Image Processing and Interpretation in Structural Electron Microscopy

J. B. Heymann

National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, 50 South Dr, Bethesda, MD 20892

The aim in structural EM is to extract information from 2D micrographs and reconstruct the 3D structure of the specimen. The micrographs need to be converted into digital form by scanning film, or captured directly by digital camera. Numerous formats have been devised to store the data, and several software packages (such as Bsoft [1,2]) provide support for multiple image formats. However, users still have to deal with some conversion issues and the limitations inherent in the various image formats. In large-scale image processing projects, these details need to be taken into account for effective workflows.

One of the major techniques in 3D EM is single particle analysis, where large numbers of images of a homogenous preparation of a specimen are used to determine its 3D structure (Fig. 1). The numerous associated parameters require a database-like approach, implemented in various ways in software packages. In Bsoft [1,2], this has been encoded in parameter files using both the STAR and XML formats. Such an organization is also required to allow parallel or distributed processing over clusters of computers (e.g., the distributed processing system, Peach [3]). Ultimately, the results of such processing need to be deposited in a public database, such as the EMDB [4]. However, this necessitates an agreement on the conventions for geometry and other descriptors [5].

A second technique demanding significant computational resources is electron tomography, which offers the opportunity to obtain the 3D structures of individual particles, parts of cells and whole small cells. However, the low signal-to-noise ratio and the missing wedge impose severe limitations on interpreting tomograms. Both limitations can be overcome by averaging subvolumes containing identical particles. In cases where averaging is not possible, components of complex structures can be located, isolated by segmentation and modeled (Fig. 2). The caveat in segmentation and in interpretation in general is that it is subjective, and the more detail extracted from a tomogram, the more likely that the operator's bias will affect the outcome. The goal is therefore to develop good practices for modeling, such as matching built models with original maps and referencing external information.

References
[1] J. B. Heymann and D. M. Belnap *J Struct Biol* **157** (2007), 3-18.
[2] J. B. Heymann et al. *J Struct Biol* **161** (2008), 232-242.
[3] P. A. Leong et al. *Structure* **13** (2005), 505-511.
[4] http://www.ebi.ac.uk/msd/emdb/
[5] J. B. Heymann et al. *J Struct Biol* **151** (2005), 196-207 (Corrigendum: *J Struct Biol* **153**, 312)
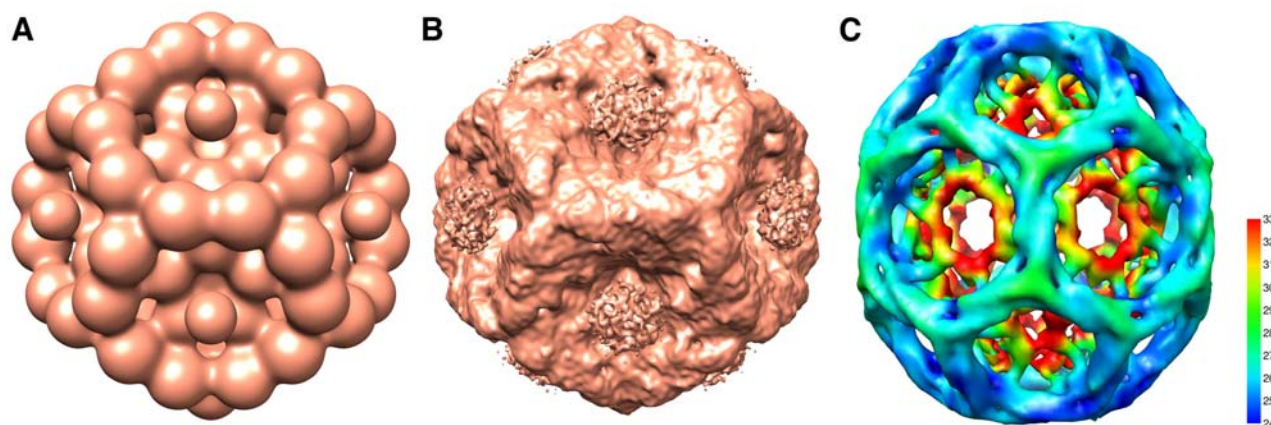[6] This work was supported by the Intramural Research Program of NIAMS/NIH.

Figure 1: (A) Single particle analysis is often done with a reference-based method, requiring an initial reference map that looks somewhat similar to the specimen. This starting reference for the *Pseudomonas syringae* bacteriophage φ6 procapsid was produced with only 4 spheres in the asymmetric, followed by icosahedral symmetrization. (B) The φ6 procapsid structure after several iterations of single particle orientation determination and reconstruction. (C) The resolution of a single particle reconstruction varies throughout the map. Estimation of the local resolution ($FSC_{0.5}$) of a reconstruction of a D_6 clathrin basket, indicates lower resolution for the more flexible inner shell heavy chain N-termini (red/yellow).
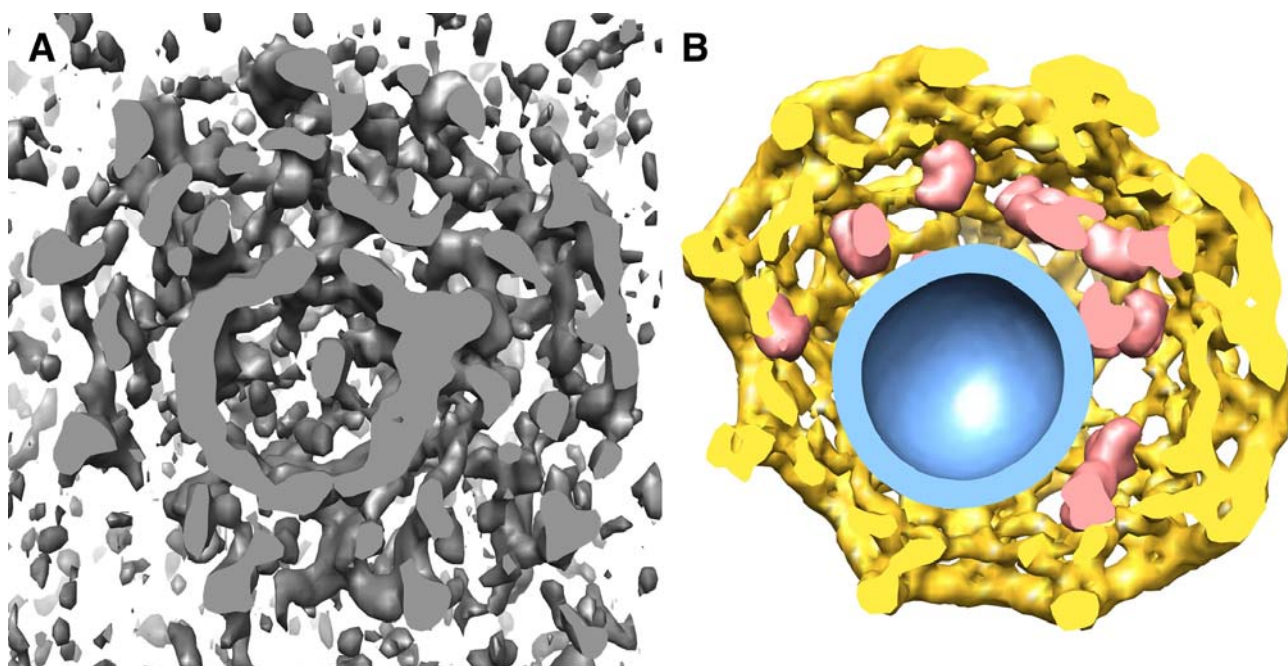

Figure 2: (A) A typical denoised tomogram still contains a lot of noise and material other than the particle of interest. This map of a clathrin-coated vesicle from bovine brain shows a central vesicle surrounded by accessory proteins and the clathrin polyhedral network. (B) Components of the particle were modeled to reflect the densities: blue, membrane (built as a point set); pink, adaptor protein cores (fitted using the atomic structure of AP-2); yellow, polyhedral clathrin coat (built from spars obtained from a single particle reconstruction).