Structure, Vol. 13, 339–341, March, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.str.2005.02.010

Macromolecular Assemblies Highlighted Editorial

The functional units in cells are often assemblies of macromolecules, including proteins and nucleic acids. Therefore, the knowledge of structure, dynamics, and energetics of macromolecular complexes is necessary for a mechanistic description of biochemical and cellular processes. This special issue of *Structure*, "Macromolecular Assemblies Highlighted," presents ten research articles and five reviews on the determination, representation, visualization, archival, and dissemination of assembly structure and dynamics.

Genome sequencing projects have provided nearly complete lists of macromolecules present in an organism. Following in the footsteps of genome sequencing, structural genomics was initiated to determine a sufficient number of protein structures by X-ray crystallography and NMR spectroscopy to enable structure characterization of most of the remaining protein sequence by computation. Through the NIH Protein Structure Initiative (http://www.nigms.nih.gov/psi/) and comparable efforts in Europe and Japan, the corresponding infrastructure that currently delivers hundreds of unique structures a year has been built. It is estimated that the ongoing scale-up of the structural genomics pipeline will allow us to substantially achieve the structural genomics goal within five to ten years.

Given the complete sets of protein sequences and their individual structures, the next challenge is to describe how the individual proteins interact in space and time. Macromolecular complexes vary widely in their activity and sizes, and play crucial roles in most cellular processes. They are often depicted as molecular machines, a metaphor that accurately captures many of their characteristic features, such as modularity, complexity, cyclic functions, and energy consumption.

Technical advances on several frontiers have expanded the applicability of existing methods to integrate structural information gathered at multiple levels of the biological hierarchy—from atoms to cells—into a common framework. The goal is a comprehensive description of the multitude of interactions between molecular entities, which in turn is a prerequisite for the discovery of general structural principles that underlie all cellular processes. This special issue of *Structure* focuses on the methods for characterization, analysis, visualization and annotation of the structures of complexes as well as their applications to specific cases.

X-ray crystallography continues to be a robust tool for determining the structures of individual proteins and small complexes. However, often the complex of interest cannot be crystallized. In such cases, it is increasingly possible to fit the atomic structures of the subunits into a medium-resolution map obtained by electron cryomicroscopy of the whole complex. These maps can be generated for complexes in different functional states, thus allowing a mechanistic description of the assembly activity. Moreover, the resulting models can give information about the molecular interactions at greater detail than the experimental resolution of the density map. Michael Chapman (Florida State University) reviews various approaches to building an atomic model of a complex from high-resolution structures of its components and an electron microscopy density map of the whole assembly, including his own. These methods are now being expanded by fitting flexible as well as rigid subunits.

Joachim Frank (Wadsworth Center) also describes a flexible fitting method to interpret medium-resolution electron-cryomicroscopy maps in terms of atomic subunit structures. The quality of the fitting is determined as a function of the number of rigid pieces into which the model is divided; suitable global and local quality-of-fit indicators include cross-correlation and R-factor. An application to the states of the ribosome involved in translocation illustrates how this method facilitates an atomic-resolution analysis of the functionally important conformational changes.

And finally, Michael G. Rossmann (Purdue University) reviews the combination of electron cryomicroscopy, which can describe large biological assemblies at low resolution, with crystallography, which can determine near atomic structures of assembly fragments. The integrated technique can be used to study dynamic processes involving large complexes and to help in selecting homogenous particles for subsequent averaging. Factors affecting the quality of electron-cryomicroscopy maps and limits of accuracy in fitting known structural fragments are discussed. The technique is illustrated by its application to a number of viruses.

Another exciting development in the structural characterization of assemblies is electron cryomicroscopy at subnanometer resolutions. Wah Chiu (Baylor College of Medicine) reviews advances that have enabled the structural determination of large biological machines, such as the rice dwarf virus and the acrosomal bundle, at subnanometer resolutions. Specifically, both computational and visualization methods for analysis of the electron-cryomicroscopy maps are described, including component segmentation, averaging based on local symmetry among components, density connectivity trace, bioinformatics analysis, and fitting of high-resolution component data. Such analyses can accurately identify the secondary structure segments of an assembly as well as suggest mechanistic models of its function.

Keiichi Namba (Osaka University) describes the structure of the bacterial R-type flagellar filament, obtained by electron cryomicroscopy at ~4 Å resolution. This filamentous helical propeller for bacterial locomotion consists of a highly ordered helical assembly of a single protein, flagellin. The density map shows the features of the alpha-helical backbone and some large side chains, thus allowing the authors to build a complete atomic model solely by electron-cryomicroscopy image analysis.

Ordinarily, structural biology provides static snapshots of a complex. Frequently, however, the components of a complex undergo conformational and configurational changes that are an essential part of the activity of the

complex. Computation began to play a key role in elucidating such assembly dynamics. Jianpeng Ma (Baylor College of Medicine) reviews normal mode analysis for describing large-amplitude molecular deformations that are ubiquitously involved in the functions of biological macromolecules based on low-resolution electron-cryomicroscopy maps. The review focuses on the strengths and weaknesses of the normal mode analysis by several well-established methods.

Ivet Bahar (University of Pittsburgh) applies the theory of elasticity to study large conformational changes involved in the maturation of the bacteriophage HK97 capsid. The most cooperative motions of the procapsid are found to be consistent with the observed change in conformation that takes place during maturation. Moreover, a few dominant modes of motion are sufficient to describe the anisotropic expansion that accompanies maturation. Based upon these modes, maturation is proposed to occur via an overall expansion and reconfiguration of the capsid initiated by puckering of the pentamers, followed by flattening and crosslinking of the hexameric subunits, and finally crosslinking of the pentameric subunits.

Whole cells are made up of thousands of macromolecular complexes, some transient and some stable. The technique that bridges the molecular and cellular levels of structure is electron cryotomography. Wolfgang Baumeister (Max Planck Institute for Biochemistry) obtained tomograms of isolated mammalian excitatory synapses in frozen-hydrated state by electron cryotomography. An automated procedure for the segmentation of complexes based on thresholding and connectivity is developed and applied to the synaptic adhesion complexes to calculate several of their morphological characteristics. It was found that the complexes are extensively connected along the synaptic cleft, forming a complex topology.

Structural characterization of macromolecular assemblies is usually difficult, justifying simultaneous consideration of all available information about a given assembly, irrespective of whether it comes from different experiments, theories, or statistical rules. Andrej Sali (University of California, San Francisco) suggests that complementing assembly shape determined by electron cryomicroscopy with information about subunit proximity determined by affinity purification provides a way to bridge the resolution gap between the assembly shape and the subunit configuration. To achieve this aim, structure characterization is expressed as a problem in satisfaction of spatial restraints that (1) represents subunits as spheres; (2) encodes information about the assembly shape, subunit excluded volume, and pulldowns in a scoring function; and (3) finds subunit configurations that satisfy the input restraints by an optimization of the scoring function. Testing of the approach with model systems suggests its feasibility.

Because of the complexity of multicomponent assemblies, it is challenging to convey a structure and its features. Several papers address assembly visualization and feature extraction. Michel Sanner (The Scripps Research Institute) describes several problems in the interactive visualization of large biological assemblies, including the development of multiresolution representations, new interaction methods for navigating and analyzing these complex systems, and flexible software environments that will facilitate the integration and interoperation of computational models and techniques. A number of software components based on the high-level, object-oriented, interpretive programming language Python are integrated and illustrated with the aid of several examples.

Chandrajit Bajaj (University of Texas at Austin) introduces a unified, compressed volumetric representation for macromolecular structures at varying feature resolutions. This compressed volumetric representation is derived using a custom-designed hierarchical wavelet basis scheme. Accuracy of molecular surfaces is retained even at very high-compression ratios, which allows for efficient exploratory visualization of the macromolecular structure and properties, particularly on desktop computers.

Thomas Ferrin (University of California, San Francisco) considers four software problems that arise in interactive visualization and analysis of large assemblies: how to represent multimers efficiently, how to make cartoon representations, how to calculate contacts efficiently, and how to select subassemblies. The methods presented here are proposed as features to add to existing visualization programs and to include in next-generation software for easy exploration of assemblies containing tens to thousands of macromolecules. These features have been implemented in the Multiscale extension of the UCSF Chimera molecular graphics program.

Arthur Olson (The Scripps Research Institute) describes a combination of computer autofabrication (i.e., 3D printing), which makes it possible to produce physical models for biological molecules and assemblies, and "augmented reality," which is a computer interface technology for perceived mixing of real-world objects and computer generated graphics. With the aid of the Python Molecular Viewer (PMV), the physical models provide a powerful, intuitive interface for manipulating the computer models, streamlining the interface between human intent, the physical model, and the computational activity.

David S. Goodsell (The Scripps Research Institute) reviews illustrations of molecular models for the study and dissemination of molecular structure and function. Several metaphors are commonly used to create these illustrations, each capturing a relevant aspect of the molecule and omitting other aspects. Standard representations are widely used for illustrating molecular structure, but methods for representing molecular properties are still an area of active research.

Finally, all structures need to be annotated, archived, and distributed to the biological community at large. Helen Berman (Rutgers University) reviews the composition of the growing number of large macromolecular complexes in the Protein Data Bank. Some of the challenges in representing, archiving, visualizing, and analyzing these structures are discussed along with possible means to overcome them.

In conclusion, structural biology is a great unifying discipline of biology. A comprehensive structural description of proteins and their assemblies may help us to discover the principles that underlie cellular pro-

cesses, bridging the gaps between genome sequencing, functional genomics, proteomics, and systems biology. Such structural characterization will require information from a number of experimental methods, physical theories, and biological databases. The goal seems daunting, but the consensus is that the prize will be commensurate with the effort invested, given the importance of molecular machines and functional networks in biology and medicine.

Acknowledgments

Based partly on: A. Sali, R. Glaeser, T. Earnest, W. Baumeister. (2003). From words to literature in structural proteomics. *Nature* **422**, 216-225. We also thank NIH National Center for Research Resources and National Science Foundation for their financial support.

Andrej Sali and Wah Chiu

Contributing Authors

Ivet Bahar
University of Pittsburgh
http://www.ccbb.pitt.edu/Research/Bahar_Lab/
Bacteriophage capsid dynamics (paper)

Chandrajit Bajaj
University of Texas at Austin
http://www.ices.utexas.edu/CCV
Volumetric Representations of Macromolecules (paper)

Wolfgang Baumeister
Max Planck Institute for Biochemistry
http://www.biochem.mpg.de/baumeister/
Synaptic cleft complexes by tomography (paper)

Helen M. Berman
Rutgers, The State University of New Jersey
http://rutchem.rutgers.edu/content_dynamic/faculty/helen_m_berman.shtml
Macromolecular complexes in the Protein Data Bank (review)

Michael S. Chapman
Florida State University
http://www.sb.fsu.edu/~chapman
Fitting subunit structures into density map (paper)

Wah Chiu
Baylor College of Medicine
http://ncmi.bcm.tmc.edu
Electron cryomicroscopy of biological machines (review)

Thomas E. Ferrin
University of California, San Francisco
http://www.cgl.ucsf.edu
Interactive visualization of complexes (paper)

Joachim Frank
Howard Hughes Medical Institute, Wadsworth Center
http://www.hhmi.org/research/investigators/frank_bio.html
Fitting subunit structures into density maps, applied to the ribosome (paper)

David S. Goodsell
The Scripps Research Institute
http://www.scripps.edu/mb/goodsell
Visualization methods from atoms to cells (mini-review)

Jianpeng Ma
Baylor College of Medicine
http://sigler.bioch.bcm.tmc.edu/MaLab
Normal mode analysis for dynamics of complexes (review)

Keiichi Namba
Osaka University
http://www.fbs.osaka-u.ac.jp/en/seminar/09a.html
An atomic structure of bacterial flagellar filament by electron cryomicroscopy (paper)

Arthur Olson
The Scripps Research Institute
http://www.scripps.edu/pub/olson-web
Combining 3D printing and augmented reality for complex visualization (paper)

Michael G. Rossmann
Purdue University
http://bilbo.bio.purdue.edu/~viruswww/Rossmann_home/
Combining X-ray crystallography and electron cryomicroscopy (review)

Andrej Sali
University of California, San Francisco
http://salilab.org/
Structure characterization based on assembly shape and subunit proximity (paper)

Michel F. Sanner
The Scripps Research Institute
http://www.scripps.edu/~sanner/software
Software for visualizing macromolecular assemblies (paper)

Also see supplementary movies of macromolecular assemblies, cryo-EM structures of macromolecular machines, density maps of a flagellar filament, and autofabrication methods and use of models in an augmented reality environment.