Identification of Ligand Binding Site and Protein-Protein Interaction Area

FOCUS ON STRUCTURAL BIOLOGY

Volume 8

Series Editor

ROBERT KAPTEIN

Bijvoet Center for Biomolecular Research,
Utrecht University, The Netherlands

For further volumes:
http://www.springer.com/series/5990

Irena Roterman-Konieczna
Editor

Identification of Ligand Binding Site and Protein-Protein Interaction Area

Springer

Editor
Irena Roterman-Konieczna
Department of Bioinformatics and Telemedicine
Jagiellonian University
Medical College
Cracow, Poland

ISSN 1571-4853
ISBN 978-94-007-5284-9 ISBN 978-94-007-5285-6 (eBook)
DOI 10.1007/978-94-007-5285-6
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2012951181

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

The successful conclusion of the Human Genome Sequencing project, along with rapid progress in the development of analytical methods and high-performance computing solutions, has given rise to numerous biological databases of ever increasing volumes. Huge datasets, which nevertheless remain publicly accessible and affordable, are a crucial element of modern science. On the one hand, the ease with which research can be conducted is a great boon; on the other hand, however, one may feel somewhat overwhelmed by the immense quantity of available data. Such data is usually quite precise and detailed in nature, to the extent that modern scientific equipment and measuring devices allow. Information systems which assist in processing such data appear adequate, and their storage and processing capabilities – sufficient to meet the needs of modern researchers. Even so, further scientific breakthroughs are hindered by the relative lack of analysis methods targeted at large-scale datasets. This problem is particularly acute in analytical science, where it manifests itself as a general dearth of broad-scope methods with which to derive information (in the form of generalized models) approximating natural phenomena.

The above issue is the principal challenge in systems biology – a discipline which aims to develop comprehensive methods for simulating living organisms, so as to enable *in silico* experimentation on such organisms. A suitable system, properly reflecting the interactions and interdependencies observed in biological constructs, would support further research on specific anomalies, pathologies and diseases, well known to any clinician.

Before such a system can be designed and implemented, a fundamental biological axiom has to be addressed – namely, the relation between genetic information (genome) and the broad spectrum of active proteins, each of which facilitates a biological process, which, together, combine to form the extremely complex structure known as the organism.

Achieving this goal requires modeling three-dimensional structures of active proteins on the basis of their aminoacid sequences. The challenge lies not so much in predicting the structure itself, but rather in proposing a mechanism which leads to the generation of such structures. Another important issue, still waiting to

be addressed, is the challenge of determining the biological function of a given protein. We would expect numerical methods (capable of predicting ligand binding sites or catalytic centers, where reaction substrates are processed) to also suggest the means by which such "active" sites are generated.

This handbook presents a review of numerical techniques used to identify ligand binding and protein complexation sites. It should be noted that there are many other theoretical studies devoted to predicting the activity of specific proteins and that useful protein data can be found in numerous databases. The aim of advanced computational techniques is to identify the active sites in specific proteins and moreover to suggest a generalized mechanism by which such protein-ligand (or protein-protein) interaction can be effected.

The project EFI similar to CASP and CAPRI has been initiated in regard to enzymatic active site recognition (http://enzymefunction.org).

Developing such tools is not an easy task – it requires extensive expertise in the area of molecular biology as well as a firm grasp of numerical modeling methods. Thus, it is often viewed as a prime candidate for interdisciplinary research. Gatenby R.A. and Maini P.K. (2003) postulate the creation of an entirely new branch of science called "mathematical ontology" (see "Cancer summed up", *Nature*, 421, p. 321), which would bring together representatives of both – seemingly unconnected – disciplines. It is hoped that such close collaboration would lead to new systems enabling scientists to better simulate the properties and functioning of living organisms.

Cracow, 2012 Irena Roterman-Konieczna

Contents

Contributors

Paweł Alejster Department of Bioinformatics and Telemedicine, Jagiellonian University – Medical College, Cracow, Poland

Mateusz Banach Department of Bioinformatics and Telemedicine, Jagiellonian University – Medical College, Cracow, Poland

Adam S.Z. Belloum The Informatics Institue, University of Amsterdam, Amsterdam, The Netherlands

Emmmanuel Bettler Université Lyon 1, CNRS, UMR 5086; Bases Moléculaires et Structurales des Systèmes Infectieux, Lyon, France

Marian Bubak AGH University of Science and Technology Krakow, Poland and the Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

Julie-Anne Chemelle Université Lyon 1, CNRS, UMR 5086; Bases Moléculaires et Structurales des Systèmes Infectieux, Lyon, France

Christophe Combet Université Lyon 1, CNRS, UMR 5086; Bases Moléculaires et Structurales des Systèmes Infectieux, Lyon, France

Reginald Cushing The Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

Gilbert Deléage Université Lyon 1, CNRS, UMR 5086; Bases Moléculaires et Structurales des Systèmes Infectieux, Lyon, France

Christophe Geourjon Université Lyon 1, CNRS, UMR 5086; Bases Moléculaires et Structurales des Systèmes Infectieux, Lyon, France

Victor Guevara-Masis The Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

Bingding Huang Systems Biology Division, Zhejiang-California International NanoSystems Institute, Zhejiang University, Hangzhou, China

Bioinformatics Group, Biotechnology Center, Technical University of Dresden, Dresden, Germany

Joël Janin IBBMC, Université Paris-Sud, Orsay, France

Wiktor Jurkowski Computational Biology Group, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-Belval, Luxembourg

Leszek Konieczny Department of Bioinformatics and Telemedicine, Jagiellonian University – Medical College, Cracow, Poland

Vladimir Korkhov Faculty of Applied Math and Control Processes, St. Petersburg State University , Saint Petersburg , Russia

Spiros Koulouzis The Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

Damian Marchewka Department of Bioinformatics and Telemedicine, Jagiellonian University – Medical College, Cracow, Poland

Astronomy and Applied Computer Science, Jagiellonian University, Cracow, Poland

Irena Roterman-Konieczna Department of Bioinformatics and Telemedicine, Jagiellonian University – Medical College, Cracow, Poland

Raphaël Terreux Université Lyon 1, CNRS, UMR 5086; Bases Moléculaires et Structurales des Systèmes Infectieux, Lyon, France

Dmitry Vasunin Faculty of Applied Math and Control Processes, St. Petersburg State University , Saint Petersburg , Russia

Zhiming Zhao The Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands