

# Structural biology of cellular machines

Wah Chiu<sup>1</sup>, Matthew L. Baker<sup>1</sup> and Steven C. Almo<sup>2</sup>

<sup>1</sup>National Center for Macromolecular Imaging and Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA

<sup>2</sup>Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY 10461, USA

**Multi-component macromolecular machines contribute to all essential biological processes, from cell motility and signal transduction to information storage and processing. Structural analysis of assemblies at atomic resolution is emerging as the field of structural cell biology. Several recent studies, including those focused on the ribosome, the acrosomal bundle and bacterial flagella, have demonstrated the ability of a hybrid approach that combines imaging, crystallography and computational tools to generate testable atomic models of fundamental biological machines. A complete understanding of cellular and systems biology will require the detailed structural understanding of hundreds of biological machines. The realization of this goal demands a concerted effort to develop and apply new strategies for the systematic identification, isolation, structural characterization and mechanistic analysis of multi-component assemblies at all resolution ranges. The establishment of a database describing the structural and dynamic properties of protein assemblies will provide novel opportunities to define the molecular and atomic mechanisms controlling overall cell physiology**

## Introduction

The enormously successful genome-sequencing projects continue to provide complete genetic descriptions of an ever-increasing number of model organisms [1–4]. At present, these efforts have yielded the raw genetic blueprints that govern all the biological processes within 150 organisms [5] and have given rise to the field of functional genomics, which aims to provide a functional characterization of each gene and its product(s). However, individual gene products rarely function independently, and it is often large multi-component protein assemblies that are the ultimate effectors of complex cellular functions. Accordingly, one of the outstanding challenges in the post-genomics era is the quantitative description of the regulatory circuits that control the spatial and temporal assembly, organization and function of these complex biological machines [6–8]. Here, we discuss current approaches to the structural characterization of large cellular machines and a perspective for the comprehensive and systematic investigation of these machines that dominate biological processes.

## Single particle machines

Based on the analysis of genome sequences, it has been suggested that life depends on 200–300 core biological processes [9]. Each of these processes involves multiple proteins, often organized into large heterogeneous assemblies, with a wide range of morphologies, functions and complexity. To elucidate the overall organization and atomic details, structural biologists have used many techniques, including X-ray crystallography (Box 1), NMR, cryo-electron microscopy (cryo-EM) (Box 2), mass spectrometry, fluorescence imaging and computational modeling, to complement biochemical and genetic information and reveal functional mechanisms for cellular assemblies.

In terms of composition, assemblies can be built from multiple copies of a single polypeptide chain such as the GroEL chaperone [10], the proteasome [11] and the calcium release channel [12]. Other assemblies are made up of numerous distinct polypeptides, such as the F1F0 ATPase [13], the Arp2/3 actin nucleation complex [14], RNA polymerase [15] and the nuclear pore [16].

The quintessential biological machine is the ribosome. The bacterial 70S ribosome (~2.5 MDa) is a machine consisting of three rRNA chains and >50 proteins. Both cryo-EM and X-ray crystallography have been used to attempt to solve the 50S ribosome structure independently. However, it was not until the low resolution cryo-EM density of the 50S ribosome was used during phasing to establish a molecular envelope that facilitated the solution of the first crystal structure [17]. Moreover, the ribosomal subunits have become a biological driving force for various technology developments for studying the structure of large cellular machines, such as new heavy atoms clusters for X-ray crystallography [18] and structural fitting for cryo-EM maps [19].

Cryo-EM and X-ray structures of the ribosome have provided important and complementary insights into the interactions of individual protein and RNA subunits, and the elucidation of the mechanisms of protein synthesis. X-ray structures enabled the folds and the spatial relationship of all the molecular components to be mapped (reviewed in Ref. [20]) and cryo-EM structures provided insight into the dynamic structural alterations required to support polypeptide initiation and elongation on the ribosome (reviewed in Ref. [21]). The cryo-EM structures of ribosome subunits bound with different ligands in different functional states revealed the binding sites and trajectories of tRNAs, mRNA and nascent polypeptides through the ribosome. Ratchet-like motions have been

Corresponding author: Chiu, W. (wah@bcm.edu).

Available online 3 February 2006

### Box 1. X-ray crystallography

X-ray crystallography [57–60] provides structural detail of macromolecules and their assemblies at or near atomic resolution. Purified proteins are examined under a wide range of solution conditions to generate crystals, which are then exposed to X-rays for the collection of X-ray diffraction data. These data are processed to yield 3D electron density maps into which the amino acid sequence of the molecule(s) of interest is fitted. This initial model is then refined so as to optimize the agreement between the model and the experimental diffraction data. Resolution is typically in the 2–3-Å range, in which not only all protein atoms (C,N,O,S), but also solvent molecules (e.g. water) and ions, which are crucial for structure and function, can be modeled. In cases with strong diffraction (i.e. ~1.5-Å resolution or better), true atomic resolution can be achieved. In exceptional cases (~1.0-Å resolution or better), the data enable the identification of deviations from ideal bonding geometries that are crucial for structure and function, as well as the modeling of hydrogen atoms, which provides the opportunity for a detailed analysis of hydrogen bond interactions.

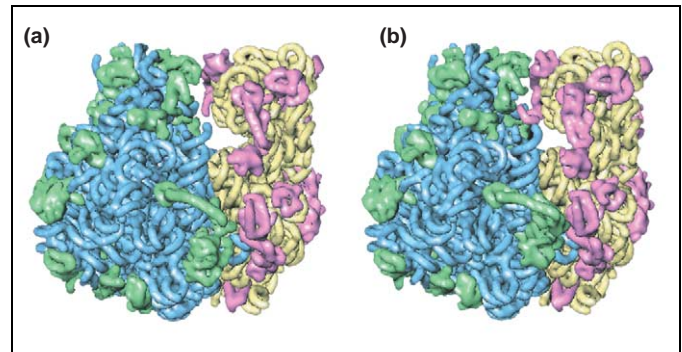
proposed to occur during the translation process [22] (Figure 1).

Although these discoveries have significantly enhanced our knowledge of ribosome structure and function, many key questions continue to challenge structural biologists. The cryo-EM structures elucidated so far have been at moderate resolutions; a higher-resolution cryo-EM structure would provide finer chemical details for each of the steps in protein translation. Additionally, high resolution structures of the ribosome in complex with other cellular factors are probable future targets. Work has already begun to elucidate the structural and functional roles of the elongation factors, producing models for the ribosomal

### Box 2. Cryo-electron microscopy (cryo-EM)

Cryo-EM has been used to study cellular assemblies in different forms, symmetries, sizes and shapes. A biochemically purified assembly is rapidly frozen and transferred to a standard transmission electron microscope equipped with a cryo-specimen holder to keep the specimen at a temperature near liquid nitrogen (-196°C) during microscopic observation [61]. Images are recorded under low dose without damaging the specimens. Generally, the images are noisy and require extensive image processing to combine images of specimens of different angles and to eliminate all the imaging artifacts [62]. To this end, several specialized software packages are available to handle different imaging techniques and particle symmetries [63].

Although the resolution of the 3D structure is key in determining the necessary number of 2D images, the inherent symmetry of the assembly also affects the number of images required for a 3D reconstruction. Assemblies with high symmetry (e.g. viruses) might only require a few thousand particle images, whereas low-symmetry objects (e.g. the ribosome) require one to two orders of magnitude more to reach subnanometer resolutions. Furthermore, the final 3D maps could have different resolutions depending on the homogeneity of the specimen, the quality of the data and the methods used to perform the reconstruction. For specimens that can form 2D crystal or helical array, the structure can be retrieved to 3–4-Å resolution, from which atomic models of the components can be derived [64,65]. For specimens existing as a single particle, tens of thousands of the particle images can be combined coherently to yield structures up to 6–9-Å resolution, where long  $\alpha$  helices and  $\beta$  sheets can frequently be recognized [66,67]. Additionally, by imaging multiple functional states and comparing the resulting structures, cryo-EM can uncover the structural dynamics of the large assembly [22,68,69].



**Figure 1.** Ribosome in different functional states. The structure of the 70S ribosome is shown in two functional states: initiation-like state (a) and an elongation state with EF-G-GTP bound (b). In both images, 50S RNA is shown in blue, 30S RNA is shown in yellow, 50S protein is shown in green and 30S protein is shown in purple. (Image courtesy of J. Frank at HHMI, Albany.)

activities that trigger the GTPase activity of the elongation factors (reviewed in Ref. [23]). Furthermore, given that the bacterial ribosome is a natural target for drug development, X-ray crystallography is ideally suited to illustrate the atomic details of ribosome–antibiotic complexes, which could in turn facilitate structure-based design of antibiotics for bacterial infection (reviewed in Ref. [24]). The most exciting opportunity in cell structural biology lies in the visualization of functional ribosomes within the cell. Cryo-electron tomography is well suited for this task, and could result in a better understanding of the function and localization of ribosomes within a cell under various physiologically relevant conditions [25].

### Filamentous machines

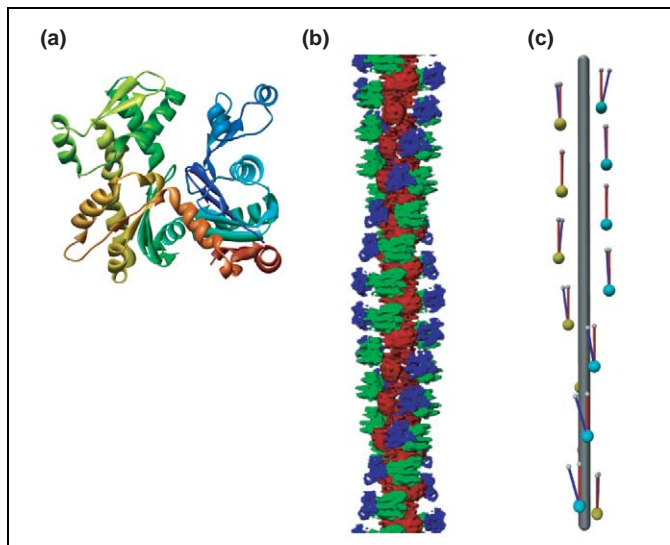
A unique family of biological machines relies on the ability of the individual subunits to form linear or helical arrangements that are required for function. These types of assembly can function as simple scaffolds to support higher-order assemblies or use the intrinsic properties of the ordered array to generate biological function. For example, the  $\alpha$ - $\beta$ -tubulin dimer forms a regular tubular array that supports basic cell morphology and serves as the track for numerous molecular motors [26]. Importantly, the formation of ordered arrays might also be associated with ‘loss of function’ or pathology, such as the hemoglobin fibers associated with sickle cell disease or the amyloid fibers that are a hallmark of Alzheimer’s disease [27].

As these types of complex are much larger than single molecules, the approaches taken to study them must either overcome the complexity of the assembly or take advantage of the inherent symmetry of the complex. Because of their structural organization, it is rarely possible to study them in their native filamentous or tubular configuration using X-ray crystallography. In cryo-EM, structure determination takes advantage of symmetry, reducing the total number of images required as symmetry increases. The divide-and-conquer approach, a process of solving individual components or domains of larger structures using crystallography or NMR and entire intact assembly by cryo-EM, has been an effective strategy. In the case of tubulin, the atomic model was derived from 2D electron crystallography [28]. When

combined with the cryo-EM structure of the intact microtubule, a clearer picture of native microtubules, their assembly, interactions and biological properties was realized [29].

One of the best examples of linearly ordered arrays is actin, which has provided insight into the fundamental structural and functional properties associated with repetitive polymeric structures. The high-resolution crystal structures of monomeric actin (Figure 2a) have been determined in different ligand-bound forms and from different species [272 actin and actin-binding proteins currently in the Protein Data Bank (PDB)]. The inherent flexibility of filamentous actin (F-actin) [30], and shortcomings in current image processing techniques to handle this elasticity, combined with the difficulty in purifying a structurally homogenous actin filament in its native biological context, has hindered the detailed structural description of F-actin, although models have been proposed [31,32].

A subnanometer resolution cryo-EM structure for the acrosomal bundle composed of actin and scruin from *Limulus polyphemus* sperm was solved recently [33]. In this case, the actin filament in the bundle becomes rigid by binding scruin at a 1:1 ratio, which forms a quasi-crystalline bundle  $\sim 1000$  Å in width and tens of microns in length. This preparation represented a unique and highly ordered filament that is biologically active. The structure revealed for the first time that the actin filament in a biological context does not conform to a simple helical description as in standard F-actin models (Figure 2b and c). As such, these samples presented novel challenges and required the development of new approaches for data



**Figure 2.** Actin. (a) Crystal structure of the actin monomer colored from N (blue) to C (red). (b) Cryo-EM structure of the asymmetric unit of the actin-scrutin filament (14 pairs of subunits) computationally segmented from the 9.5-Å map of the quasi-crystalline acrosomal bundle. Each of the actin subunits is well resolved with distinct spatial orientations along the filament within the asymmetric unit of the bundle. (c) Relative orientations of actin from the acrosomal bundle (blue vector) are shown in comparison to the Holmes F-actin model (red vector). Each strand of actin is indicated with either a cyan or yellow sphere. This type of structural variation at the 'quaternary structure' level is likely to occur in actin filaments inside the cell, where the filament would interact with many cellular factors to affect its nucleation and disassembly. (Image courtesy of M. F. Schmid at Baylor College of Medicine.)

processing and subsequent computational analysis to yield the final and biologically meaningful structure [33]. Individual actin monomers adopt distinct orientations with respect to the filament axis (Figure 2c). This lack of strict helical symmetry persists throughout the genuine biological bundle in a periodic fashion and results in the formation of a quasi-crystalline bundle. The structure shows that a single two-domain scruin molecule wraps around two consecutive actin molecules along the filament and interacts with scruin molecules in neighboring filaments. This structural organization provides a mechanistic rationale for the force generation associated with the spring-like transformation of the acrosomal bundle from a coiled to an extended form during the fertilization process in *L. polyphemus* [34]. Additionally, this structure directly illustrates the plasticity of individual actin monomers within a filament and highlights the potential contribution of this inherent structural property to specific biological functionality.

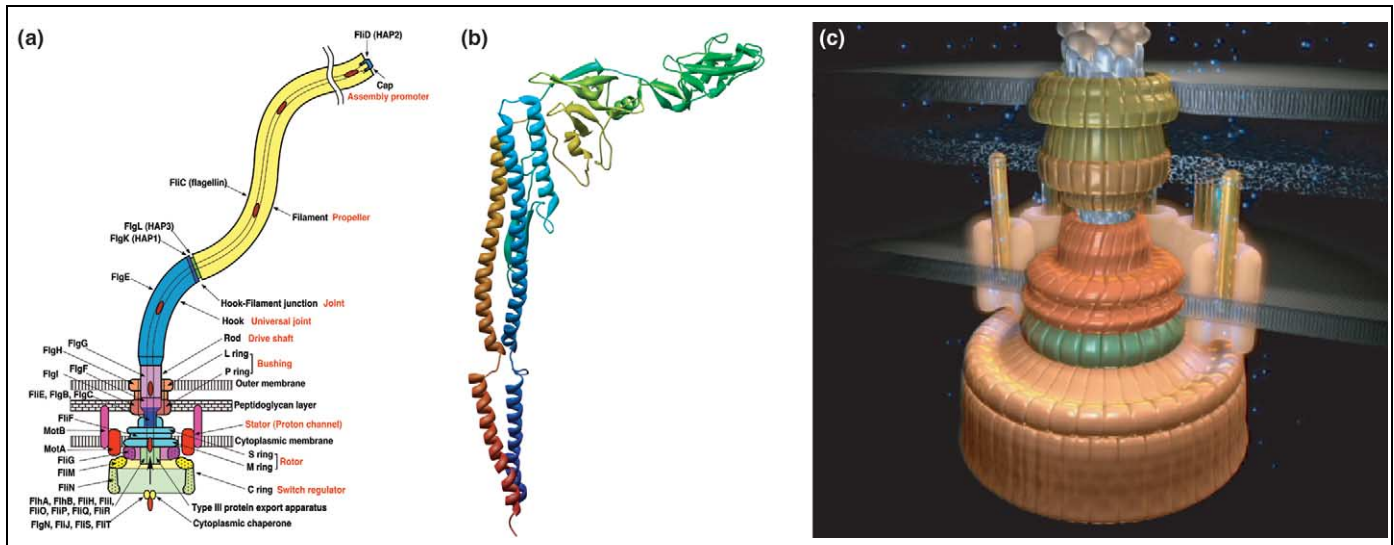
### Assemblies of assemblies

Some of the cellular machines are assemblies of assemblies. Typifying this type of large multi-tiered machinery is the flagella (Figure 3a), which is the organelle used for propulsion in many single-celled organisms. In *Escherichia coli*, 42 distinct protein types comprise the 15- $\mu$ m assembly. In general, flagellin forms a hollow helical tube that functions as the propeller, which is coupled to the motor through a shaft and a series of rings. A 'hook' attaches the flagellin propeller to the outer cell membrane, forcing the propeller to point away from the cell. The rotary engine that drives the flagella movement is located on the inner cell membrane and is powered by a proton motive force, ATP hydrolysis and  $\text{Na}^+$  ion gradients, reaching 200–1000 rpm with an attached flagella filament [35].

The size and complexity of flagella, and their ability for self assembly, have made structural studies challenging. Crystallographers have determined the high-resolution structure of flagellin by truncating almost 100 amino acids from the termini ( $\sim 53$  from the amino end and 43 from the carboxyl end) of flagellin [36] (Figure 3b). Additionally, a 4-Å cryo-EM map of the intact R-type straight bacterial flagella filament defined the organization of all protein subunits within the flagella filament [37]. Furthermore, cryo-EM has yielded low resolution (22-Å and 25-Å resolution) structures of the rotor of the bacterial flagellar motor [38] (Figure 3c). From such studies, it is now possible to model the entire organelle and design testable hypotheses to understand the molecular and atomic mechanisms that drive cellular movement in bacterial taxis. Most importantly, this combined structural approach, which spans all resolution ranges, provides a powerful paradigm for dissecting the structural, dynamic and mechanistic features of all biological machines.

### Future cellular machines for structural studies

Although the ribosome, acrosomal bundle and bacteria flagella represent remarkable achievements in structural and cellular biology, numerous challenges remain. In addition to a handful of obvious targets, for example, the



**Figure 3.** Flagella. A schematic model of the multi-component bacteria flagella is shown in (a). The structure of flagellin, the primary component of the bacteria flagella, is shown in (b). An image derived from cryo-EM of the bacteria flagella motor is shown in (c). From top to bottom in (c), the grey represents the propeller, the driving shaft is cyan, the bushing is yellow and dark orange, the stator is peach with an orange cylinder, the rotor is dark orange and green, and the switch regulator is in pale orange. (Image courtesy of K. Namba at Osaka University, Osaka.)

nuclear pore complex [39], the mitotic spindle apparatus [40,41], the mammalian axoneme [42] and the protein-folding machinery with its substrate [43], it is difficult to predict what might be the next great focus of structural biology, as there will be an increasing number of such machinery complexes identified as proteomic research progresses. However, it is inevitable that larger and more complex assemblies will be at the forefront of structural biology. Entire pathways, including both stable and transient complexes, will become targets for structure determination. As these pathways grow more complete and intricate, electron cryo-tomography could be used to image larger cellular complexes and processes in organelles or even in whole cells. This approach could result in the imaging of biological processes at resolution ranges that span from single molecules to entire cells, and would be fully complementary to emerging large-scale light microscopy initiatives.

New developments in heterologous expression systems offer promise in achieving the relatively large-scale expression and purification of multi-component assemblies. For example, Tan and co-workers have developed an efficient strategy for the generation of multicistronic vectors for expression in *E. coli* [44]. This strategy has recently been used for the production and purification of biochemically active DASH complex, a ten-component assembly that is involved in the binding of microtubules to kinetochores in yeast [45,46]. An efficient strategy for the generation of recombinant baculovirus has been recently reported that supports the expression of multi-component assemblies in a eukaryotic system [47]. These approaches are likely to be of continued importance, especially as complexes with relatively low natural abundance are targeted.

Additionally, recent reports of the efficient large-scale isolation of numerous diverse complexes from *Saccharomyces cerevisiae* forecast the systematic structural analysis of the entire ensemble of multi-component

protein assemblies that are responsible for complex biological function [48,49]. Large numbers of machinery complexes from *S. cerevisiae* can be purified by a tandem affinity purification approach (TAP-Tag) and characterized in terms of individual components [50]. These studies use homologous recombination to generate fusion proteins carrying affinity tags (targets) that can then be introduced into the cognate host cell or organism. This enables efficient and rapid purification of the entire ensemble of proteins (i.e. complexes) associated with the tagged molecule, potentially providing material for initial cryo-EM and X-ray structural determinations. Combined with mass spectrometry and functional assays, affinity purification methods can provide a rapid means to analyze the composition and possibly the mechanistic features of many assemblies. The placement of affinity tags might affect protein stability/folding, assembly and expression and thus different strategies could be employed in tag placement. Additionally, it is preferable to maintain expression of the construct at or close to normal levels. A particular advantage of this approach in *S. cerevisiae* is that expression of the tagged protein is driven from the normal promoter, thus avoiding complications associated with overexpression. At present, only approximately one third of the *S. cerevisiae* genome has been reported, suggesting that a large number of additional complexes remain to be defined for structural analysis [51]. Other genetically tractable organisms, such as *Schizosaccharomyces pombe* and *Dictyostelium*, could enable the characterization of homologous or unique assemblies.

The ability to examine related assemblies from multiple organisms provides multiple specimens with distinct biochemical and biophysical properties for crystallization and cryo-EM, thus maximizing the likelihood of any individual structure determination. Furthermore, the comparison of complexes from distinct organisms will enable direct functional and evolutionary comparisons. Beyond eukaryotic model systems, the systematic

examination of protein complexes from prokaryotic and viral sources will provide novel opportunities. For example, the examination of complexes from related clinically important and non-infectious laboratory model organisms (i.e. *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*) could provide information on the organization, assembly and function of protein complexes, in both pathogen–pathogen and pathogen–host interactions, that are directly responsible for human disease. Furthermore, the analysis of assemblies from the simplest organisms (e.g. *Mycoplasma genitalium*) will enable the identification and characterization of those that are fundamental or essential for life [52]. Describing the assemblies involving human pathogens will be crucial for drug and vaccine design against various infectious agents responsible for cold sores to bioterrorism threats. Ultimately, protein assemblies derived from diverse human-cell types, both normal and pathological, must be examined. These studies could directly identify cell-type-specific differences (e.g. liver versus kidney) and provide clues relevant to the complex regulatory mechanisms that are involved in normal physiology and a wide range of pathologies, including cancers and inherited genetic diseases, as well as the cellular and molecular responses to various types of therapeutic intervention.

Beyond naturally occurring complexes, synthetic assemblies (i.e. nano-machines) will benefit from the structural tools used in studying natural complexes. Recent work has already demonstrated the feasibility of the construction of DNA aptamers as potential nano-delivery devices [53]. Similarly, synthetic viral constructs are also amenable to structural studies [54,55]. As such, structural biology has extended its reach beyond traditional structural and functional studies into nano-medicine and protein engineering.

### Infrastructure to study the structures and functions of large machines

Given the central roles of cellular machines in biology and nano-medicine, how should structural biology be used and further developed to address these complex biological entities? There is a need to develop collaborative and individual research with diverse expertise that would encompass the following activities in the most cost-effective manner: (i) proteomics and bioinformatics capabilities to identify and characterize the individual molecular components that contribute to specific transient and stable assemblies; (ii) molecular biology and protein purification facilities for the generation of assemblies in quantities sufficient to support biophysical analysis; (iii) advanced cryo-EM software for the automation of data collection and structure solution at the highest-achievable resolution with partially pure samples; (iv) efficient approaches for the crystallization and structure solution of complexes, which might be available in relatively small amounts; (v) the development of novel computational algorithm of flexible docking of crystal structures of individual domains or components within medium-resolution cryo-EM maps (Box 3); (vi) complementary informatics and experimental approaches to validate proposed structural models of protein assemblies;

### Box 3. Hybrid methods in structural and computational biology

Both cryo-EM and crystallography are well suited for studying large assemblies, although both approaches have intrinsic limitations. X-ray crystallography is capable of producing high resolution structures of an entire assembly or components in a static state. However, it is often difficult to capture the dynamic structure of the assembly in different functional states using X-ray crystallography. Conversely, cryo-EM can capture entire complex assemblies in different functional states, although only at low–medium resolutions [68]. In combining these two disciplines, computational approaches have enabled a more in-depth analysis of large macromolecular machine [70,71]. When low–medium resolution cryo-EM structures are available, several computational approaches have been developed to rigidly [72–75] or flexibly [76] fit high resolution structures of individual polypeptides and domains derived from NMR and crystallographic analyses. These approaches enable the placement of individual components within the cryo-EM density to yield an atomic resolution model for the entire multi-component assembly that can be experimentally validated. Additional computational tools have also enabled the identification of secondary structure elements [75,77] and nucleic acid in medium-resolution cryo-EM structures [78]. In general, these types of tool use feature recognition algorithms to analyze and ‘mine’ small structural features from the larger complex. Such tools have provided a method for building pseudo-atomic models for the components or entire machine [79]. Beyond direct structural analysis, there are efforts to combine the sequence-based structure prediction with the medium resolution cryo-EM structures to build a pseudo-atomic model of these complexes [80,81]. These hybrid approaches will have a central role in understanding the functions of complex cellular machinery.

(vii) complementary informatics and experimental approaches to examine and reveal the dynamic features that are central to the biological function of protein assemblies; (viii) the integration of biochemical, genetic and cell biological information that spans all temporal and spatial resolution ranges to complement the high-resolution structural studies; and (ix) a robust and easily interrogated database that will support individual investigators in the analysis of macromolecular assemblies that define the functions and mechanisms relevant to normal physiology and human pathology.

There is debate among biologists over the most cost-effective means to accomplish these goals. In view of the complexities described, a well coordinated and systematic effort by a large and highly interdisciplinary team of investigators is required. Significant progress has already been made towards the establishment of this type of infrastructure. The 3D-EM Network (<http://www.3dem-noe.org/>) and Spine initiatives (<http://www.spineurope.org>) in Europe are examples of forums for various structural biology fields to coordinate research and training, as well as to foster research and industry collaborations in structural biology. Additionally, 3D-EM Repertoire (<http://www.3drepertoire.org>) focuses on multidisciplinary approaches to studying large biological complexes.

In the USA, the Department of Energy has initiated a large-scale research program to identify and characterize biological complexes in microbial organisms relevant to alternative energy production and bioremediation (<http://www.doenomestolife.org/research/index.shtml>). The

National Institutes of Health has also established the Protein Structure Initiative (PSI; <http://www.nigms.nih.gov/Initiatives/PSI>), which currently supports the activities of several structural genomics centers to establish an efficient high-throughput structure discovery pipeline with a long-term goal of describing the 2000–3000 unique folds and producing meaningful structural models for all individual globular protein domains. However, there is no specific plan for coordinating these efforts, although significant interest exists within the community [7]. Regardless of the means of support, these efforts will continue to evolve in the next decades when cellular machines will be the prime targets for elucidating not only structures, but also functions in cellular environments.

## Conclusion

There is now the opportunity and the need to expand existing efforts on the structural characterization of complex biological machines. The required informatics, computational and experimental tools are available and being further developed for the identification, isolation and structure solution of assemblies at multiple-resolution ranges. Worldwide structural genomics efforts, including the PSI [56], are rapidly expanding the available structural database of individual proteins and domains. The time is ripe for these activities to support the development of high-resolution structural studies of a wide range of cellular machines. Most important is the great increase in the amount of biological data describing strong and transient macromolecular interactions and their regulation and integration into the overall physiology of the cell. It is clear that the most exciting times for structural biology lie ahead.

## Acknowledgements

We thank Joachim Frank, Michael F. Schmid and Keichii Namba for providing figures 1–3 respectively. We acknowledge the support of grants from NIH (P20RR020647, P41RR02250, PN2EY016525 and U54GM074945), NSF (EIA-0325004) and the Keck Center for Computational and Structural Biology of the Gulf Coast Consortia (NIH 5T15LM07093 to M.L.B.).

## References

- Olivier, M. *et al.* (2001) A high-resolution radiation hybrid map of the human genome draft sequence. *Science* 291, 1298–1302
- Celniker, S.E. and Rubin, G.M. (2003) The *Drosophila melanogaster* genome. *Annu. Rev. Genomics Hum. Genet.* 4, 89–117
- Venter, J.C. *et al.* (2001) The sequence of the human genome. *Science* 291, 1304–1351
- Yu, J. *et al.* (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296, 79–92
- Hallin, P.F. and Ussery, D.W. (2004) CBS Genome Atlas Database: a dynamic storage for bioinformatic results and sequence data. *Bioinformatics* 20, 3682–3686
- Levchenko, A. (2001) Computational cell biology in the post-genomic era. *Mol. Biol. Rep.* 28, 83–89
- Sali, A. (2003) NIH workshop on structural proteomics of biological complexes. *Structure* 11, 1043–1047
- Alberts, B. and Miske-Lye, R. (1992) Unscrambling the puzzle of biological machines: the importance of the details. *Cell* 68, 415–420
- Martin, A.C. and Drubin, D.G. (2003) Impact of genome-wide functional analyses on cell biology research. *Curr. Opin. Cell Biol.* 15, 6–13
- Braig, K. *et al.* (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371, 578–586
- Lowe, J. *et al.* (1995) Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science* 268, 533–539
- Ludtke, S.J. *et al.* (2005) The pore structure of the closed RyR1 channel. *Structure* 13, 1203–1211
- Abrahams, J.P. *et al.* (1996) The structure of bovine F1-ATPase complexed with the peptide antibiotic efrapeptin. *Proc. Natl. Acad. Sci. U. S. A.* 93, 9420–9424
- Robinson, R.C. *et al.* (2001) Crystal structure of Arp2/3 complex. *Science* 294, 1679–1684
- Cramer, P. (2004) RNA polymerase II structure: from core to functional complexes. *Curr. Opin. Genet. Dev.* 14, 218–226
- Wente, S.R. (2000) Gatekeepers of the nucleus. *Science* 288, 1374–1377
- Ban, N. *et al.* (1998) A 9 Å resolution X-ray crystallographic map of the large ribosomal subunit. *Cell* 93, 1105–1115
- Clemons, W.M., Jr. *et al.* (2001) Crystal structure of the 30 S ribosomal subunit from *Thermus thermophilus*: purification, crystallization and structure determination. *J. Mol. Biol.* 310, 827–843
- Tama, F. *et al.* (2003) Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9319–9323
- Moore, P.B. and Steitz, T.A. (2003) The structural basis of large ribosomal subunit function. *Annu. Rev. Biochem.* 72, 813–850
- Frank, J. (2003) Toward an understanding of the structural basis of translation. *Genome Biol.* 4, 237
- Gao, H. *et al.* (2003) Study of the structural dynamics of the *E. coli* 70S ribosome using real-space refinement. *Cell* 113, 789–801
- Nilsson, J. and Nissen, P. (2005) Elongation factors on the ribosome. *Curr. Opin. Struct. Biol.* 15, 349–354
- Hermann, T. (2005) Drugs targeting the ribosome. *Curr. Opin. Struct. Biol.* 15, 355–366
- Medalia, O. *et al.* (2002) Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. *Science* 298, 1209–1213
- Kikkawa, M. *et al.* (1994) Direct visualization of the microtubule lattice seam both *in vitro* and *in vivo*. *J. Cell Biol.* 127, 1965–1971
- Turner, M.S. *et al.* (2003) Twisted protein aggregates and disease: the stability of sickle hemoglobin fibers. *Phys. Rev. Lett.* 90, 128103
- Nogales, E. *et al.* (1998) Structure of the  $\alpha$   $\beta$  tubulin dimer by electron crystallography. *Nature* 391, 199–203
- Wang, H.W. and Nogales, E. (2005) Nucleotide-dependent bending flexibility of tubulin regulates microtubule assembly. *Nature* 435, 911–915
- Egleman, E.H. and Orlova, A. (1995) Allostery, cooperativity, and different structural states in F-actin. *J. Struct. Biol.* 115, 159–162
- Holmes, K.C. *et al.* (1990) Atomic model of the actin filament. *Nature* 347, 44–49
- Orlova, A. *et al.* (2001) Probing the structure of F-actin: cross-links constrain atomic models and modify actin dynamics. *J. Mol. Biol.* 312, 95–106
- Schmid, M.F. *et al.* (2004) Structure of the acrosomal bundle. *Nature* 431, 104–107
- Mahadevan, L. and Matsudaira, P. (2000) Motility powered by supramolecular springs and ratchets. *Science* 288, 95–100
- Ryu, W.S. *et al.* (2000) Torque-generating units of the flagellar motor of *Escherichia coli* have a high duty ratio. *Nature* 403, 444–447
- Samatey, F.A. *et al.* (2001) Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. *Nature* 410, 331–337
- Yonekura, K. *et al.* (2003) Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* 424, 643–650
- Suzuki, H. *et al.* (2004) Structure of the rotor of the bacterial flagellar motor revealed by electron cryomicroscopy and single-particle image analysis. *J. Mol. Biol.* 337, 105–113
- Beck, M. *et al.* (2004) Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science* 306, 1387–1390
- Bloom, K. (2005) Chromosome segregation: seeing is believing. *Curr. Biol.* 15, R500–R503
- O'Toole, E.T. *et al.* (1999) High-voltage electron tomography of spindle pole bodies and early mitotic spindles in the yeast *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 10, 2017–2031
- McEwen, B.F. *et al.* (2002) Use of frozen-hydrated axonemes to assess imaging parameters and resolution limits in cryoelectron tomography. *J. Struct. Biol.* 138, 47–57

- 43 Saibil, H.R. *et al.* (2001) Allostery and protein substrate conformational change during GroEL/GroES-mediated protein folding. *Adv. Protein Chem.* 59, 45–72
- 44 Tan, S. *et al.* (2005) The pST44 polycistronic expression system for producing protein complexes in *Escherichia coli*. *Protein Expr. Purif.* 40, 385–395
- 45 Miranda, J.J. *et al.* (2005) The yeast DASH complex forms closed rings on microtubules. *Nat. Struct. Mol. Biol.* 12, 138–143
- 46 Westermann, S. *et al.* (2005) Formation of a dynamic kinetochore-microtubule interface through assembly of the Dam1 ring complex. *Mol. Cell* 17, 277–290
- 47 Berger, I. *et al.* (2004) Baculovirus expression system for heterologous multiprotein complexes. *Nat. Biotechnol.* 22, 1583–1587
- 48 Sali, A. *et al.* (2003) From words to literature in structural proteomics. *Nature* 422, 216–225
- 49 Russell, R.B. *et al.* (2004) A structural perspective on protein-protein interactions. *Curr. Opin. Struct. Biol.* 14, 313–324
- 50 Rigaut, G. *et al.* (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17, 1030–1032
- 51 Aloy, P. *et al.* (2004) Structure-based assembly of protein complexes in yeast. *Science* 303, 2026–2029
- 52 Adams, M. *et al.* (2004) Meeting review: 2003 NIH protein structure initiative workshop in protein production and crystallization for structural and functional genomics. *J. Struct. Funct. Genomics* 5, 1–2
- 53 Shih, W.M. *et al.* (2004) A 1.7-kilobase single-stranded DNA that folds into a nanoscale octahedron. *Nature* 427, 618–621
- 54 Chatterji, A. *et al.* (2004) New addresses on an addressable virus nanoblock; uniquely reactive Lys residues on cowpea mosaic virus. *Chem. Biol.* 11, 855–863
- 55 Nemerow, G.R. and Stewart, P.L. (1999) Role of  $\alpha(v)$  integrins in adenovirus cell entry and gene delivery. *Microbiol. Mol. Biol. Rev.* 63, 725–734
- 56 Norvell, J. and Berg, J.M. (2005) The protein structure initiative, five years later. *Scientist* 19, 30
- 57 Hammond, C. (2001) *The Basics of Crystallography and Diffraction*, Oxford University Press
- 58 Ladd, M.F.C. and Palmer, R.A. (2003) *Structure Determination by X-ray Crystallography*, Kluwer Academic/Plenum Publishers
- 59 Massa, W. (2004) *Crystal Structure Determination*, Springer
- 60 McPherson, A. (2003) *Introduction to Macromolecular Crystallography*, Wiley-Liss
- 61 Lepault, J. *et al.* (1983) Electron microscopy of frozen biological suspensions. *J. Microsc.* 129, 89–102
- 62 Frank, J. *et al.* (2002) Three-dimensional imaging of biological complexity. *J. Struct. Biol.* 138, 85–91
- 63 Carragher, B. and Penczek, P.A. (2003) Analytical methods and software tools for molecular microscopy. *J. Struct. Biol.* 144, 1–3
- 64 Grigorieff, N. *et al.* (1996) Electron crystallographic refinement of the structure of bacteriorhodopsin. *J. Mol. Biol.* 259, 393–421
- 65 Miyazawa, A. *et al.* (2003) Structure and gating mechanism of the acetylcholine receptor pore. *Nature* 423, 949–955
- 66 Ludtke, S.J. *et al.* (2004) Seeing GroEL at 6 Å resolution by single particle electron cryomicroscopy. *Structure (Camb)* 12, 1129–1136
- 67 Chiu, W. *et al.* (2005) Electron cryomicroscopy of biological machines at subnanometer resolution. *Structure (Camb)* 13, 363–372
- 68 Subramaniam, S. and Henderson, R. (2000) Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature* 406, 653–657
- 69 Saibil, H.R. (2000) Conformational changes studied by cryo-electron microscopy. *Nat. Struct. Biol.* 7, 711–714
- 70 Fotin, A. *et al.* (2004) Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* 432, 573–579
- 71 Orlova, E.V. *et al.* (1996) Two structural configurations of the skeletal muscle calcium release channel. *Nat. Struct. Biol.* 3, 547–552
- 72 Wriggers, W. and Birmanns, S. (2001) Using situs for flexible and rigid-body fitting of multiresolution single-molecule data. *J. Struct. Biol.* 133, 193–202
- 73 Volkman, N. and Hanein, D. (1999) Quantitative fitting of atomic models into observed densities derived by electron microscopy. *J. Struct. Biol.* 125, 176–184
- 74 Rossmann, M.G. *et al.* (2001) Combining electron microscopic with x-ray crystallographic structures. *J. Struct. Biol.* 136, 190–200
- 75 Jiang, W. *et al.* (2001) Bridging the information gap: computational tools for intermediate resolution structure interpretation. *J. Mol. Biol.* 308, 1033–1044
- 76 Tama, F. *et al.* (2004) Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis. *J. Mol. Biol.* 337, 985–999
- 77 Kong, Y. and Ma, J. (2003) A structural-informatics approach for mining  $\beta$ -sheets: locating sheets in intermediate-resolution density maps. *J. Mol. Biol.* 332, 399–413
- 78 Spahn, C.M. *et al.* (2000) A method for differentiating proteins from nucleic acids in intermediate-resolution density maps: cryo-electron microscopy defines the quaternary structure of the *Escherichia coli* 70S ribosome. *Structure Fold. Des.* 8, 937–948
- 79 Zhou, Z.H. *et al.* (2001) Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus. *Nat. Struct. Biol.* 8, 868–873
- 80 Baker, M.L. *et al.* (2003) Architecture of the herpes simplex virus major capsid protein derived from structural bioinformatics. *J. Mol. Biol.* 331, 447–456
- 81 Topf, M. and Sali, A. (2005) Combining electron microscopy and comparative protein structure modeling. *Curr. Opin. Struct. Biol.* 15, 578–585

### Forthcoming EMBO FEBS FEMS Summer School

**Molecular basis of bacterial virulence and survival within infected hosts and the environment**  
Island of Spetsai, Greece  
5–15 September 2006

The course offers a unique blend of informal Mediterranean atmosphere and a scientific program at the cutting edge of microbiology, cell biology and pathophysiology. There will be 80 students accepted on the course and those who are interested should contact the organizers (Pascale Cossart, Institut Pasteur, France, e-mail: pcossart@pasteur.fr; Roberto Kolter, Harvard Medical School, USA, e-mail: rkolter@hms.harvard.edu) for additional information.

The deadline for applications is 15 April 2006 and accepted students will be notified by 15 June.

For more information visit: <http://spetses.med.harvard.edu/>