these structures may be redundant, or even appear uninteresting at first sight, many are of the highest technical quality, of fundamental and/or medical importance and, taken overall, provide a valuable database. Moreover, it has been reported that, in 2005, structures arising out of structural genomics and structural proteomics efforts accounted for 44% of the total number of novel structures reported. Although many of the novel structures solved by the SG/SP centers were, on the surface, low-hanging fruit, which filled gaps either in a given proteome or in fold space without yielding novel functional information, other targets have been, as already mentioned, of great fundamental and/or medical importance. Furthermore, filling in fold space provides a robust body of templates for homology modeling, which can rapidly take advantage of these templates as computational techniques become more sophisticated, and computing power increases. Thus, in our view, whatever policy decisions are taken with respect to funding of large-scale SG/SP projects, what has been achieved so far will have a lasting impact on biological and biomedical research. Indeed, the EC, through a Specific Support Action (SSA), established the Forum for European Structural Proteomics (**FESP**, see <http://www.ec-fesp.org>) to assess the current status and make recommendations for future European infrastructure requirements in the SG/SP area. We feel, therefore, that it is timely to publish a book in which these achievements are presented with a look to their potential impact on biological and biomedical research in general.

In this volume, we have tried to bring together experts capable of addressing all aspects of the SG/SP effort, from target selection, through the various techniques for expressing and purifying proteins and protein complexes and the methodologies for solving their structures, to their impact on drug design and on coping with emerging diseases. In view of the ongoing debate on SG/SP funding, we have also included a special chapter dealing with policy, which includes sections written by several scientists and officials who have been closely associated with the decision-making processes.

List of Contributors

Enrique E. Abola

Department of Molecular Biology
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla
California 92037
USA

Mariana Babor

Department of Plant Sciences
Weizmann Institute of Science
Rehovot 76100
Israel

Amos Bairoch

Swiss Institute of Bioinformatics
Centre Medical Universitaire
1, rue Michel-Servet
1211 Geneve 4
Switzerland

Edward N. Baker

School of Biological Sciences
University of Auckland
Private Bag 92019
Auckland
New Zealand

Lucia Banci

Centro Risonanze Magnetiche
University of Florence
Via Luigi Sacconi 6
Sesto Fiorentino
Florence 50019
Italy

Wolfgang Baumeister

Max Planck Institute of Biochemistry
Am Klopferspitz 18a
Martinsried D-82152
Germany

Florence Bedez

Institut de Génétique et de Biologie Moléculaire et Cellulaire
UMR 7104, 1 rue Laurent Fries
BP 10142, 67404 Illkirch Cedex
France

David M. Belnap

Department of Chemistry and Biochemistry
Brigham Young University
Provo, UT 84602
USA

Jeremy Berg

National Institute of General Medical Sciences
National Institutes Health
Bethesda, MD 20B92-6200
USA

Stephen Burley

SGX Pharmaceuticals Inc
10505 Roselle Street
San Diego, CA 92121
USA

Shuai Chen

Institute of Biochemistry
Center for Structural and Cell Biology in Medicine
University of Lübeck
Ratzeburger Allee 160
23538 Lübeck
Germany

A. Chesneau

EMBL Grenoble Outstation
6 rue Jules Horowitz
BP181, 38042 Grenoble Cedex 9
France

Mirosław Cygler

Department of Biochemistry
McGill University
845 Sherbrooke Street West
Montreal, Quebec H3A 2T5
Canada

Susan Daenke

Division of Structural Biology
Wellcome Trust Centre for Human Genetics
University of Oxford
Roosevelt Drive
Oxford, OX3 7BN
UK

Marvin Edelman

Department of Plant Sciences
Weizmann Institute of Science
Rehovot 76100
Israel

Josefina Enfedaque

European Commission
Research Directorate General
BE-1049 Brussels
Belgium

D. J. Hart

EMBL Grenoble Outstation
6 rue Jules Horowitz
BP181, 38042 Grenoble Cedex 9
France

Udo Heinemann

Max-Delbrück-Center for Molecular Medicine
Robert-Roessle-Str 10
Berlin D-13125
Germany

Rolf Hilgenfeld

Institute of Biochemistry
Center for Structural and Cell Biology in Medicine
University of Lübeck
Ratzeburger Allee 160
23538 Lübeck
Germany

Ursula Hinz

Swiss Institute of Bioinformatics
Centre Medical Universitaire
1, rue Michel-Servet
1211 Geneve 4
Switzerland

Zongchao Jia

Department of Biochemistry
Queen's University
Kingston, ON
K7L 3N6
Canada

Hualiang Jiang

Drug Discovery and Design Center
State Key Laboratory of Drug Research
Shanghai Institute of Materia Medica
Chinese Academy of Sciences
Zuchongzhi
Rd. 555, Shanghai 201203
China

Andrzej Joachimiak

Biosciences Division
Midwest Center for Structural Genomics and
Structural Biology Center
Argonne National Laboratory
9700 S Cass Ave. Argonne
IL 60439

E. Yvonne Jones

Division of Structural Biology
Wellcome Trust Centre for Human Genetics
University of Oxford
Roosevelt Drive
Oxford, OX3 7BN
UK

Christina Kiel

EMBL-CRG Systems Biology partnership Unit
Centre de Regulacio Genomica (CRG)
Dr Aiguader 88
08003 Barcelona
Spain

Saša Jenko Kokalj

Department of Biochemistry and Molecular Biology
Jozef Stefan Institute
Jamova 39
1000 Ljubljana
Solvenija

Seiki Kuramitsu

Department of Biological Sciences
Graduate School of Science
Osaka University
1-1 Machikaneyama-cho
Toyonaka, Osaka 560-0043
Japan

Arthur M. Lesk

Department of Biochemistry and Molecular Biology, and
the Huck Institute for Genomics, Proteomics and Bioinformatics
The Pennsylvania State University
University Park, PA 16802
USA

Ronen Levy

Department of Plant Sciences
Weizmann Institute of Science
Rehovot 76100
Israel

Allan Matte

Biotechnology Research Institute
6100 Royalmount Ave.
Montreal QC H4P 2R2
Canada

Dino Moras

Institut de Génétique et de Biologie Moléculaire et Cellulaire
UMR 7104, 1 rue Laurent Fries
BP 10142, 67404 Illkirch Cedex
France

John Moulton

Center for Advanced Research in Biotechnology
University of Maryland
Biotechnology Institute
9600 Gudelsky Drive
Rockville, MD 20850
USA

John Norvell

National Institute of General Medical Sciences
National Institutes Health
Bethesda, MD
USA

Yuan-Ping Pang

Computer-Aided Molecular Design Laboratory
Mayo Clinic
Rochester, Minnesota
USA

Helen Parkinson

European Bioinformatics Institute
Wellcome Trust Genome Campus Hinxton
Cambridgeshire CB10 1SD
United Kingdom

Arnaud Poterszman

Institut de Génétique et de Biologie Moléculaire et Cellulaire
UMR 7104, 1 rue Laurent Fries
BP 10142, 67404 Illkirch Cedex
France

Jacques Remacle

European Commission
BE-1049 Brussels
Belgium

Christophe Romier

IGBMC
1 rue Laurent Fries
B.P. 10142, 67404 Illkirch Cedex

Vineet Sangar

Department of Biochemistry and Molecular Biology, and
the Huck Institute for Genomics, Proteomics and Bioinformatics
The Pennsylvania State University
University Park, PA 16802
USA

Gunter Schneider

Karolinska Institutet
Scheelevägen 2
SE-171 77 Stockholm
Sweden

Luis Serrano

EMBL-CRG Systems Biology partnership Unit
Centre de Regulacio Genomica (CRG)
Dr Aiguader 88
08003 Barcelona
Spain

Xu Shen

Drug Discovery and Design Center
State Key Laboratory of Drug Research
Shanghai Institute of Materia Medica
Chinese Academy of Sciences
Zuchongzhi
Rd. 555, Shanghai 201203
China

Israel Silman

Department of Neurobiology
Weizmann Institute of Science
Rehovot 76100
Israel

Barbara Skene

Formerly Head of Department, Molecules, Genes and Cells
The Wellcome Trust
London NW1 2BE
UK

Vladimir Sobolev

Department of Plant Sciences
Weizmann Institute of Science
Rehovot 76100
Israel

Alasdair C. Steven

Laboratory of Structural Biology
National Institute of Arthritis, Musculoskeletal, and Skin Diseases
National Institutes of Health
Bethesda, MD 20892
USA

Raymond C. Stevens

Department of Molecular Biology
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla
California 92037
USA

David I. Stuart

Division of Structural Biology
Wellcome Trust Centre for Human Genetics
University of Oxford
Roosevelt Drive
Oxford, OX3 7BN
UK

Michael D. Suits

Department of Biochemistry
Queen's University
Kingston, ON
K7L 3N6
Canada

Joel L. Sussman

Department of Structural Biology
Weizmann Institute of Science
Rehovot 76100
Israel

Thomas Szyperski

816 Natural Sciences Complex
Chemistry Department
State University of New York at Buffalo
Buffalo, NY 14260
USA