Mini-Review

Molecular Machines: Putting the Pieces Together

Eva Nogales* and Nikolaus Grigorieff‡

*Howard Hughes Medical Institute, Department of Molecular and Cell Biology, University of California Berkeley and Life Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720; and ‡Howard Hughes Medical Institute, W.M. Keck Institute for Cellular Visualization, Rosenstiel Basic Medical Research Center, Brandeis University, Waltham, Massachusetts 02454

The use of EM and image analysis techniques, particularly the single particle methodology, has opened the door to the structural characterization of large protein complexes in the megadalton range. This review covers the basic principles of single particle EM and recent examples of its application to the study of large molecular machines in the eukaryotic cell. These examples will give the reader a perspective on how this emerging technique can be applied to the study of large protein complexes to determine their overall shape, or the location of subunits and domains within them, and to follow conformational changes accompanying intermediates in their functional cycle.

Structural Studies of Molecular Machines

We are entering a new and exciting time in biology. The advent of full genome information for several organisms, including humans, provides a new tool to probe cellular structure and function. Structural genomics is now on its way and promises to accelerate the production of structural models at the atomic resolution for many proteins and nucleic acids. However, even the most optimistic structural biologist realizes that the extrapolation from structure to function will be difficult to ascertain. Numerous approaches, most importantly microchip technology for genomic expression, will work in parallel with the structural effort on the task of linking genes to cellular function.

In this postgenomic era, the structure determination of large, complex protein ensembles will pose a particularly difficult problem for structural biologists. It is now clear that most functions in the cell are not carried out by single protein enzymes, colliding randomly within the cellular jungle, but by macromolecular complexes containing multiple subunits with specific functions (Alberts, 1998). Many of these complexes are described as "molecular machines." Indeed, this designation captures many of the aspects characterizing these biological complexes: modularity, complexity, cyclic function, and, in most cases, the consumption of energy. Examples of such molecular machines are the replisome, the transcriptional machinery, the spliceosome, and the ribosome. The size of these macromolecular complexes, alone, often makes them inaccessible to x-ray crystallography, since crystals have to be obtained first and phasing such large structures is difficult. In most cases, with the clear exception of the ribosome, these complexes are relatively sparse in the cell, and purification yield only small amounts of sample, usually insufficient for crystallization trials. Reconstitution of large complexes from recombinant proteins create other difficulties, such as maintaining correct stoichiometry and order of assembly, and often, the composition of a complex is simply not known. As a result, many efforts have concentrated on determining the structures of individual subunits and domains within the machines. Consequently, information on the organization of the assembly subunits, their interactions, and sometimes their precise function within the context of the fully functional complex is lost. The most promising technique for the structural characterization of large molecular machines is EM and image reconstruction. In particular, single particle techniques overcome the need for crystals, making this method generally applicable to the structure determination of molecular machines. Furthermore, several reconstructions of complexes differing only slightly in composition or conformation can be obtained relatively quickly, and even conformational changes (on a millisecond time scale) can be trapped by rapid freezing. In this review, we summarize some of the principles underlying the potential of this methodology. We discuss examples of present and future applications to the characterization of molecular machines. Finally, we consider the prospects of improving these methods to reach high resolution and high throughput.

Single Particle Reconstruction

In diffraction-based methods, the interaction of the probe (x-rays or electrons) with biological matter results in a scattering process that carries information on the atomic organization in the sample. Unfortunately, biological macromolecules are very sensitive to radiation damage during this process. For EM, this means the number of electrons passing through the sample has to be kept small. The resulting low signal to noise ratio must be overcome by aver-

Address correspondence to Eva Nogales, Howard Hughes Medical Institute, LSA 355, UC Berkeley, CA 94720-3200. Tel.: (510) 642-0557. Fax: (510) 643-6791. E-mail: Enogales@lbl.gov
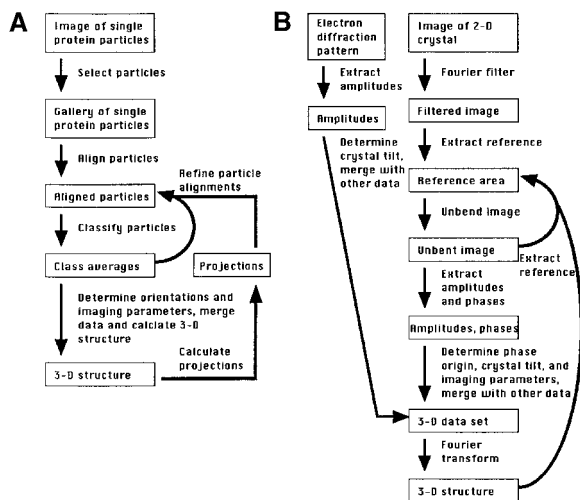
Figure 1. Processing of images of single protein particles and 2-D crystals. (a) After selecting particles from images scanned into a computer, the particles are aligned and classified according to shape and orientation. The classification yields averages with improved signal to noise ratio that can be used as new references to realign the particles. After a few iterations of alignment and classification, the orientations corresponding to the views represented by the class averages are determined. The data is merged with data from other images after determining the imaging parameters (image defocus and magnification). Assuming the data set contains a sufficient number of different orientations, a 3-D structure is calculated. (b) The image of a 2-D crystal is scanned into a computer, it is Fourier-transformed, the noise between diffraction spots is masked out, and the filtered image is back-transformed. To correct for distortions in the crystal (unbending), a reference area containing between 10 and 100 unit cells is selected from the filtered image. An improved reference area is derived from the unbent image, and the original image is unbent a second time. Spot amplitudes and phases are then extracted from the unbent image and merged with data extracted from other images after determining the sample parameters (phase origin and crystal tilt) and imaging parameters (defocus, beam tilt, and amplitude scaling). A 3-D structure is then calculated using a simple Fourier transformation. If the crystals are sufficiently large ($\sim$10 μm), amplitudes can also be obtained from electron diffraction patterns. After determining the crystal tilt, the diffraction amplitudes are merged with image data where they replace the image amplitudes (diffraction amplitudes are more reliable than amplitudes derived from images).

aging many images of protein particles of identical composition and conformation. Crystallographic methods (using x-rays or electrons) take advantage of the molecular alignment provided by the crystal to obtain information averaged over many molecules. The principle of single particle averaging can be used to substitute the crystallographic ordering by computational alignment of imaged molecules (Frank et al., 1988), sometimes referred to as crystallization "in silico." Images contain both amplitude and phase information. A 3-dimensional (D)[1] structure can be calculated using basic mathematical principles to determine the relative orientations of different 2-D views of one object. The order and size of a protein crystal determine the

---

[1]*Abbreviations used in this paper:* D, dimensional; EF-G, elongation factor G; NPC, nuclear pore complex; RC, regulatory complex.

---

power and resolution of the diffraction data. Similarly, the accuracy of the alignment (the angular relationship between different views of the molecule) and the number of images averaged determine the reliability and resolution of the EM reconstruction from single particles. Fig. 1 shows a schematic of image analysis procedures in 2-D electron crystallography and in single particle reconstructions.

The single particle method is especially well suited for studying large biological complexes. In fact, the larger the complex, the easier it is to align different particles to achieve higher resolution. In addition, only a small amount of protein is required, on the order of a few picomole for a 25-Å reconstruction. At this resolution, the overall shape of a complex can be determined and domains may become visible.

Single particle reconstruction is not without problems, the most obvious being its limited resolution (currently, a resolution better than 15 Å is considered state of the art for particles without symmetry). Whereas theoretical calculations show that atomic resolution can in principle be obtained from single particles (Henderson, 1995; Glaeser, 1999), the resolution has been limited by technical problems. These include image contrast lower than predicted by theory, limited accuracy of alignment procedures, and the sheer magnitude of data and computational time required to reach 3 Å. Furthermore, certain conformations of domains or loops in a protein might be stabilized in a crystal if they are involved in crystal contacts. Such contacts are absent in isolated molecules and complexes, giving them a larger degree of conformational freedom and further limiting the resolution obtained using single particle averaging. Of all EM techniques, only electron crystallography of 2-D crystals has yielded atomic resolution for a handful of proteins (Henderson et al., 1990; Kühlbrandt et al., 1994; Kimura et al., 1997; Nogales et al., 1998; Murata et al., 2000).

What Can Be Learned from Single Particle Reconstructions?

Within the present range of attainable resolution, image reconstruction can provide not only an overall shape of the protein complex under study, but also information on the assembly process of the full molecular machinery, the distribution of subunits within it, and the overall conformational changes underlying its function. We have chosen examples from recent literature to illustrate how this methodology is used to elucidate the structure–function relationship in a variety of molecular machines in a cell. We will consider: (a) the determination of 3-D structures, (b) the mapping of subunits and active sites within a complex, (c) the visualization of different assembly stages of a complex, (d) conformational changes accompanying biological activity, and (e) a movie-like series of snapshots of functional states.

3-D Structure Determination

The structure of a large protein assembly resolved at 20-Å resolution provides the overall shape and domain architecture of the complex. The outline of the complex can be useful to correlate biochemical and functional information
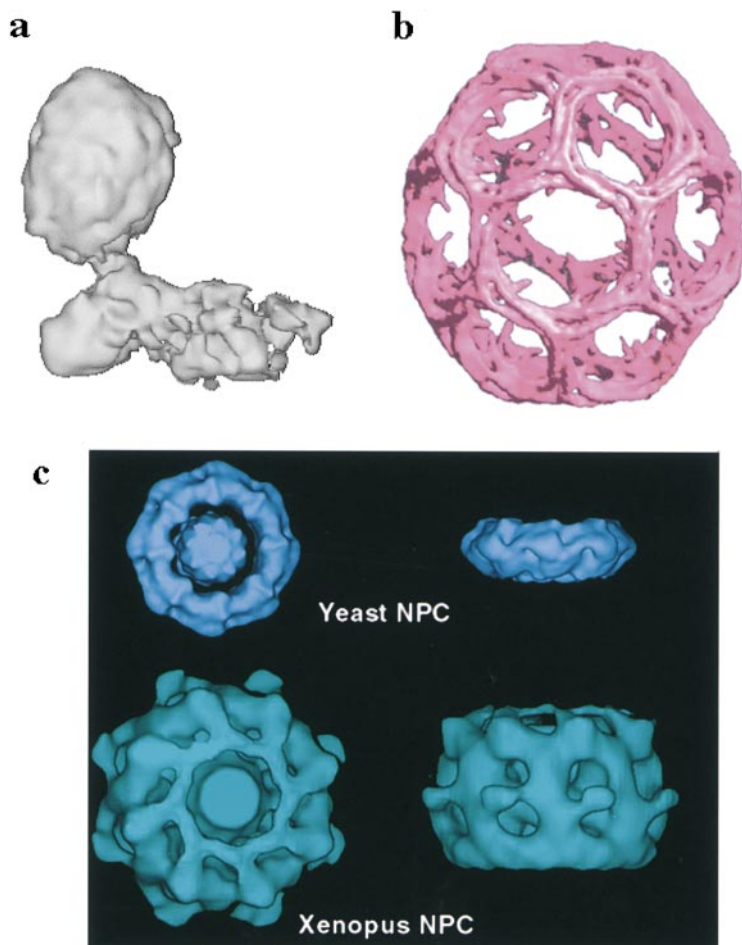
a **b**

c

Yeast NPC

Xenopus NPC

Figure 2. 3-D reconstructions of (a) complex I (Grigorieff, 1998), (b) clathrin cage (Smith et al., 1998) (C. Smith, Birbeck College, London, UK), and (c) yeast and vertebrate nuclear pore (Yang et al., 1998) (C. Akey, Boston University, Boston, MA).

with the coarse structural features visible at the 20 Å level, even if the location of individual subunits cannot yet be determined.

Mitochondrial Complex I. Complex I is the entry point for electrons in the electron transport chain of mitochondria. The bovine complex consists of 43 different subunits with a total molecular mass of $\sim$900 kD. It can be dissociated into two subcomplexes, Iα and Iβ, containing 23 and 17 subunits, respectively. Iα contains mostly hydrophilic subunits, including those that bind NADH, FMN, and all the Fe-S clusters identified by electron paramagnetic resonance. Iβ is part of the membrane domain of complex I, since it contains at least two of the hydrophobic subunits. The 3-D structure of bovine complex I (Grigorieff, 1998) (Fig. 2 a), complex I from *Neurospora crassa* (Guénebaut et al., 1997), and *Escherichia coli* (Guénebaut et al., 1998) have been studied by single particle EM. At 17-Å resolution (Fig. 2 a), the bovine complex revealed two separate domains linked by a thin stalk. The domains were identified as membrane and matrix domains by comparing their shape to other complex I structures determined earlier at a lower resolution, and by correlating their volume with the approximate volumes calculated for subcomplexes Iα and Iβ. The structure implies that the thin stalk contains part of the electron transfer pathway through complex I.

Clathrin Cages. Clathrin is the major coat protein for vesicles involved in endocytosis and in the latest stage of the biosynthetic pathway. The clathrin coat is capable of

deforming the membrane into an invaginated bud in which selected recycling receptors bearing appropriate signals are concentrated. The 3-D structure of a clathrin cage at 21-Å resolution reveals a hexagonal barrel detailing the packing of entire clathrin molecules as they interact to form the cage (Smith et al., 1998) (Fig. 2 b). This structure provides a structural basis for the assembly and disassembly of clathrin coats on a membrane. Each triskelion contributes to two connecting edges of the polyhedral cage to form a tightly packed structure that maximizes the area of contact between the legs. The terminal domains continue inward towards the adaptor complex, forming "hook-like" structures and a network with other legs. The terminal domain density could be interpreted using the x-ray structure of the 55-kD NH$_2$-terminal fragment (Musacchio et al., 1999). The structure surrounding the threefold vertices is fairly open and may be accessible to other proteins.

Nuclear Pore Complex. The nuclear pore complex (NPC) creates an aqueous channel across the nuclear envelope through which macromolecular transport between nucleus and cytoplasm occurs. Nucleocytoplasmic traffic is bidirectional and involves diverse substrates, including protein and RNA. The NPC comprises multiple copies of $\sim$50 different nucleoporins, which form a scaffold for proteins involved in binding and translocation of macromolecules across the nuclear membrane. Low resolution models of vertebrate and yeast NPCs were derived using single particle methods. EM revealed important differences,
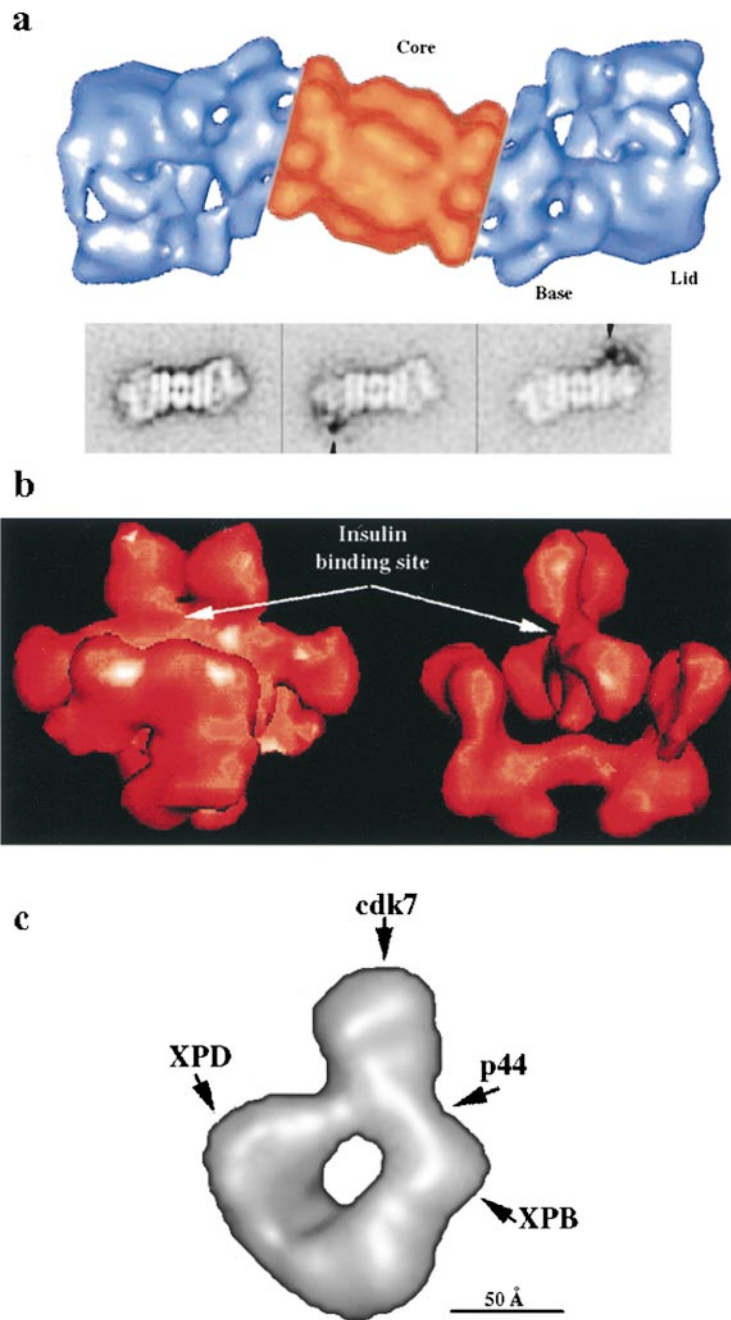
a

Core

Base   Lid

b

Insulin
binding site

c

cdk7

XPD   p44

XPB

50 Å

Figure 3. Localization of subunits and active sites within large protein complexes. (a) 26S proteosome (core region in orange, lids in blue) (Walz et al., 1998): the use of gold-conjugated substrate was used to locate the position of the deubiquitinylating enzyme p37A (bottom) (Hölzl et al., 2000) (W. Baumeister, Max-Planck-Institute, Martinsried, Germany). (b) Insulin receptor: the location of the insulin-binding site was obtained using gold-labeled insulin (Luo et al., 1999) (F.P. Ottensmeyer, University of Toronto, Ontario, Canada). (c) TFIIH: four of its subunits were located using subunit-specific antibodies (Schultz et al., 2000) (P. Schultz, Centre National de la Recherche Scientifique, Illkirck, France).

within the overall similarity of the structures, that have arisen during evolution and may be correlated with differences in nuclear transport regulation (Akey, 1995; Yang et al., 1998) (Fig. 2 c). The NPC has eightfold rotational symmetry and a multidomain-spoke complex surrounding a central transporter framed by cytoplasmic and nucleoplasmic coaxial rings. Eight thin fibers involved in protein import (Rutherford et al., 1997) project from the NPC cytoplasmic ring, and a basket structure that participates in mRNP export (Kiseleva et al., 1998) extends from the nucleoplasmic ring.

Mapping Subunits and Active Sites

Ideally, the 3-D envelope of a multisubunit complex is combined with information about the location of individual subunits within it. Such information could be obtained,

for example, by labeling subunits using subunit-specific antibodies. If the complex is obtained from a recombinant source, it is sometimes also possible to introduce cysteines at specific sites that can be labeled with gold clusters. These labels produce additional density in the 3-D structure and allow a mapping of subunit locations. Often, atomic models obtained by x-ray crystallography or nuclear magnetic resonance are available for some of the subunits. These can be used to interpret the EM map by fitting the models into the low resolution density, a process referred to as "docking." Thus, even if a high resolution structure of the entire complex cannot be obtained due to the lack of suitable crystals, an approximate atomic structure can still be derived by fitting together components. However, an atomic model derived this way is likely to contain some degree of error, since some subunits may as-
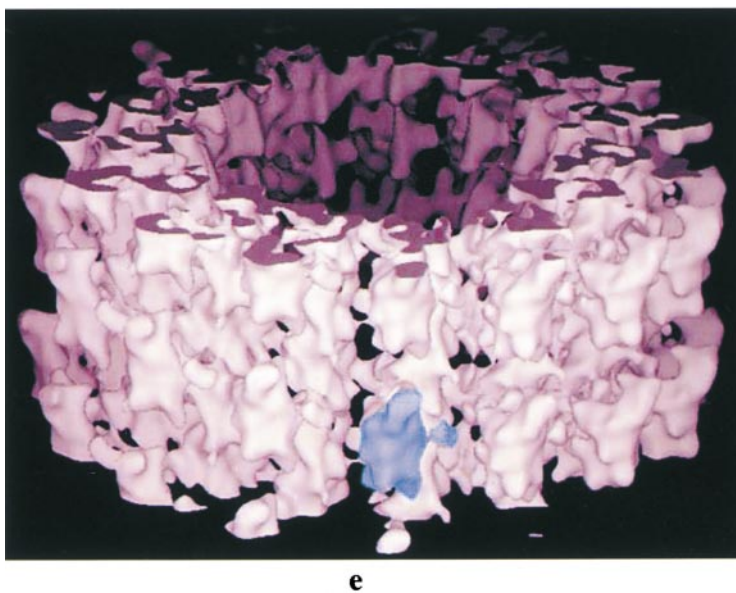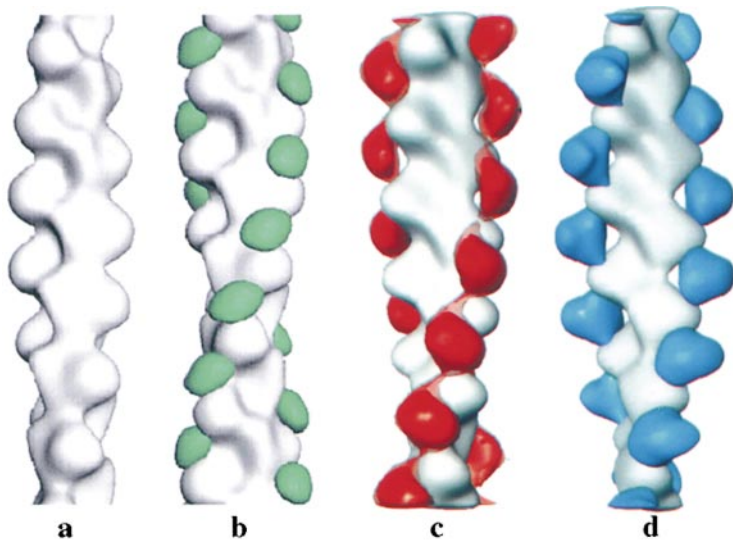
Figure 4. Structure of F-actin and actin-binding cyto-skeletal polymers and interacting proteins. (a) Undeco-rated F-actin, (b) F-actin decorated with cofilin (the he-lical screw is significantly altered) (McGough et al., 1997), (c) F-actin decorated with gelsolin, and (d) F-actin decorated with α-actinin (McGough et al., 1998) (A. McGough, Purdue University, West Lafayette, IN). (e) Helical reconstruction of a microtubule decorated with the motor domain of monomeric kinesin KIF1A (Kikkawa et al., 2000). One motor domain is shown in blue (M. Kikkawa, University of Tokyo, Tokyo, Japan).

sume different conformations in isolation and in states bound to other subunits in the complex.

Proteosome. Most eukaryotic proteins are degraded in the ubiquitin pathway by the 26S proteosome (2–3 MD) in an ATP-dependent process. Although the proteosome 20S core is well characterized (Fig. 3 a, orange), little is known about the regulatory complexes (RCs) that associate at both ends of the core and are involved in recognizing and unfolding proteins before degradation (Fig. 3 a, blue). The RCs contain hexameric AAA ATPases and up to 12 other subunits. One of these subunits is the deubiquitinylating enzyme p37A. Using a nonhydrolyzable substrate ana-logue, Hölzl and colleagues (2000) recently mapped its lo-cation. This is the first time that one of the RC's specific functions, deubiquitinylating activity, has been localized, thus providing new insights into the sequence of events leading to substrate degradation. Gold-conjugated inhibi-tor was mapped to the interface between the two RC sub-complexes, the base and the lid (Fig. 3 a, bottom).

Insulin Receptor. The insulin receptor is a constitutively dimeric receptor of the tyrosine kinase family. The 3-D structure of the insulin receptor bound to its ligand was determined by cryoEM and image reconstruction (Fig. 3 b) (Luo et al., 1999). Gold-labeled insulin served to locate the insulin-binding domain, whereas available high res-olution domain substructures were used to obtain a de-tailed model of this heterotetrameric receptor. This study showed that both subunits participate in insulin binding and that the kinase domains are in a juxtaposition that permits autophosphorylation in the first step of receptor activation.

Transcription Factor TFIIH. TFIIH is a complex of 12 different subunits shared by the transcription initiation and the transcription-coupled DNA repair machinery. Studies of 2-D crystals of a yeast core TFIIH containing five of the subunits have rendered an 18-Å resolution view of this part of the complex (Chang, 2000). The structure of the complete human TFIIH was obtained at a resolution of 38 Å using the single particle methodology (Schultz et al., 2000). The TFIIH is organized into a ring-like structure from which a large protein domain protrudes (Fig. 3 c). Im-munolabeling experiments have shown the relative posi-
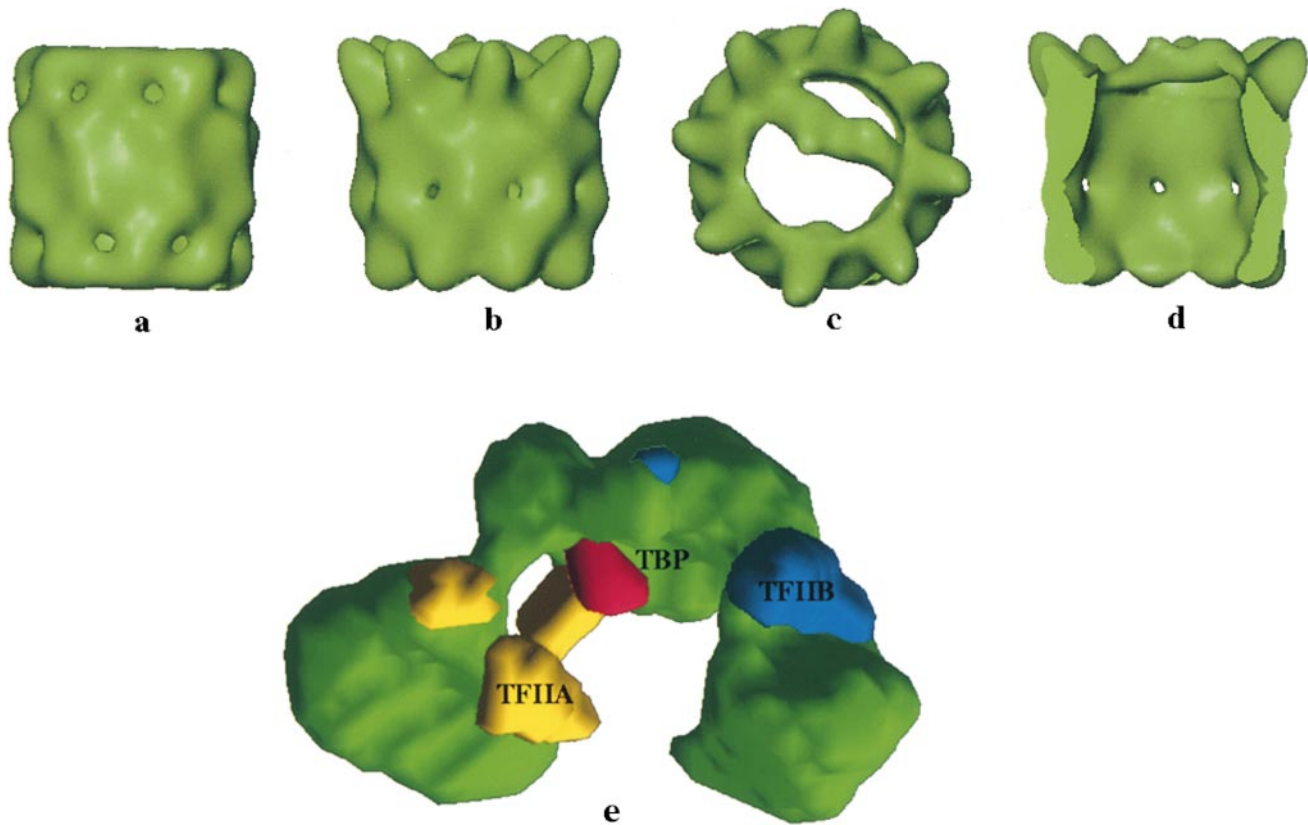
Figure 5. Reconstruction of CCT bound to α-actin. The partially folded actin molecule binds at the top of the apical domains of two of the subunits in the top cavity. (a) Unbound CCT; (b and c) orthogonal views of the actin-bound CCT; and (d) section through the actin-bound CCT (Llorca et al., 1999a) (J.M. Valpuesta, Centro Nacional de Biotechnologia, Madrid, Spain). (e) Reconstruction of human TFIID (green) bound to TFIIA (yellow) and TFIIB (blue). The position of TBP within TFIID was localized using a TBP-specific antibody (magenta) (Andel et al., 1999).

tions of some of the subunits within the complex. The helicases XPB and XPD are located roughly on opposite sides of the cyclic part of the structure, which surrounds a central cavity that could accommodate double-stranded DNA. The protrusion accommodates the kinase involved in phosphorylation of the COOH-terminal domain of Pol II.

Actin-binding Proteins. The use of EM and image reconstruction has been fruitful in the study of the actin cytoskeleton, particularly on the structural characterization of the binding interfaces of actin-binding proteins and their effect on the actin filament structure. These studies have generally used helical reconstruction to make use of the natural symmetry present in these cytoskeletal fibers to visualize F-actin decorated with different interacting proteins. This approach, combined with the docking of crystallographic structures, was most successfully used to determine acto-myosin interactions (Rayment et al., 1993). Since then, it has been used to describe the interactions of actin with α-actinin (McGough et al., 1994) (Fig. 4 d), cofilin (McGough et al., 1997) (Fig. 4 b), gelsolin (McGough et al., 1998) (Fig. 4 c), fimbrin (Hanein et al., 1997, 1998), tropomyosin, calponin (Hodgkinson et al., 1997), and scruin (Schmid et al., 1995), to cite some examples (for a review see McGough, 1998). These studies have provided invaluable information for understanding how these proteins cap, sever, cross-link, or destabilize the actin filament. Although helical reconstruction differs from single particle methods, both techniques deal with naturally occurring, physiologically relevant protein assemblies. Moreover, single particle methods are becoming increasingly useful to compensate for irregularities in helical structures.

Microtubule Motors. CryoEM and helical reconstruction have also been used extensively to obtain intermediate resolution structures of kinesin motors bound to microtubules. The crystal structures of the motors and tubulin have been docked into cryoEM envelopes to obtain detailed information on the contacts between these proteins (for review see Amos, 2000). A case of special interest is the study of microtubules decorated with KIF1A, a monomeric, yet processive, motor that differs from conventional kinesin by an insertion of a lysine-rich segment in loop L11 (Fig. 4 e). Docking of the tubulin and kinesin structures into the 15-Å reconstruction shows that the position of an extra density coincides with the position of the L11 loop and faces the acidic COOH-terminal end of the tubulin monomer (Kikkawa et al., 2000). This strongly suggests that an ionic interaction is responsible for the processivity of this motor. The interaction creates a tethering force during the motor step and avoids its diffusion away from the microtubule wall, even in the absence of a second head.

Conformational Changes

EF0F1. H$^+$-ATPases of the F0F1-type catalyze ATP synthesis coupled with proton transport in the membrane of mitochondria, chloroplasts, and bacteria. They have a
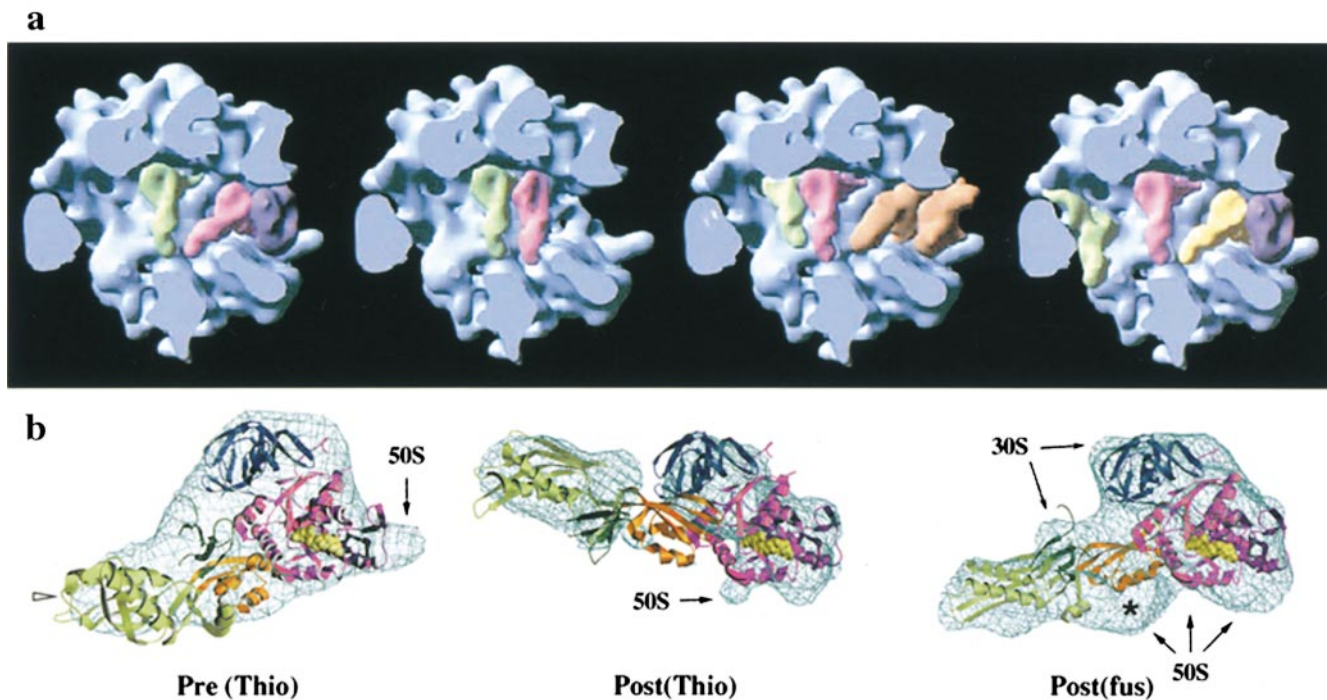
a

b

Pre (Thio)   Post(Thio)   Post(fus)

Figure 6. CryoEM studies of the ribosome and the translocation cycle. (a) Location of tRNAs (green, magenta, and yellow) and elongation factors (orange and purple) during the translation cycle (Frank et al., 1999) (J. Frank, Wadsworth Center, Albany, NY). (b) Conformations of EF-G (Stark et al., 2000). Orthogonal views of difference densities fitted with EF-G. The blue mesh shows the density attributed to EF-G from difference mapping. The structure of EF-G was fitted using the crystal structure of EF-G–GDP, but allowing for movements of domains 3, 4, and 5 (dark green, light green, and orange) relative to domains 1 and 2 (magenta and blue). Areas of contact with the 30S or 50S ribosomal subunits are indicated. Density due to the ribosome is marked by an asterisk (Stark et al., 2000) (W. Wintermeyer, Universität Witten, Witten, Germany).

transmembrane component, F0, involved in proton transport and a soluble part, F1, that contains the ATP-binding sites. Single particle image processing showed that F1 is connected by a prominent stalk to a more peripheral part of F0 (Böttcher et al., 2000). The 3-D maps of the enzyme both in the absence and presence of AMP-PNP showed that the nucleotide does not induce any significant changes in F0, whereas major conformational changes were observed in F1: decrease in diameter, appearance of a pointed cap, and a significant change in its inner cavity.

Chaperonins. Chaperonin complexes are required for the efficient folding of many proteins in a cellular context. Bacterial chaperonin GroEL and its coactivator GroES have been extensively characterized by x-ray and cryoEM (for review see Saibil, 2000). The latter has been particularly useful in the study of the chaperonin complex bound to substrate or in different nucleotide states (Chen et al., 1994; Rye et al., 1999).

The eukaryotic chaperonin CCT is formed by two stacked rings, each with eight different subunits, in contrast with the simpler, homooligomer GroEL. Valpuesta and coworkers have used single particle methods with images of frozen–hydrated samples of CCT to obtain 3-D reconstructions of both the apo and ATP-bound complexes (Llorca et al., 1999b). Whereas the apo-CCT has a symmetric, barrel-like structure, the binding of ATP generates an asymmetric particle. One of the rings in the ATP-CCT undergoes substantial movement in the apical domains that rotate and point towards the cylinder axis. The same authors have characterized the binding to CCT of α-actin,

one of the few specific substrates of this chaperonin (Llorca et al., 1999a) (Fig. 5, a–d). Using antibodies specific to two of the eight subunits, they have shown that the small domain of actin binds to CCTd and the large domain to CCTb or CCTe, clearly indicating that the binding of actin to CCT is subunit specific.

Complex Assembly

Some complexes go through distinct stages of assembly during their biochemical cycle. A classic example is the spliceosome, with at least four biochemically distinct assembly intermediates needed to facilitate the two chemical steps of splicing: lariat formation and exon ligation. The enzymatic activity of such complexes can be studied by trapping the intermediates, either biochemically or by rapid freezing. The 3-D structure of each intermediate can then be determined by EM, producing many snapshots to visualize the entire pathway. In some cases, the assembly pathway of a large supracomplex can be followed by sequentially adding components to a nucleating unit. The position of the additional elements can then be followed together with any conformational changes that occur.

Transcription Preinitiation. The general transcription factor TFIID contains about 12 different subunits involved in core promoter recognition and interaction with gene-specific activators. The binding of TFIID to the DNA promoter is followed by the recruitment of the rest of the general transcriptional machinery and RNA Pol II. The 3-D structure of this complex shows a horseshoe shape formed by three lobes, with a central cavity that could accommo-

date dsDNA or even a whole nucleosome (Andel et al., 1999; Brand et al., 1999) (Fig. 5 e). The lobes are connected by narrow regions of density, which suggests that relative movement of the lobes (that would change the size and shape of the central cavity) can be easily accomplished, perhaps upon binding of the complex to DNA or upon interaction with activators or other GTFs (general transcription factors). Reconstructions of multicomplexes that recapitulate the first steps in transcriptional machinery assembly have shown the binding sites of TFIIB and TFIIA on the large TFIID (Andel et al., 1999) (Fig. 5 e). The positions of these sites, together with the existing knowledge of the interaction between different TFIID components, suggest a hypothetical distribution of subunits on the trilobal TFIID structure.

Moving Machines: The Ribosome

Structural analysis of the ribosome and the translational cycle has driven the development of single particle EM, and the ribosome is the most successful example of this methodology. Joachim Frank and coworkers have obtained reconstructions of the ribosome in different stages of the translation cycle with the ribosome bound to different cofactors (tRNAs) (Agrawal et al., 1996), elongation factors (Agrawal et al., 1998), or the translocon channel (Beckmann et al., 1997). Similar efforts have been carried out by van Heel and coworkers (Stark et al., 1995, 1997, 2000). The EM efforts to elucidate the structure of the ribosome have developed in parallel with those of x-ray crystallography and have proven how the two methodologies are highly complementary. EM reconstructions of the ribosome at ∼20-Å resolution were initially used to phase x-ray data (Ban et al., 1998; Harms et al., 1999), whereas high resolution crystal structures of ribosomal components have been "docked" into the EM reconstructions of the full ribosome to gain atomic detail of some aspects of the translation cycle (Malhotra et al., 1998). The entire translation cycle has been visualized by obtaining a complete series of intermediate structures of the ribosome using single particle EM (Frank et al., 1999) (Fig. 6 a).

Recent articles have been dedicated to the detailed analysis of 3-D reconstructions of the ribosome intermediate states throughout translocation. Stark and co-workers have shown that elongation factor G (EF-G) has different conformations before and after translocation, resulting in different conformations of the ribosome in these two states (Stark et al., 2000). The authors docked the crystal structure of EF-G–GDP to the density attributed to the factor in both states and found that the movement of domains 1 and 2 with respect to domains 3, 4, and 5 is required to obtain a good fit (Fig. 6 b).

Frank and Agrawal have most recently shown that both binding of EG-F and GTP hydrolysis cause rotations of the 30S subunit with respect to the 50S subunit, in addition to various conformational changes in both subunits (Frank and Agrawal, 2000). The binding of the EG-F and an initial anti-clockwise rotation are accompanied by a widening of the mRNA channel. Upon hydrolysis, the tRNAs translocate, and then a clockwise rotation occurs that narrows the mRNA channel again.

Future Samples and Perspectives

The capabilities of single particle EM will certainly increase within the next few years. Only a few years ago, a resolution of ∼7 Å for hepatitis B virus particles (Böttcher et al., 1997; Conway et al., 1997) was celebrated as a breakthrough for this technique. The high resolution was greatly facilitated by the 60-fold symmetry of the virus. However, since then, work on asymmetrical particles, such as the ribosome, has progressed to a resolution of ∼10 Å (Gabashvili et al., 2000), impressive indeed, considering that only five years ago such structures yielded resolutions between 20 and 30 Å. This jump in resolution is most dramatically illustrated by realizing that at 10-Å resolution secondary structure (α-helices) becomes visible, whereas at 20–30 Å a complex could only be described as an arrangement of blobs. The improvement is mainly due to technical advances in EM, namely the introduction of the field emission electron source that generates a highly coherent beam to enhance image contrast. Further improvements in instrumentation are also likely to increase image quality. Instruments are now commercially available that combine field emission with increased voltage up to 300 kV (higher voltage reduces image distortions due to specimen charging), more stable specimen stages (some are cooled to liquid helium temperature for better sample cryoprotection), and better vacuum to reduce or avoid sample contamination. In the next 5 to 10 yr, current developments will also make it easier to collect large data sets. Microscopes will be equipped with fast sample loading mechanisms, automated data collection (Potter et al., 1999), and CCD cameras, avoiding cumbersome film development and scanning. Automation of data collection is particularly promising since this is often a rate-limiting step, especially when data sets containing a few hundred thousand particle images must be collected for high resolution. Progress is also being made in data processing. Not only is the computing power of a workstation increasing every year, but better algorithms for particle classification, alignment, and reconstruction (Grigorieff, 1998; Sigworth, 1998; Lanzavecchia et al., 1999; Ludtke et al., 1999; Li et al., 2000; Pascual et al., 2000), using correct statistical weighting schemes, are key to high resolution. Automation and robust, easy to use image processing software will increase the throughput of structures determined by single particle EM. New algorithms will also be developed to dock atomic models of subunits into a lower resolution map (Volkmann and Hanein, 1999; Wriggers et al., 1999). This procedure will become more important as the number of available atomic models increases, developing into a standard technique linking x-ray crystallography, nuclear magnetic resonance, and EM. Low resolution structures determined by EM will also be used to derive phases for x-ray diffraction data (Prasad et al., 1999; Reinisch et al., 2000). Hybrid techniques like docking and low resolution phasing combine the strengths of all structural techniques in common use today, making them the most powerful tools in the structural analysis of large complexes in the cell.

The number of macromolecular complexes of great biological relevance that await structural characterization is still very large. In many cases, such complexes are available in quantities too small for crystallization trials, or are

too unstable or perhaps too flexible to form crystals at all. The complexes may be modular with different components joining the assembly at different stages in its functional cycle. Examples of such complexes are the replisome machinery involved in DNA replication, the spliceosome complexes needed for RNA processing, or chromatin remodeling complexes required to allow gene replication. The big unknown of the molecular motors, dynein, and its interaction with dynactin, are in the minds of many biologists interested in the cytoskeleton. The development of cryoEM techniques and of single particle image processing methods for single particles will undoubtedly parallel the discovery and biochemical purification of many new and essential molecular machines. Our expectations for higher throughput and higher resolution should make possible the simultaneous biochemical and structural characterization of these fascinating macromolecular assemblies in our effort to elucidate their mechanisms of function.

References

Agrawal, R.K., P. Penczek, R.A. Grassucci, Y. Li, A. Leith, K.H. Nierhaus, and J. Frank. 1996. Direct visualization of A-, P-, and E-site transfer RNAs in the *Escherichia coli* ribosome. *Science.* 271:1000–1002.

Agrawal, R.K., P. Penczek, R.A. Grassucci, and J. Frank. 1998. Visualization of elongation factor G on the *Escherichia coli* 70S ribosome: the mechanism of translocation. *Proc. Natl. Acad. Sci. USA.* 95:6134–6138.

Akey, C.W. 1995. Structural plasticity of the nuclear pore complex. *J. Mol. Biol.* 248:273–293.

Alberts, B. 1998. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell.* 92:291–294.

Amos, L.A. 2000. Focusing-in on microtubules. *Curr. Opin. Struct. Biol.* 10:236–241.

Andel, F., A.G. Ladurner, C. Inouye, R. Tjian, and E. Nogales. 1999. Three-dimensional structure of the human TFIID-IIA-IIB complex. *Science.* 286:2153–2156.

Ban, N., B. Freeborn, P. Nissen, P. Penczek, R.A. Grassucci, R. Sweet, J. Frank, P.B. Moore, and T.A. Steitz. 1998. A 9 angstrom resolution x-ray crystallographic map of the large ribosomal subunit. *Cell.* 93:1105–1115.

Beckmann, R., D. Bubeck, R. Grassucci, P. Penczek, A. Verschoor, G. Blobel, and J. Frank. 1997. Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science.* 278:2123–2126.

Böttcher, B., S.A. Wynne, and R.A. Crowther. 1997. Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature.* 385:88–91.

Böttcher, B., I. Bertsche, R. Reuter, and P. Gräber. 2000. Direct visualization of conformational changes in EF(0)F(1) by electron microscopy. *J. Mol. Biol.* 296:449–457.

Brand, M., C. Leurent, V. Mallouh, L. Tora, and P. Schultz. 1999. Three-dimensional structures of the TAFII-containing complexes TFIID and TFTC. *Science.* 286:2151–2153.

Chang, W.-H. 2000. Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. *Cell.* 102:609–613.

Chen, S., A.M. Roseman, A.S. Hunter, S.P. Wood, S.G. Burston, N.A. Ranson, A.R. Clarke, and H.R. Saibil. 1994. Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature.* 371:261–264.

Conway, J.F., N. Cheng, A. Zlotnick, P.T. Wingfield, S.J. Stahl, and A.C. Steven. 1997. Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature.* 385:91–94.

Frank, J., and R.K. Agrawal. 2000. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature.* 406:318–322.

Frank, J., A. Verschoor, T. Wagenknecht, M. Radermacher, and J.-M. Carazo. 1988. A new non-crystallographic image-processing technique reveals the architecture of ribosomes. *Trends Biochem. Sci.* 13:123–127.

Frank, J., A.B. Heagle, and R.K. Agrawal. 1999. Animation of the dynamical events of the elongation cycle based on cryoelectron microscopy of functional complexes of the ribosome. *J. Struct. Biol.* 128:15–18.

Gabashvili, I.S., R.K. Agrawal, C.M.T. Spahn, R.A. Grassucci, D.I. Svergun, J. Frank, and P. Penczek. 2000. Solution structure of the *E. coli* 70S ribosome at 11.5 angstrom resolution. *Cell.* 100:537–549.

Glaeser, R.M. 1999. Review: electron crystallography: present excitement, a nod to the past, anticipating the future. *J. Struct. Biol.* 128:3–14.

Grigorieff, N. 1998. Three-dimensional structure of bovine NADH: ubiquinone oxidoreductase (complex I) at 22 Å in ice. *J. Mol. Biol.* 277:1033–1046.

Guénebaut, V., R. Vincentelli, D. Mills, H. Weiss, and K.R. Leonard. 1997. Three-dimensional structure of NADH-dehydrogenase from *Neurospora crassa* by electron microscopy and conical tilt reconstruction. *J. Mol. Biol.* 265:409–418.

Guénebaut, V., A. Schlitt, H. Weiss, K. Leonard, and T. Friedrich. 1998. Consistent structure between bacterial and mitochondrial NADH:ubiquinone oxidoreductase (complex I). *J. Mol. Biol.* 276:105–112.

Hanein, D., P. Matsudaira, and D.J. DeRosier. 1997. Evidence for a conformational change in actin induced by fimbrin (N375) binding. *J. Cell Biol.* 139:387–396.

Hanein, D., N. Volkmann, S. Goldsmith, A.M. Michon, W. Lehman, R. Craig, D. DeRosier, S. Almo, and P. Matsudaira. 1998. An atomic model of fimbrin binding to F-actin and its implications for filament crosslinking and regulation. *Nat. Struct. Biol.* 5:787–792.

Harms, J., A. Tocilj, I. Levin, I. Agmon, H. Stark, I. Kolln, M. van Heel, M. Cuff, F. Schlunzen, A. Bashan, F. Franceschi, and A. Yonath. 1999. Elucidating the medium-resolution structure of ribosomal particles: an interplay between electron cryo-microscopy and x-ray crystallography. *Structure.* 7:931–941.

Henderson, R. 1995. The potential and limitations of neutrons, electrons and x-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* 28:171–193.

Henderson, R., J.M. Baldwin, T.A. Ceska, F. Zemlin, E. Beckman, and K.H. Downing. 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* 213:899–929.

Hodgkinson, J.L., M. el-Mezgueldi, R. Craig, P. Vibert, S.B. Marston, and W. Lehman. 1997. 3-D image reconstruction of reconstituted smooth muscle thin filaments containing calponin: visualization of interactions between F-actin and calponin. *J. Mol. Biol.* 273:150–159.

Hölzl, H., B. Kapelari, J. Kellermann, E. Seemuller, M. Sumegi, A. Udvardy, O. Medalia, J. Sperling, S.A. Muller, A. Engel, and W. Baumeister. 2000. The regulatory complex of *Drosophila melanogaster* 26S proteasomes. Subunit composition and localization of a deubiquitylating enzyme. *J. Cell Biol.* 150:119–130.

Kikkawa, M., Y. Okada, and N. Hirokawa. 2000. 15 Å resolution model of the monomeric kinesin motor, KIF1A. *Cell.* 100:241–252.

Kimura, Y., D.G. Vassylyev, A. Miyazawa, A. Kidera, M. Matsushima, K. Mitsuoka, K. Murata, T. Hirai, and Y. Fujiyoshi. 1997. Surface of bacteriorhodopsin revealed by high-resolution electron crystallography. *Nature.* 389:206–211.

Kiseleva, E., M.W. Goldberg, T.D. Allen, and C.W. Akey. 1998. Active nuclear pore complexes in *Chironomus*: visualization of transporter configurations related to mRNP export. *J. Cell Sci.* 111:223–236.

Kühlbrandt, W., D.N. Wang, and Y. Fujiyoshi. 1994. Atomic model of plant light-harvesting complex. *Nature.* 367:614–621.

Lanzavecchia, S., P.L. Bellon, and M. Radermacher. 1999. Fast and accurate three-dimensional reconstruction from projections with random orientations via radon transforms. *J. Struct. Biol.* 128: 152–164.

Li, W., F.P. Ottensmeyer, and G. Harauz. 2000. Angular reconstitution of the *Staphylothermus marinus* phosphoenolpyruvate synthase. *Microsc. Res. Tech.* 49:233–244.

Llorca, O., E.A. McCormack, G. Hynes, J. Grantham, J. Cordell, J.L. Carrascosa, K.R. Willison, J.J. Fernandez, and J.M. Valpuesta. 1999a. Eukaryotic type II chaperonin CCT interacts with actin through specific subunits. *Nature.* 402:693–696.

Llorca, O., M.G. Smyth, J.L. Carrascosa, K.R. Willison, M. Radermacher, S. Steinbacher, and J.M. Valpuesta. 1999b. 3D reconstruction of the ATP-bound form of CCT reveals the asymmetric folding conformation of a type II chaperonin. *Nat. Struct. Biol.* 6:639–642.

Ludtke, S.J., P.R. Baldwin, and W. Chiu. 1999. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* 128: 82–97.

Luo, R.Z., D.R. Beniac, A. Fernandes, C.C. Yip, and F.P. Ottensmeyer. 1999. Quaternary structure of the insulin-insulin receptor complex. *Science.* 285: 1077–1180.

Malhotra, A., P. Penczek, R.K. Agrawal, I.S. Gabashvili, R.A. Grassucci, R. Jünemann, N. Burkhardt, K.H. Nierhaus, and J. Frank. 1998. *Escherichia coli* 70 S ribosome at 15 Å resolution by cryo-electron microscopy: localization of fMet-tRNAfMet and fitting of L1 protein. *J. Mol. Biol.* 280:103–116.

McGough, A. 1998. F-actin-binding proteins. *Curr. Opin. Struct. Biol.* 8:166–176.

McGough, A., B. Pope, W. Chiu, and A. Weeds. 1997. Cofilin changes the twist of F-actin: implications for actin filament dynamics and cellular function. *J. Cell Biol.* 138:771–781.

McGough, A., W. Chiu, and M. Way. 1998. Determination of the gelsolin binding site on F-actin: implications for severing and capping. *Biophys. J.* 74:764–772.

McGough, A., M. Way, and D. DeRosier. 1994. Determination of the alpha-

actinin–binding site on actin filaments by cryoelectron microscopy and image analysis. *J. Cell Biol.* 126:433–443.

Murata, K., K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J.B. Heyman, A. Engel, and Y. Fujiyoshi. 2000. Structural determinants of water permeation through aquaporin-1. *Nature.* 407:599–605.

Musacchio, A., C.J. Smith, A.M. Roseman, S.C. Harrison, T. Kirchhausen, and B.M. Pearse. 1999. Functional organization of clathrin in coats: combining electron cryomicroscopy and x-ray crystallography. *Mol. Cell.* 3:761–770.

Nogales, E., S.G. Wolf, and K.H. Downing. 1998. Structure of the αβ tubulin dimer by electron crystallography. *Nature.* 391:199–203.

Pascual, A., M. Barcena, J.J. Merelo, and J.M. Carazo. 2000. Mapping and fuzzy classification of macromolecular images using self-organizing neural networks. *Ultramicroscopy.* 84:85–99.

Potter, C.S., H. Chu, B. Frey, C. Green, N. Kisseberth, T.J. Madden, K.L. Miller, K. Nahrstedt, J. Pulokas, A. Reilein, D. Tcheng, D. Weber, and B. Carragher. 1999. Leginon: a system for fully automated acquisition of 1000 electron micrographs a day. *Ultramicroscopy.* 77:153–161.

Prasad, B.V., M.E. Hardy, T. Dokland, J. Bella, M.G. Rossmann, and M.K. Estes. 1999. X-ray crystallographic structure of the Norwalk virus capsid. *Science.* 286:287–290.

Rayment, I., H.M. Holden, M. Whittaker, C.B. Yohn, M. Lorenz, K.C. Holmes, and R.A. Milligan. 1993. Structure of the actin-myosin complex and its implications for muscle contraction. *Science.* 261:58–65.

Reinisch, K.M., M.L. Nibert, and S.C. Harrison. 2000. Structure of the reovirus core at 3.6 Å resolution. *Nature.* 404:960–967.

Rutherford, S.A., M.W. Goldberg, and T.D. Allen. 1997. Three-dimensional visualization of the route of protein import: the role of nuclear pore complex substructures. *Exp. Cell Res.* 232:146–160.

Rye, H.S., A.M. Roseman, S. Chen, K. Furtak, W.A. Fenton, H.R. Saibil, and A.L. Horwich. 1999. GroEL-GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. *Cell.* 97:325–338.

Saibil, H. 2000. Molecular chaperones: containers and surfaces for folding, stabilizing or unfolding proteins. *Curr. Opin. Struct. Biol.* 10:251–258.

Schmid, M.F., J. Jakana, W. Chiu, and P. Matsudaira. 1995. A 7-A projection map of frozen, hydrated acrosomal bundle from *Limulus* sperm. *J. Struct. Biol.* 115:209–213.

Schultz, P., S. Fribourg, A. Poterszman, M. Chipoulet, V. Mallouh, D. Moras, and J.M. Egly. 2000. Molecular architecture of human TFIIH. *Cell.* 102:599–607.

Sigworth, F.J. 1998. A maximum-likelihood approach to single-particle image refinement. *J. Struct. Biol.* 122:328–339.

Smith, C.J., N. Grigorieff, and B.M. Pearse. 1998. Clathrin coats at 21 A resolution: a cellular assembly designed to recycle multiple membrane receptors. *EMBO (Eur. Mol. Biol. Organ.) J.* 17:4943–4953.

Stark, H., F. Müller, E.V. Orlova, M. Schatz, P. Dube, T. Erdemir, F. Zemlin, R. Brimacombe, and M. van Heel. 1995. The 70S *Escherichia coli* ribosome at 23 Å resolution: fitting the ribosomal RNA. *Structure.* 3:815–821.

Stark, H., M.V. Rodnina, J. RinkeAppel, R. Brimacombe, W. Wintermeyer, and M. vanHeel. 1997. Visualization of elongation factor Tu on the *Escherichia coli* ribosome. *Nature.* 389:403–406.

Stark, H., M.V. Rodnina, H.J. Wieden, M. van Heel, and W. Wintermeyer. 2000. Large-scale movement of elongation factor G and extensive conformational change of the ribosome during translocation. *Cell.* 100:301–309.

Volkmann, N., and D. Hanein. 1999. Quantitative fitting of atomic models into observed densities derived by electron microscopy. *J. Struct. Biol.* 125:176–184.

Walz, J., A. Erdmann, M. Kania, D. Typke, A.J. Koster, and W. Baumeister. 1998. 26S proteasome structure revealed by three-dimensional electron microscopy. *J. Struct. Biol.* 121:19–29.

Wriggers, W., R.A. Milligan, and J.A. McCammon. 1999. Situs: a package for docking crystal structures into low-resolution maps from electron microscopy. *J. Struct. Biol.* 125:185–195.

Yang, Q., M.P. Rout, and C.W. Akey. 1998. Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications. *Mol. Cell.* 1:223–234.