**Title:**

Deriving *de novo* atomic models by cryo electron microscopy

**Authors:**

Peng Ge1, Nicole Poweleit1,2, Robert P. Gunsalus2,3, Z. Hong Zhou1,2,\*

**Author Affiliation:**

Ge, Peng 1,

1Electron Imaging Center for Nanomachines (EICN),

California NanoSystems Institute,

University of California,

Los Angeles

[gepeng1983@gmail.com](mailto:gepeng1983@gmail.com)

Poweleit, Nicole 1,2,

1Electron Imaging Center for Nanomachines (EICN),

California NanoSystems Institute,

University of California,

Los Angeles

2Department of Microbiology,

Immunology and Molecular Genetics,

University of California,

Los Angeles

[poweleit@ucla.edu](mailto:poweleit@ucla.edu)

Gunsalus, Robert P. 2,3,

2Department of Microbiology,

Immunology and Molecular Genetics,

University of California,

Los Angeles

3UCLA Institute of Genomics and Proteomics,

Los Angeles, CA

[robg@microbio.ucla.edu](mailto:robg@microbio.ucla.edu)

Zhou, Z. Hong 1,2,\*

1Electron Imaging Center for Nanomachines (EICN),

California NanoSystems Institute,

University of California,

Los Angeles

2Department of Microbiology,

Immunology and Molecular Genetics,

University of California,

Los Angeles

[Hong.Zhou@UCLA.edu](mailto:Hong.Zhou@UCLA.edu)

**Keywords:**

Electron microscopy, cryo electron microscopy, atomic modeling, TMV, CPV, eman, iterative helical real space reconstruction, single particle reconstruction

**\*Corresponding Author:**

Z. Hong Zhou,

[Hong.Zhou@UCLA.edu](mailto:Hong.Zhou@UCLA.edu)

**Short Abstract:**

Cryo electron microscopy (cryoEM) can be employed to derive *de novo* atomic models of macromolecular complexes in solution. The steps involved in high resolution cryoEM of biological molecules, from image recording, to data processing, to atomic modeling based on the resulting cryoEM density map, are illustrated.

**Long Abstract:**

Cryo electron microscopy is a structural biology technique for determining three-dimensional structures of supramolecular complexes in solution. In cryoEM, samples in their physiological, non-crystalline state are embedded in vitreous ice during low dose imaging. A three-dimensional density map of the sample is obtained by combining tens of thousands of cryoEM images. In recent years, the resolution of cryoEM has improved steadily, reaching a resolution sufficient for building atomic models from the cryoEM density map alone.

In this video article, we illustrate how such atomic models can be derived by cryoEM. Others have described the detailed procedures for preparing frozen hydrated samples, image acquisition and the basic steps of data processing of cryoEM [1](#_ENREF_1),[2](#_ENREF_2). We will focus on several criticle technical aspects important for reaching 3-4 Å resolution and for building reliable atomic models. We use the tobacco mosaic virus (TMV) and cytoplasmic polyhedrosis virus (CPV) as representative complexes with helical and icosahedral objects, respectively. *De novo* atomic models of these objects have previously been determined by cryoEM and published, and their biological relevance is described in previous publications[3-6](#_ENREF_3).

**Introduction:**

Beginning in the early 1970s, electron microscopy has been used in structural biology to investigate macromolecular assemblies at a “biologically significant” resolution [7](#_ENREF_7). Atomic resolution protein structures determined by x-ray crystallography or NMR were docked into the lower resolution (30-8 Å) electron microscopy density maps (for example [8](#_ENREF_8),[9](#_ENREF_9)). In this way, interactions between subunits of a large complex could be assessed to give hints at the overall function of a macromolecular assembly. While this approach is still frequently used for complexes with low or no symmetry, advances over the last twenty years in sample preparation, microscope design, data collection, and data processing have made *de novo* atomic model building, based on the EM density map alone, possible for certain types of samples [3](#_ENREF_3),[6](#_ENREF_6).

Three broad types of sample are suitable for *de nove* atomic model building by cryoEM: 2D crystalline arrays, such as bacteriorhodopsin [10](#_ENREF_10) and tubulin [11](#_ENREF_11), helical assemblies, such as tobacco mosaic virus [3](#_ENREF_3) and the bacterial flagellum [12](#_ENREF_12), and complexes with high symmetry, such as icosahedral viruses [4-6](#_ENREF_4),[13](#_ENREF_13) and GroEL [14](#_ENREF_14). Our laboratory has been successful in deriving several atomic models for helical assemblies and icosahedral viruses [3](#_ENREF_3),[5](#_ENREF_5),[6](#_ENREF_6). Based on past experiences we present the current article, which is intended to highlight some essential steps which allow for the electron density map to be refined to 3-4 Å resolution.

Below, we provide a working protocol with detailed steps for determining icosahedral and helical structures to about 3-Å resolution and for deriving atomic models. We also provide representative results and a discussion focusing on crucial steps in sample preparation and image processing for achieving atomic resolution structures. For in-depth understanding of the underlying theory and recent progress, viewers are advised to read recent reviews on this topic [15-17](#_ENREF_15" \o "Grigorieff, 2011 #63).

**Protocol:**

A general workflow of atomic model determination by cryoEM is given in Fig. 1. In the following protocol, items 1, 2 and 4 are shared in both single particle- and helical-structural determination. Item 3.1 is for single particle structures while item 3.2 is for helical structures. Therefore, one either will follow 1, 2, 3.1, 4 or 1, 2, 3.2, 4 depending on the nature of the sample being studied.

**1) Grid preparation, sample preparation, and image collection**

1.1) Place commercial cryoEM grids in a petri dish lined with several pieces of filter paper, and add 1,2-dichloro-ethane (ethylenedichloride, EDC) so that the liquid level is about one half to one centimeter. Gently push any floating grids into the liquid. Close the petri dish with a glass cover and seal with plastic film. Leave the sealed dish in a fume hood for a week, then remove from EDC and dry.

**Note:** Specific microscope settings will vary based on the type of microscope being used and the control software associated with the microscope. Results may also vary based on the microscope being used.

1.2) Coat dried grids with a fine layer of carbon by placing the grids on a flat open surface in a vacuum evaporator and establishing vacuum. Adjust a piece of graphite so it is held in the clamps above the grids and contact sparks will create fine carbon particles which coat the grids.

1.3) Layer carbon-coated grids on a closed style film image plate. Put this plate, with its projection chamber open, into the imaging plane of a 100kV electron microscope. Remove all apertures and irradiate with the smallest size beam which is sufficient to cover all the grids. Leave the grids in the microscope to “bake” for 10-50 hours. This approach removes all contaminants on the grids like oil from the oil pump of the vacuum evaporator and reduces charging of the grid during imaging.

1.4) Plunge freeze samples using an automated or manual freezing apparatus. Perform sample perpetration as demonstrated in other JoVE articles [1](#_ENREF_1),[2](#_ENREF_2).

1.5) Prepare the electron microscope for imaging by cryo-cycling at least overnight, and load the grid(s) into the electron microscope. To do this click the “cryocyle” button the microscope user interface.

**NOTE:** Cryocycle a function of the microscope, which warms up the microscope and cools it down to cryogenic temperature again.

1.6) Carefully perform direct alignment of the microscope, including coma free alignment (manually, not with the software). A well aligned microscope is essential for obtaining high resolution data. Two things are especially important for a better cryoEM image: minimal coma and higher coherence of the beam. We typically check the alignment every one to two days.

1.7) Take images on films or using a direct electron detector either manually (with low dose) or with leginon automation software. Use 1.5 µm defocus for larger particles and up to 3 µm defocus for particles which are difficult to see [3](#_ENREF_3),[5](#_ENREF_5),[6](#_ENREF_6).

1.8) Develop and scan the films to obtain an electronic copy of the data if appropriate or transfer electronic files to another location for data processing (See Fig. 2 for examples of good and bad data).

**2. Preprocessing of images**

Many of the steps in this and the following procedures require a large commitment of computer processing: transfer the files to a computer cluster if available.

2.1) If the data files are in a format other than .mrc, convert the data files to .mrc.

2.2) Determine the contrast transfer function (CTF) of each image using ctfffind3.exe [18](#_ENREF_18).

2.3) Create a filtered version of the data where the particles are clearly visible. There are many types of filters available through the proc2d command in EMAN [19](#_ENREF_19), use a Wiener filter (according to the ctf), a bilateral filter and a low pass filter at first, but use other filters or different filters if necessary to see the particles. Filtered images are ONLY used during pre-processing.

2.4) Examine the 2D power spectra of each image. Images that are well outside the defocus target value set while imaging (e.g., autofocus failed), that have any serious charging or drifting problem, or that are very astigmatic, can be ignored henceforth.

2.5) Examine the real space image. Any image with apparent ice contamination can be ignored henceforth.

**3. Image Processing**

**3.1) Image Processing for Icosahedral Single Particles (Windows or Linux).**

This protocol utilizes highly optimized programs in the IMIRS package[20](#_ENREF_20) and consists of three main processing steps -- orientation determination, orientation refinement and 3D Fourier inversion -- executed in an iterative fashion (Fig. 3).

3.1.1) For all particles picked, run *AutoOrtElim* to generate a list of selected particles with initial orientation and center parameters using common lines.

3.1.2) Run *hrefine* program to refine the orientation and center parameters for all particles with an initial orientation estimate. The program is based on minimization of phase residuals between the particle image and a set of projection images

3.1.3) Run *batch\_reconstruct\_ISAF* or a GPU-accelerated 3D reconstruction program, *g3d* (Ref. [21](#_ENREF_21)), to merge the Fourier transforms of all “good” particles according their refined orientation and center parameters and generate a 3D density map. Steps 3.1.1 and 3.1.2 are iterated until no further improvement in the 3D density map is obtained.

**3.2 Image Processing for Helical Structures (Linux)**

This protocol assumes that the reader has an in-depth understanding of EMAN software package (further than using graphic interface and default parameters). The work flow is outlined in Fig. 4.

3.2.1) Preparing segmented particles and 2D analysis

3.2.1.1) Box the particles using the helixboxer command of EMAN. The box size should be 3-4 times larger than the diameter of the particle for thin helixes and 60% larger for very thick helixes. The box size should be factorable to simple primes (e.g. 192, 256, 432, etc.)

3.2.1.2) Do not box particles or portions of particles which: are on the carbon, are on the rings at the edge of the micrograph, are off the micrograph, are obscured by contaminants, are overlapped by other particles, or have bad ice.

3.2.1.3) Create a directory to house the particles which will be boxed out and run the helixbatchboxer command.

3.2.1.4) Each particle should overlap the previous particle by one helixal turn. If the size of the helical turn is not known or too large, start with a 90% overlap.

3.2.1.5) Do phase-flipping ctf correction for each particle based on the astigmatic CTF parameters given by 2.2. Create a directory for the phase flipped particles and copy to the particles the ctf parameters which were determined in 2.2.

3.2.1.6) Create a refine2d directory, and copy the refine2d.py script to this location. Create or copy other accessory scripts which will be needed to run the refine2d.py script such as a script file (R2D.sh) for submission into the scheduler in the refine2d directory. Similarly, use script to group refinement commands too.

3.2.1.7) Create a start.hed file using lstcat.py.

3.2.1.8) Create class averages for the boxed particles. Use the start.hed file to run refine2d.py, this will result in reference free classification of all the particles. Set up the R2D.sh script to run the averaging process several times (about 20 iterations). For helices, a specially modified version of refine2d.py can be used to align the filament vertically.

3.2.1.9) Create an averaged Fourier transform of the final iteration of class averaging. Based on the resulting pattern, determine the helicity of the sample. This is traditional helical cryoEM work [22](#_ENREF_22).

3.2.1.10) To determine helicity, create a grid on the FFT using the perimerdian line and periequitorial line as the defining pattern. From this grid determine n,l, and m numbers for the equation l = nt + um. Use one line in the perimerdian and one in the periequitorial to set up a system of equations and solve for the unknowns, t and u; where t is the number of turns and u is the number of subunits (i.e. u/t = subunits/turn).

3.2.2) Generate the first structure

3.2.2.1) Use the *make3d* command from *EMAN* to create a preliminary 3D density map using the classaverages with random angular assignment.

3.2.2.2) Create a symdoc.dat file based on an initial estimation of the helical parameter. This file is in *Spider* document format.

3.2.2.3) Use *himpose* of *IHRSR* package [23](#_ENREF_23) to apply the helicity; then convert it back to mrc for further use in *EMAN*.

**NOTE:** If the helicity of the sample is not clear, several models will have to be built using different symmetry parameters and used as starting models for trial runs. The trial runs may or may not converge to the same helicity. The most plausible result will be used henceforth (plausible as judged empirically).

3.2.3) Refine the starting structure

3.2.3.1) Follow the guidelines given by the EMAN authors for refining a cryoEM structure. Between every two iterations, IHRSR programs are used to refine the helicity and the newly refined helicity is applied to the refinement result. maxshift=<pixel> (set to (100% - overlap) \* boxsize / 2) is appended to the refine command-line to prevent moving the new data (100% - overlap) of a segment out of the refinement box (volume).

3.2.3.2) Use EMAN style CTF correction (option ctfc=<resolution> in EMAN refine program). To simplify the CTF determination, EMAN style ctf parameter is built directly from ctffind3.exe results. Set amplitude to 1, bfactor to 0, noise baseline to a constant 0.

3.2.3.3) Use B-factor (Nikolaus Grigorieff group) program to boost high resolution factors of the structure. A typical refinement uses a 0 B-factor in initial refinement, a ~160 1/Å2 B-factor for further atomic resolution refinement. B-factor can be adjusted based on the microscope’s performance.

3.2.3.4) Reduce hsearch range by angular increments as the structure refines.

3.2.3.5) Set resolution cutoff and gradually increase the resolution as the structure refines.

3.2.3.6) Gradually decrease angular interval for projection as prescribed by EMAN authors. Projections can be limited to between 90 and 70-80 degrees altitude to save computational load, since helical objects have very limited out-of-plane tilt in the ice. (The filament aligns to the Z direction in the volume. When projected with 90 degrees altitude, the projection image should see a perfect side view of the filament.) A Vernier based angular sampling scheme can be used to reduce the number of necessary projections [3](#_ENREF_3" \o "Ge, 2011 #6).

**NOTE:** A typical refinement with 64,000 segments of 640x640 box can take 1000 cpu\*days for one iteration in the last few iterations. Reduce the overlap between adjacent segments to save computational load if necessary.

3.2.3.7) When the helicity refinement converges, use the final helicity (may be average from last few iterations) to refine the structure for several more iterations with Fourier space helicizing followed by real space symmetrizing (by himpose) [3](#_ENREF_3" \o "Ge, 2011 #6). This approach tremendously reduces the number of projections if the helical asymmetric unit is small enough (in terms of angular span) and enables refinement with even smaller angular intervals.

**4) Building the Atomic Model**

First build the atomic models for cryoEM structures with Coot and then refine the model structure with CNS [24](#_ENREF_24) and Phenix [25](#_ENREF_25). An example (CPV) of atomic modeling is illustrated in Fig. 5.

4.1)Start building the atomic level model before the refined structureis final. Use the preliminary atomic model to generate a theoretical radially averaged (one dimensional) structure factor for the purpose of scaling Fourier amplitudes [26](#_ENREF_26" \o "Zhang, 2010 #60). A B-factor of 40-100 1/Å2 can be used to sharpen the structure after scaling. Then continue to build and refine the atomic models to achieve the completed model.

4.1.1) Start the atomic model when the reconstruction maps achieve 4Å resolution: 4.5Å resolution maps can also be modeled but are more challenging (See Representative Results).

4.1.2) Trace C-alpha backbone using “baton-building” in Coot. When tracing the C-alpha, it is advisable to pay attention to the sequence as well and trying to have matching residue registration between the map and the sequence particularly at this step. This will simplify the next several steps.

**NOTE:** At a region of poor density, try to register the sequence before and after it. Then one will be able to more accurately trace the residues corresponding to this region.

4.1.3) Generate backbone model (poly A) with the appropriate program, typically Coot [27](#_ENREF_27) or REMO [28](#_ENREF_28).

4.1.4) Mutate the backbone model into full atomic model with the protein sequence using “Mutate Residue Range” in Coot.

4.1.5) Convert the cryoEM map into “hkl” pseudo-crystallographic reflections (CNS: em\_map\_to\_hkl.inp). Before conversion, the map is typically tailored (the map is cut into a smaller volume with “Tools -> Volume Data -> Volume Viewer -> Sub-region selection” in Chimera to save memory and segmented to include the subunits of interest only).

4.1.6) Mark out the testing set (5-10%) with CNS (make\_cv.inp, resulting in a .cv file).

4.2) Refine the automatically generated full atomic model with CNS against the reflections (.cv file). This time the refinement is limited to a single copy of a single protein at a time (segmentation needed). The refinement process is illustrated in Fig. 6. It starts with a crude model (Fig. 6a, green model) that is resulted from 4.1.4 above.

4.2.1) Use “vector residue” target instead of amplitude-only targets if there is reliable phase information. The refinement is more powerful when phases are also included.

4.2.2) Use simulated annealing to boil down the structure and make an initial good fit (CNS: anneal.inp). The key of not losing the fit is to lock down the C-alpha atoms with harmonic restraints. Use 20 as the constant (result in Fig. 6b, green model).

4.2.3) In Coot, manually fit any misfit residue (result in Fig. 6c, green models). (Typically, only <5% of all the residues are misfit.)

4.2.4) Refine group B-factor (CNS: bgroup.inp).

4.2.5) Minimize the structure (CNS: minimize.inp) with reduced (10) harmonic constant. Restrain both C-alpha and C-beta atoms this time.

4.2.6) Refine group B-factor again (CNS: bgroup.inp).

4.2.7) Minimize the structure again with more reduced (5) harmonic constant.

4.2.8) Refine group B-factor again (result in Fig. 6d, green model).

4.2.9) Put all refined structures of single proteins together. Use non-crystallographic symmetry (NCS) to deal with symmetry. Use either NCS constraint or restraint. When using restraint, make multiple copies of the proteins and dock them into symmetry-related copies. With the next iteration, release all harmonic restraints.

4.2.10) Manually adjust the structure if serious clashes appear after applying symmetry. These clashes typically result from long residues fitted to densities that belong to other copies. Use openmp, parallel version of CNS to reduce calculation time since refinement against an entire virus can be very time-consuming. Typical refinement of a virus takes a few days on 8 cpu cores.

4.2.11) Optionally, improve the structure by refining it in Phenix with Ramachandran restraints.

**Representative Results:**

The steps of cryoEM structure determination include sample purification and vitrification, low-dose imaging, orientation determination and refinement, 3D reconstruction, and atomic model building. First, samples particularly suitable for high-resolution cryoEM analysis are those with adequate size (>1MDa molecular weight, < 150 nm in thickness) such that sufficient contrast for visualization is provided, and with structural uniformity and integrity such that all particles are structurally identical for averaging. We show here an example of a good image of a suitable sample (Fig. 2c). (Notice that we intentionally selected a good image with some scanning defect – the Newton rings at the top-left corner of the image.) Second, a good cryoEM image should be without any ice contamination or drifting / charging. Figure 2a and 2b show an examples of ice-contaminated sample and an example of an image with drifting / charging, respectively. In an electron microscope, without applying any correction, one should see crystal lattices of crystal ice in the sample (Fig. 2a, top, inset), which give rise to the reflections in corresponding Fourier transforms (Fig. 2a, bottom, inset).

Although modern cryoEM instruments are capable of recording images at ~1Å resolution, when high-resolution cryoEM is concerned, the coherence of the beam plays a critical role in the resulting resolution. Therefore, we test our microscope alignment by taking test images with a CCD camera in two steps: first, the optical alignment is tested by imaging carbon film at 250,000x magnification with eight times low-dose dosage, under these conditions, the Thon rings should reach the point resolution of the microscope; second, the coherence of the beam is tested by imaging carbon film at 120,000x magnification with the working dose for cryoEM low-dose imaging, under these conditions, a single 4kx4k CCD frame should show Thon rings reaching 3Å resolution. Since the dosage is low, the Thon rings would have been hidden by noise if the signal over resolution drops off rapidly: an indication of poor beam coherence.

When reconstructed to different resolution, a cryoEM structure shows different features. Sometimes, the resolution of a structure can be judged empirically by its features, such judgments match resolution given by pseudo-crystallographic R-factors in the majority of cases. Figure 7 shows the TMV structure reconstructed (with resolution limit) to different (2.7Å, 3.3Å, 3.6Å, 4Å, 4.5Å, 5Å, 6Å, 7Å, 8Å, 10Å) resolutions. Notice that at 3.3Å resolution, side-chain features are very visible, sometimes showing difference between Val, Leu and Ile. At 3.6Å, loops are resolved, strands in β sheets are well separated and small side-chains start to blur. At 4.5Å resolution, strands become hard to resolve and only big (i.e., aromatic, arginine) side-chains can be seen. 5.5Å is the first resolution in which decent alpha-helical grooves are visible. At 6-7Å resolution, alpha-helices show up as sausages, although the size and length of them match well with that of the atomic model. At 8Å resolution, alpha-helices start to blur, and they are marginally distinguishable at 10Å resolution.

When building the model, the phase information in a cryoEM structure helps tremendously to automatically build the model, as compared to the amplitude-only data of X-ray crystallography. At 3.5Å resolution, atomic models can be built with ease. First, a c-alpha trace is built from the density map; be very sure that the amino acid registration is correct at this stage. When building loops that are not easily traced, be sure to consult structure elements before and after the loops so that there are no extra or missing residues in the loop (Fig. 6a, red). Then simply convert the tracing into a poly-alanine main-chain model and mutate the main-chain model to match the known sequence (Fig. 6a, green). Then run a simulated annealing with c-alpha atoms restrained by harmonic potentials (result: Fig. 6b, green). Fix any largely misfit residues (Fig. 6c, red to green). Then refine the atomic structure several times, gradually reducing the harmonic restraint potential until zero (result: Fig. 6d, green).

**Figure Legends:**

**Figure 1 Workflow of *de novo* atomic model generation by cryoEM**

This flowchart briefly lists the steps of atomic structure determination by cryoEM from sample preparation to final product. Light blue boxes represent sample preparation, red boxes represent imaging, yellow boxes represent micrograph processing, green boxes represent initial model building, dark blue boxes represent model refinement, and the purple box represents applications for the model.

**Figure 2 Examples of problematic micrographs**

(a) An image of an ice-contaminated sample. This kind of images contains diffraction contrasts. In a 300kV, high-end microscope, the lattices of the crystalline-ice areas are visible at higher magnification (inset). (b) A drifted image. The corresponding Fourier transform shows its characteristic envelope. (c) An acceptable image. Notice that we intentionally included some area with digitizing artifact for illustrative purpose (top left corner). The Fourier transform of this image has Thon rings reaching 3.5 Å resolution.

**Figure 3 Single-particle reconstruction**

Flow chart showing the basic steps of single-particle reconstruction implemented in the IMIRS package 17 (Protocol section 3.1 or the yellow boxes from Figure 1)

**Figure 4 Helical reconstruction**

Flow chart showing the basic steps of helical reconstruction using the Iterative Helical Real Space Reconstruction protocol (Protocol section 3.2 or the yellow boxes from figure 1.) Cyan boxes are typical of the IHRSR protocol. Yellow boxes mark improved steps for high resolution structure determination.

**Figure 5 Atomic modeling of a cryoEM reconstruction**

(a) CryoEM reconstruction of CPV. (b) Atomic models built for individual subunits. (c) Density maps with atomic models superimposed, showing good match between the features of the densities and their corresponding models. Most of the side-chain densities in this panel match well with their models.

**Figure 6 Refinement of atomic models against cryoEM density maps**

In each of the panels, red models stand for the result of the last step; green models stand for the result of this step; density maps are colored navy blue. (a) From Cα tracing to full atomic model. At this step, the atomic models don’t fully match the densities. (b) Simulated annealing, protocol 4.2.2. (c) Manually adjust misfit models. Some models are too far away from their corresponding densities to be effectively refined into these densities. Protocol 4.2.3. (d) Final refinements. Constraints are gradually released and models gradually refine into densities. Protocol 4.2.4-8.

**Figure 7 Progression of resolution during cryoEM structure refinement**

CryoEM structure of TMV reconstructed to different resolutions as marked on the panels. Notice how α helices become visible around 8 Å resolution, and large individual amino acids become identifiable around 4 Å resolution.

**Discussion:**

With advances in imaging hardware and software, cryoEM has come of age as a structural biology method enabling *de novo* atomic models based on density maps alone. However, some major limitations persist for this technique, in particular the limitations in sample types and preparations which are suitable for use in cryoEM. Thus when considering cryoEM for structural determination, the most important question to ask is still: is this the right technique for my sample [15-17](#_ENREF_15" \o "Grigorieff, 2011 #63)? While X-ray crystallography and NMR remain the best choice for small, individual proteins; cryoEM has developed into an excellent tool for obtaining structural data of macromolecular complexes, both at the level of individual proteins and protein interactions within complexes. Continued advances in physics and material science, computer science and computer processing, sample preparation procedures, and the very recent direct electron detection or counting technology, make cryoEM a growing field of interdisciplinary research which will continue to push limitations and provide unique structural insights.

To get an idea of how big the complex would be or could be, the size of our CPV and TMV structures may be referred to. CPV has an alignable protein shell of nearly 30 MDa in size, although smaller viruses like Hepatitis B Virus (HBV) capsid can also be determined to atomic resolution. On the other hand, the helical structure of TMV contains a 6.5 MDa segment of the much longer helical virus. Other unpublished structures demonstrate that one can reach the resolution necessary for a *de novo* atomic model with an alignable size as small as 4.5 MDa.

Several factors are important for reaching atomic resolution. First, the stability of the microscope plays a vital role, especially the stability of magnification. With a three condenser lens system, modern microscopes can control the parallelism of the beam very well. It was estimated that without the three condenser system, parallelism can be aligned to within 1% error; with the three condenser system, it can be aligned as well as 0.1% error. This helps to stabilize the magnification of the microscope tremendously (Wim Hagen, personal communications). However, several groups have previously achieved ~ 4Å resolution with earlier generation microscopes. Second, coma has a big negative effect on final resolution. Uncorrected residue beam tilt after coma free alignment generates considerable phase shift in images at high resolution and was considered to be the major barrier beyond 3Å resolution [29](#_ENREF_29). However, a poorly aligned beam tilt (including pivot point, a misalignment of which generates beam tilt when the beam is shifted) can be even worse. Third, the use of astigmatic CTF correction allowed us to have better defined phases in the images. This factor impacts models at about 4Å resolution. It is also interesting that it is the phases (via phase flipping) of the images that play a more critical role than the amplitudes. Proper astigmatic phase flipping with circular amplitude correction allows for a resolution sufficient for atomic modeling, whereas further application of astigmatic amplitude correction doesn’t obviously improve the resolution.

There are still more issues that are not possible to enumerate in a limited space. “Full-screen anti-aliasing” of the phase flipping process gives better features on the structure. The unceasing discussion of whether to integrate all data processing steps into a well-designed software package with optimized user-interface (as IMIRS package in protocol 3) or to leave all modules of a package accessible to expert users so that individuals can tweak parameters or even software code at the cost of poor user-interface (as combined IHRSR / EMAN 1 workflow in protocol 4) has yet to be concluded. (We consider the most recent package, EMAN 2 [30](#_ENREF_30), to be an acceptable balance between the two extremes, bearing a windowed user-interface with usable default parameters and a command-line interface for advanced users to customize the process.) In any case, specialized handling of sample and data requires far more than words to describe and working with a group or individual with considerable experience in *de novo* atomic model building may be essential.

**Acknowledgements:**

Xuekui Yu provided CPV data for analysis. We acknowledge funding from NIH (GM071940 and AI094386) and NSF (DBI-1338135), PG and NP receive funding from American Heart Association Western Affiliate (13POST17340020) and theNIH Biotechnology Training Program (T32GM067555), respectively.

**Disclosures:**

The authors declare no conflict of interests.

**References:**

1 Meng, X., Zhao, G. & Zhang, P. Structure of HIV-1 capsid assemblies by cryo-electron microscopy and iterative helical real-space reconstruction. *J Vis Exp*, doi:10.3791/3041 (2011).

2 Meyerson, J. R. *et al.* Determination of molecular structures of HIV envelope glycoproteins using cryo-electron tomography and automated sub-tomogram averaging. *J Vis Exp*, doi:10.3791/2770 (2011).

3 Ge, P. & Zhou, Z. H. Hydrogen-bonding networks and RNA bases revealed by cryo electron microscopy suggest a triggering mechanism for calcium switches. *Proc Natl Acad Sci U S A* **108**, 9637-9642, doi:10.1073/pnas.1018104108 (2011).

4 Yu, X., Ge, P., Jiang, J., Atanasov, I. & Zhou, Z. H. Atomic model of CPV reveals the mechanism used by this single-shelled virus to economically carry out functions conserved in multishelled reoviruses. *Structure* **19**, 652-661, doi:10.1016/j.str.2011.03.003 (2011).

5 Liu, H. *et al.* Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science* **329**, 1038-1043 (2010).

6 Zhang, X., Jin, L., Fang, Q., Hui, W. H. & Zhou, Z. H. 3.3 A cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell* **141**, 472-482 (2010).

7 Beer, M., Frank, J., Hanszen, K. J., Kellenberger, E. & Williams, R. C. The possibilities and prospects of obtaining high-resolution information (below 30 A) on biological material using the electron microscope. Some comments and reports inspired by an EMBO workshop held at Gais, Switzerland, October 1973. *Q Rev Biophys* **7**, 211-238 (1974).

8 Bharat, T. A. M. *et al.* Structure of the immature retroviral capsid at 8[thinsp]A resolution by cryo-electron microscopy. *Nature* **487**, 385-389, doi:10.1038/nature11169 (2012).

9 Li, S., Hill, C. P., Sundquist, W. I. & Finch, J. T. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**, 409-413, doi:10.1038/35030177 (2000).

10 Kimura, Y. *et al.* Surface of bacteriorhodopsin revealed by high-resolution electron crystallography. *Nature* **389**, 206-211, doi:10.1038/38323 (1997).

11 Nogales, E., Wolf, S. G. & Downing, K. H. Structure of the alpha beta tubulin dimer by electron crystallography. *Nature* **391**, 199-203, doi:10.1038/34465 (1998).

12 Yonekura, K., Maki-Yonekura, S. & Namba, K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**, 643-650, doi:10.1038/nature01830 (2003).

13 Zhang, X. *et al.* Structure of Sputnik, a virophage, at 3.5-A resolution. *Proc Natl Acad Sci U S A* **109**, 18431-18436, doi:10.1073/pnas.1211702109 (2012).

14 Ludtke, S. J. *et al.* De novo backbone trace of GroEL from single particle electron cryomicroscopy. *Structure* **16**, 441-448, doi:10.1016/j.str.2008.02.007 (2008).

15 Grigorieff, N. & Harrison, S. C. Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Curr Opin Struct Biol* **21**, 265-273, doi:10.1016/j.sbi.2011.01.008 (2011).

16 Zhou, Z. H. Atomic resolution cryo electron microscopy of macromolecular complexes. *Adv Protein Chem Struct Biol* **82**, 1-35, doi:10.1016/B978-0-12-386507-6.00001-4 (2011).

17 Zhou, Z. H. Structures of viral membrane proteins by high-resolution cryoEM. *Curr Opin Virol* **5**, 111-119, doi:10.1016/j.coviro.2014.04.001 (2014).

18 Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**, 334-347 (2003).

19 Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128**, 82-97, doi:10.1006/jsbi.1999.4174 (1999).

20 Liang, Y., Ke, E. Y. & Zhou, Z. H. IMIRS: a high-resolution 3D reconstruction package integrated with a relational image database. *J Struct Biol* **137**, 292-304 (2002).

21 Zhang, X. & Zhou, Z. H. Low cost, high performance GPU computing solution for atomic resolution cryoEM single-particle reconstruction. *J Struct Biol* **172**, 400-406, doi:10.1016/j.jsb.2010.05.006 (2010).

22 De Rosier, D. J. & Klug, A. Reconstruction of Three Dimensional Structures from Electron Micrographs. *Nature* **217**, 130-134 (1968).

23 Egelman, E. H. The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. *J Struct Biol* **157**, 83-94, doi:S1047-8477(06)00168-7 [pii]10.1016/j.jsb.2006.05.015 (2007).

24 Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nat. Protocols* **2**, 2728-2733, doi:10.1038/nprot.2007.406 (2007).

25 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

26 Zhang, J. *et al.* Mechanism of folding chamber closure in a group II chaperonin. *Nature* **463**, 379-383, doi:10.1038/nature08701 (2010).

27 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

28 Li, Y. & Zhang, Y. REMO: A new protocol to refine full atomic protein models from C-alpha traces by optimizing hydrogen-bonding networks. *Proteins: Structure, Function, and Bioinformatics* **76**, 665-676, doi:10.1002/prot.22380 (2009).

29 Zhang, X. & Zhou, Z. H. Limiting factors in atomic resolution cryo electron microscopy: no simple tricks. *J Struct Biol* **175**, 253-263, doi:10.1016/j.jsb.2011.05.004 (2011).

30 Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* **157**, 38-46, doi:10.1016/j.jsb.2006.05.009 (2007).