Visualization of Biological Nano-Machines at Subnanometer Resolutions

Wah Chiu[†,††], Donghua Chen[†], Joanita Jakana[†], Juan Chang[†,†], Wen Jiang[†,†††], Steven J Ludtke[†] and Matthew L Baker[†]

[†]National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology
[††]Program in Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine
[†††]Current address: Department of Biological Sciences, Purdue University

Introduction

Based on genome sequence analysis, it has been suggested that a living cell conducts 200-300 distinct biological processes (Martin and Drubin, 2003), each of which can be carried out by macromolecular complexes. These complexes, which can both be persistent and transitory, are usually relatively large (>0.5 MDa) and contain specific and complex spatial organization of components to facilitate their dynamic behaviors. As such, these complex assemblies actually are biological nano-machines. As the name implies, an analogy can be drawn to almost any conventional man-made machine in which the machine is composed of multiple components operating in a coordinated fashion to perform one or more biological functions. These biological nano-machines can be found throughout the cell, responsible for key processes such as cell division and signaling, as well as being critical in diseases such as cancer or viral infections. Nano-machines are not restricted to "large-scale" cellular processes and may be involved in genome replication and repair, protein folding, protein trafficking and ion transport (Gavin et al., 2006; Krogan et al., 2006). These machines can assume different forms, shapes and symmetries. Therefore, knowledge of the molecular structures of the biological nano-machines, their components and their interactions can provide both mechanistic descriptions of function and clues for developing therapeutics related to health and disease. However, because of all these intrinsic complexities, structure elucidation of biological nano-machines at sufficient resolutions often poses a serious challenge.

Houston, TX 77030
E-mail: wah@bcm.edu

Using Single Particle Cryo-EM to Study Nano-Machines

Electron cryo-microscopy (also known as cryo-electron microscopy or cryo-EM) is the most appropriate imaging technique to derive 3-D structures of biological nano-machines because of several reasons (Chiu et al., 2006; Frank, 2002). In single particle cryo-EM, samples do not need to be crystallized as in X-ray crystallography. Rather, specimens can be prepared in frozen, hydrated state at different sample conditions without chemical fixative or negative stain. As the conditions attempt to mimic cellular conditions, often the frozen, hydrated samples represent a heterogeneous mixture of conformationally different particles. However, while this may be a drawback for conventional structure determination, single particle cryo-EM image processing techniques are capable of computationally "purifying" samples, thus capturing several states from a single experiment. Additionally, well-defined computational protocols can determine the 3-D structure of a sample directly from low-dose electron imaging conditions. Currently, the resulting reconstructions of conformationally rigid and uniform machines have achieved subnanometer (6-10 Å) resolutions (Chiu et al., 2005; Jiang and Ludtke, 2005). In this resolution range, secondary structure elements can be identified and high-resolution structures can be fitted to provide an understanding of the biological machine and its components (Chiu et al., 2002; Rossmann et al., 2005). Despite not being able to obtain atomic resolution structures of biological samples, cryo-EM continues to gain popularity in understanding biological processes as evidenced from the increasing number of publications (**Figure 1**) and the establishment of a public cryo-EM structure repository (http://www.ebi.ac.uk/msd-srv/emsearch/index.html).

A Single Particle Cryo-EM Experiment

As cryo-EM has assumed an increasingly important role in determining biological nano-machines, the complexity of imaging the samples to higher resolutions has also increased. However, the basic procedure (Jiang and Chiu, 2006; Serysheva et al., 2006) can be summarized as having two components: an experimental and a computational component (**Figure 2**). In the experimental portion of a cryo-EM experiment, the specimen must first be biochemically purified. While not as stringent of a restriction as other structural imaging modalities, the specimen should be relatively "pure" in terms of both conformation and function, meaning that the individual particles in sample should generally have the same structure and function. The purified sample can then be suspended in a buffer depending on the sample and the intended experiments. Typically, only a small amount of material is needed. For initial testing, as little as 100 μl at concentrations of 0.1-1 mg/ml are required. Subsequent higher resolution structural studies may require additional amounts of sample, however in some favorable cases high resolution structures may be obtained from the initial imaging of the specimen.

Once a suitably purified specimen has been obtained, the sample must then be prepared for imaging in the electron cryomicroscope. This step involves freezing the specimen on a holey grid (e.g. Quantifoil grid) which has been

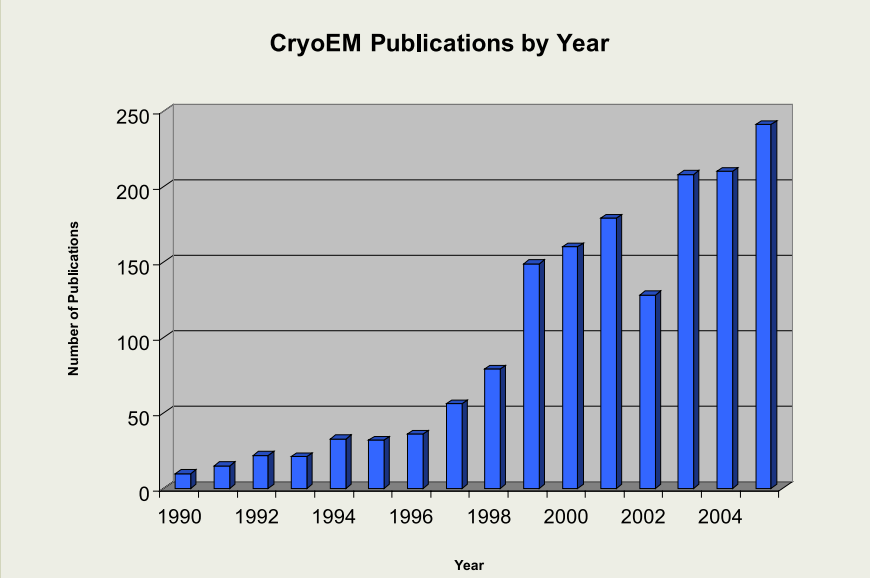

Fig. 1 Cryo-EM publications since 1990. A list of publications indexed by year is shown. This list of unique papers was generated by searching PubMed (http://www.ncbi.nlm.nih.gov/entrez/) using the Endnote software (http://www.endnote.com, The Thomson Corporation) with the following keywords: electron cryomicroscopy, cryoelectron microscopy, cryoEM and cryo-EM.

cleaned and pre-treated with glow discharging. 3-5 μl of sample can then be applied to the grid with the excess liquid blotted off with filter paper. At this point, the sample can be vitrified by plunging the grid with sample rapidly into a liquid ethane bath (Dubochet et al., 1988). To accomplish this, a variety of freezing apparatuses can be used from simple hand-operated devices, to pneumatic plungers or even temperature/humidity controlled plunging chambers from commercial companies. Once the sample has been frozen, it can then be stored almost indefinitely in a liquid nitrogen dewar or imaged immediately in the electron cryomicroscope.

The freezing step can often be the rate limiting step in the experiment. For any new specimen, it is generally necessary to perform trial-and-error freezing experiments to identify suitable conditions for obtaining a sufficiently thin layer of ice yet maintaining a reasonable number of randomly orientated particles across the holes of the holey carbon grid. Often variables such as specimen concentration, buffer, blot time, wait time, carbon support, temperature and humidity must be evaluated to achieve optimal imaging conditions. One of the more frequent challenges in this step is the presence of detergent, glycerol or sucrose which has been used to help solublize, purify or store the specimens. For imaging in the electron cryomicroscope, these types of molecules must be removed without disrupting the integrity of the complex itself.

The third step in a cryo-EM experiment is imaging frozen, hydrated specimen in the electron cryomicroscope equipped with a cryo-specimen transfer device (**Figure 1**). The specimen grid is suspended in a cryo-specimen stage that can generally be operated at either liquid nitrogen or liquid helium temperatures. Since biological specimens are highly radiation sensitive, low dose imaging is necessary to record the images. The practical dosage ranges from 10-40 electrons/Å^2 depending on the specimen temperature and the intended resolution of the study. Typically, images are recorded at 1-3 μm underfocus in order to provide sufficient contrast to visualize the particles while still maintaining high enough resolution structural information. The imaged specimen can be recorded on photographic film or directly to a CCD camera. When using film, the data has to be subsequently digitized using a scanner, such as Nikon Super Coolscan 9000ED. In addition to their convenience, CCD cameras provide an additional benefit as the low frequency component of the image is enhanced (Booth et al., 2004), making it easier to identify the particles. Conversely, the CCD has an additional modulation transfer function that may "smear out" high frequency data. Previous experience with a commercially available Gatan 4 k × 4 k CCD camera has shown that the signal up to 2/5 Nyquist frequency can be retrieved and used in the reconstruction (Booth et al., 2004), meaning that subnanometer resolution data can be obtained at moderate microscope magnifications (i.e. 80,000 ×).

In the second half of the cryoEM experiment, computational techniques are used to transform the 2-D images of the randomly orientated particles into an intermediate resolution 3-D structure, typically referred to as a reconstruction (**Figure 2**). The digital data, whether acquired directly from the CCD camera or digitized using a scanner, is subject to a series of data processing steps that determine the parameters associated with the contrast transfer function of the microscope and the envelope function of the images. These parameters serve to make corrections to the image data prior to merging the individual 2-D particle images into a single 3-D reconstruction.

There are numerous image processing packages such as Spider (Frank et al., 1996), IMAGIC (van Heel et al., 1996), MRC (Crowther et al., 1996), Frealign (Grigorieff, 1998), PFT (Baker and Cheng, 1996), XMIPP (Sorzano et al., 2004), IMIRS (Liang et al., 2002), SAVR (Jiang et al., 2001) and EMAN (Ludtke et al., 1999) that can be used to perform data processing and reconstruction steps. Conceptually, all the software packages use similar principles for image reconstruction. The reconstruction process is iterative, meaning the same procedure is applied to the raw experimental data multiple times, each time producing a slightly improved 3-D model. In the EMAN software suite (Ludtke et al., 1999), the iterative process involves first classifying each of the noisy raw particle images by comparing it to computational projections of the current 3-D model. Corrections to the effects of the contrast transfer function and envelop functions were made for each individual particle images prior to its classification. Classified particles are then rotationally and translationally aligned, and an average image is produced from the experimental data for each projection of the 3-D model. These averages are then used to construct the new 3-D model used for the next iteration of the reconstruction process. This extremely computationally intensive process is continued until the 3-D model is no longer improving. The 3-D model from the last iteration is then the final reconstruction of the specimen.

While the reconstruction represents the structure of the specimen in question, a considerable effort is required to decipher and analyze the architecture and composition of the

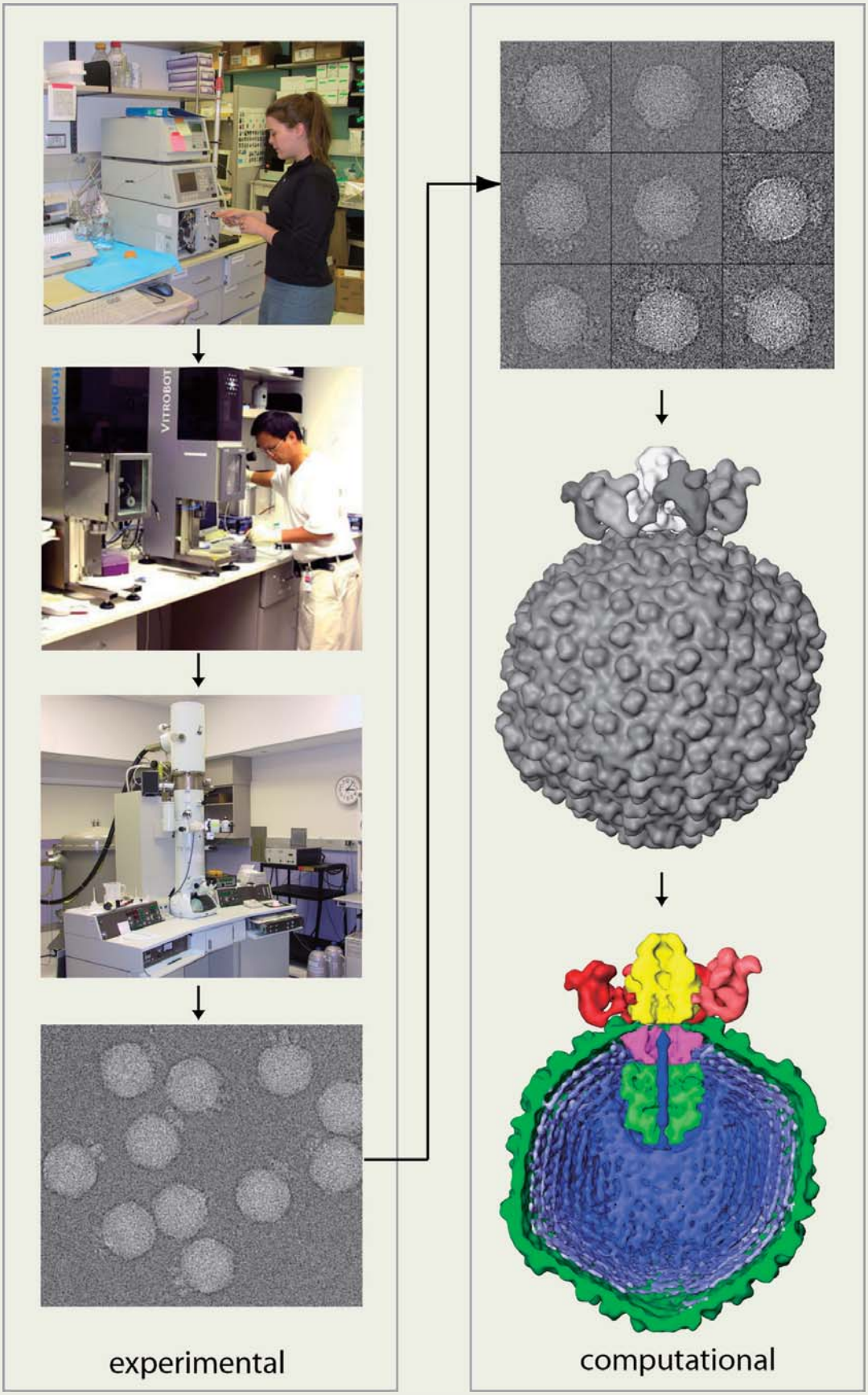

Fig. 2 From sample to model using cryo-EM. Shown is a generic flow diagram illustrating the basic procedure in a cryo-EM experiment. In the first half of the procedure (left panel), the "experimental" portion, the sample is purified, vitrified and imaged, from which raw micrographs or CCD frames of the sample are collected. Once the images have been collected, the data is subjected to the "computational" portion of the experiment which includes image processing, reconstruction, segmentation and analysis as shown on the right panel.

biological nano-machine. As these complexes can be quite large and composed of multiple molecular components, it can be challenging to interpret, let alone just visualize the reconstructed density map at subnanometer resolutions. The first in dissecting the intricate structure of the complex is to identify the individual subunits and segment them from the reconstruction. Again, numerous software packages and approaches are available for this procedure (Bajaj et al., 2003; Baker, 2006; Dougherty and Chiu, 1998; Volkmann, 2002; Yu and Bajaj, 2005). The majority of the approaches rely on the ability to isolate discrete density based by creating a mask and utilizing symmetry information when present. In many instances, the isolation of individual components is a hierarchical procedure, first identifying larger sub-groups of proteins, such as an asymmetric unit in a virus, followed by progressive finer segmentations until individual proteins or small domains have been identified (Baker, 2006).

Once the individual structural components have been isolated, analysis of the individual components can then be done. At subnanometer resolutions, this analysis includes the identification of secondary structure elements and fitting structures into the density map (Chiu et al., 2005; Chiu et al., 2002; Jiang et al., 2001). Additionally, new techniques have begun to arise that utilize the cryoEM density map as a constraint in building structural models (Mitra et al., 2005; Rossmann et al., 2005; Topf et al., 2006). These types of analysis are then combined with traditional genetic, biochemical and bioinformatics analysis to produce "higher resolution" structural models for the individual components. It is plausible that these individual structural models can then be integrated with the cryo-EM refinement, from which a more detailed structural and functional understanding of the complex can be gained.

GroEL, a Bacterial Chaperonin

With the relatively rapid growth and development of new technology in cryo-EM, it is important to demonstrate the fidelity and limitations in a cryo-EM experiment. For this purpose, it has been necessary to examine a specimen using the cryo-EM pipeline described in **Figure 2** in which the high resolution X-ray crystallographic structure is known. The bacterial chaperonin, GroEL, has been the subject of numerous structural studies by both X-ray crystallography and cryo-EM, including the highest resolution single particle cryoEM reconstruction to date at 6 Å resolution reported today (Ludtke et al., 2004). GroEL, composed of 14 identical subunits and sometimes associated with the capping protein, GroES, is a protein folding machine, with a molecular mass near 1 million Daltons. **Figure 3A** shows an image of frozen, hydrated GroEL recorded on a JEM-2010F cryomicroscope operated at 200 kV equipped with a Gatan cryo-specimen holder and 4 k × 4 k CCD camera. As discussed previously, the image contains particles randomly orientated and thus appears differently in the 2-D images. ~8000 individual GroEL particles were combined and used to compute the 9 Å resolution reconstruction (**Figure 3B**), based on the 0.5 criterion of the Fourier Shell Correlation technique, using the image processing package EMAN (Ludtke et al., 1999). At this resolution, it was possible to dissect and analyze the structure of a GroEL monomer (**Figure 3C**). Using SSEhunter (M. L. Baker, unpublished), individual a-helices greater than two turns were identified in the monomeric subunit (**Figure 3C**). Additionally, large flat density regions, corresponding to β-sheets could be readily identified. To confirm the GroEL structure and the secondary structural analysis, the GroEL X-ray structure (1OEL) (Braig et al., 1994) was fitted to the GroEL density, using FOLDHUNTER (Jiang et al., 2001), and compared (**Figure 3D**). The resulting comparison showed the relative accuracy in identifying nearly all of the secondary structure elements and illustrated the ability of the cryo-EM reconstruction to faithfully capture the native structure of GroEL. As such, the structure and the procedure used to generate the density map and model were validated. Experiments such as this one not only provide the necessary validation of the technique but also provide an excellent way to test new experimental and computational procedures.

Challenges to Higher Resolution

Modern cryo-EM instruments have a resolving power better than 3 Å even with imaging conditions used for the biological nano-machines at moderate magnification and low dose. Therefore, the question becomes if the instrumentation can achieve atomic resolutions, why is it not currently possible to reconstruct density maps to these resolutions? In fact, there are several plausible factors that currently limit high-resolution cryo-EM. First, electron beam induced specimen movement can blur the image anisotropically in a relatively unpredictable manner (Henderson, 2004). Though various attempts have been made to reduce this phenomenon, no general remedy has yet been introduced. Nevertheless, images of frozen, hydrated particles can be recorded in a liquid helium cryomicroscope with information contents beyond 6 Å as demonstrated from the visibility of the contrast transfer function rings from the power spectrum of the boxed-out particle images (**Figure 4**). Secondly, inadequacy of the image reconstruction algorithm can result in inaccurate particle orientations and/or improperly correcting the contrast transfer function and envelope function. Furthermore, conformational heterogeneity, due to the dynamic nature of the biological nano-machines, can limit the ability of the reconstruction software to accurately assign orientations and combine them in the 3-D reconstruction. As seen in X-ray crystallography map, different structural regions have different quality, usually referred to as a B-factor. Regions with high B-factors often represent regions of disorder or regions that are highly flexible. As the traditional single particle cryo-EM approach relies on averaging large numbers of individual particles, this type of flexibility, despite the same global structure, would produce reconstructions at resolutions limited by the extent of rigidity of the particle. Computational approaches have begun to address both flexibility and specimen heterogeneity and may result in more accurate and higher resolution structures.

Challenges in cryo-EM are not restricted to specimen, imaging and reconstructing the specimen. As the reconstructions achieve higher resolutions, the density maps grow in size. Current visualization software is capable of interactive visualization; however larger maps are nearing the maximum size and dimensions for standard desktop visualization software. While one would also assume that the subunits are more completely defined and thus easier to segment in higher resolution structures, the fine structural details in this resolution range in fact can blur the boundaries between the subunits and noise. Structural analysis can also be complicated by higher resolution structures. X-ray crystallography and NMR have an extensive repertoire of software designed for building structural models. However, only limited algorithms and applications have been developed by those pushing the resolution envelope in cryo-EM and as such complete packages for structural analysis do not yet exist.

Conclusion

Cryo-EM is a maturing field in structural biology that can determine structures of macromolecular complexes at a broad range of resolutions and can bridge the information gap between cell biology and crystallography/NMR. CryoEM also affords investigators a unique glimpse at these large, dynamic machines, which generally are not amenable to traditional structural approaches. As the field continues to expand, new techniques emerge and conquer previous roadblocks. The past ten years have seen cryo-EM reconstruction improved from 20-30 Å to subnanometer resolutions and now stand poised to achieve near atomic resolution structures in the near future. The realization of this goal will ultimately result in making an impact in both our understanding of biological processes and in health and medicine.

Acknowledgements

We thank the support of grants from NIH (P41RR02250, P01GM064692 and P20RR 020647) and NSF (EIA-0325004). The specimens provided by Dr. David Chuang at UTSW Medical School and Dr. Peter Weigele and Dr. Jonathan King are greatly appreciated.

References

- Bajaj, C., Yu, Z., and Auer, M. (2003). Volumetric feature extraction and visualization of tomographic molecular imaging. *J. Struct Biol.*, **144**, 132-143.
- Baker, M. L., Yu, Z., Chiu, W., and Bajaj, C. (2006). Automated segmentation of molecular subunits in electron cryomicroscopy density maps. *J. Struct Biol*., *in press*.
- Baker, T. S., and Cheng, R. H. (1996). A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J. Struct Biol.*, **116**, 120-130.
- Booth, C. R., Jiang, W., Baker, M. L., Zhou, Z. H., Ludtke, S. J., and Chiu, W. (2004). A 9 Å single particle reconstruction from CCD captured images on a 200 kV electron cryomicroscope. *J. Struct Biol*., **147**, 116-127.
- Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1994). The crystal

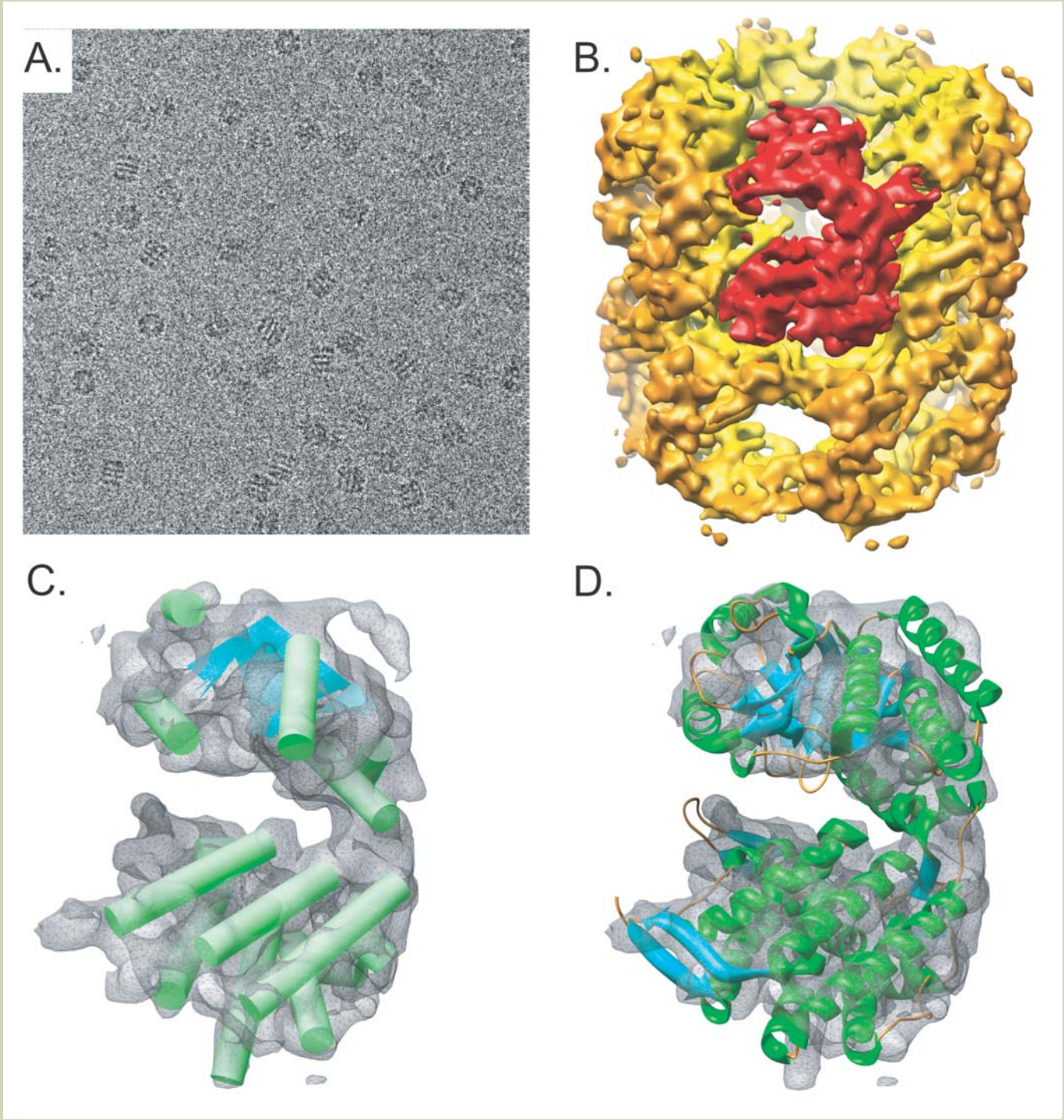

Fig. 3 The 9 Å structure of GroEL. A CCD image of GroEL from the JEOL JEM-2010F is shown in (A). ~8000 individual GroEL particles were used to reconstruct GroEL to 9 Å resolution (tilted side view and top view in (B). A single monomer is highlighted in red. A GroEL monomer was analyzed using SSEhunter, revealing helices and sheets (green and blue, respectively in C). The X-ray crystal structure of GroEL (1OEL) was fitted to the density corresponding to a single GroEL monomer using FOLDHUNTER, thus illustrating the effectiveness and accuracy of cryo-EM imaging (D).

structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature*, **371**, 578-586.

- Chiu, W., Baker, M. L., and Almo, S. C. (2006). Structural biology of cellular machines. *Tredns Cell Biol.*, **16**, 144-150
- Chiu, W., Baker, M. L., Jiang, W., Dougherty, M., and Schmid, M. F. (2005). Electron cryomicroscopy of biological machines at subnanometer resolution. *Structure*, **13**, 363-372.
- Chiu, W., Baker, M. L., Jiang, W., and Zhou, Z. H. (2002). Deriving folds of macromolecular complexes through electron cryomicroscopy and bioinformatics approaches. *Curr Opin Struct Biol.*, **12**, 263-269.
- Crowther, R. A., Henderson, R., and Smith, J. M. (1996). MRC image processing programs. *J. Struct Biol.*, **116**, 9-16.
- Dougherty, M. T., and Chiu, W. (1998). Using animation to enhance 3D visualization: a strategy for a production and environment. *Microsc and Microanal.*, **4**, 452-453.
- Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowall, A. W., and Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys.*, **21**, 129-228.
- Frank, J. (2002). Single-particle imaging of macromolecules by cryo-electron microscopy. *Annu Rev Biophys Biomol Struct.*, **31**, 303-319.
- Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A.

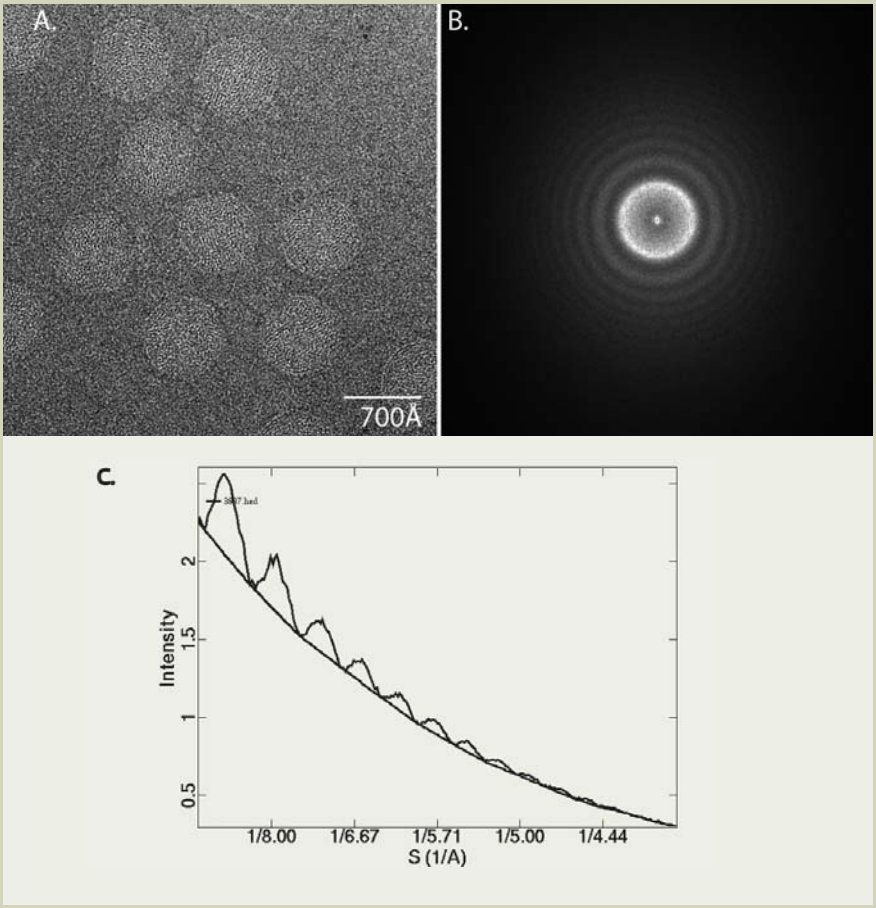

Fig. 4 Imaging Epsilon 15 bacteriophage. A micrograph at 1.3 µm defocus of Epsilon 15 from the JEOL JEM-3000SFF operated at 300 kV and specimen at 4.2K (A). FFT power spectrum of the boxed out particle images are shown (B). A circularly averaged power spectrum is shown in (C).

(1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct Biol.*, **116,** 190-199.

- Gavin, A. C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., Dumpelfeld, B., *et al.* (2006). Proteome survey reveals modularity of the yeast cell machinery. *Nature*.
- Grigorieff, N. (1998). Three-dimensional structure of bovine NADH:ubiquinone oxidoreductase (complex I) at 22 A in ice. *J. Mol Biol.*, **277**, 1033-1046.
- Henderson, R. (2004). Realizing the potential of electron cryo-microscopy. *Quart Rev Biophys.*, **37**, 3-13.
- Jiang, W., Baker, M. L., Ludtke, S. J., and Chiu, W. (2001). Bridging the information gap: computational tools for intermediate resolution structure interpretation. *J. Mol Biol.*, **308**, 1033-1044.
- Jiang, W., and Chiu, W. (2006). Cryo-electron microscopy of icosahedral virus particles, In Methods in Mol. Biol., J. Kuo, ed. (Totowa, NJ: The humana press), pp. in press.
- Jiang, W., and Ludtke, S. J. (2005). Electron cryomicroscopy of single particles at subnanometer resolution. *Curr Opin Struct Biol.*, **15**, 571-577.
- Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., *et al.* (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*, **440**, 637-643.
- Liang, Y., Ke, E. Y., and Zhou, Z. H. (2002). IMIRS: a high-resolution 3D reconstruction package integrated with a relational image database. *J. Struct Biol.*, **137**, 292-304.
- Ludtke, S. J., Baldwin, P. R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct Biol.*, **128**, 82-97.
- Ludtke, S. J., Chen, D. H., Song, J. L., Chuang, D. T., and Chiu, W. (2004). Seeing GroEL at 6 Å resolution by single particle electron cryomicroscopy. *Structure*, **12**, 1129-1136.
- 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.
- Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C. L., Ban, N., and Frank, J. (2005). Structure of the E. coli protein-conducting channel bound to a translating ribosome. *Nature*, **438**, 318-324.
- Rossmann, M. G., Morais, M. C., Leiman, P. G., and Zhang, W. (2005). Combining X-ray crystallography and electron microscopy. *Structure*, **13**, 355-362.
- Serysheva, I. I., Chiu, W., and Ludtke, S. J. (2006). Single particle electron cryomicroscopy of the ion channels in the excitation-contraction coupling junction, In Methods in Cell Biol, Cellular Electron Microscopy, J. R. McIntosh, ed. (San Diego: academic press), pp. in press.
- Sorzano, C. O., Marabini, R., Velazquez-Muriel, J., Bilbao-Castro, J. R., Scheres, S. H., Carazo, J. M., and Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct Biol.*, **148**, 194-204.
- Topf, M., Baker, M. L., Marti-Renom, M. A., Chiu, W., and Sali, A. (2006). Refinement of protein structures by Iterative comparative modeling and cryoEM density fitting. *J. Mol Biol.*, **357**, 1655-1668.
- van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R., and Schatz, M. (1996). A new generation of the IMAGIC image processing system. *J. Struct Biol.*, **116**, 17-24.
- Volkmann, N. (2002). A novel three-dimensional variant of the watershed transform for segmentation of electron density maps. *J. Struct Biol.*, **138**, 123-129.
- Yu, Z., and Bajaj, C. (2005). Automatic ultrastructure segmentation of reconstructed cryo-EM maps of icosahedral viruses. IEEE Transactions on Image Processing. **14**, 1324-1337.