

Complementing crystallography: the role of cryo-electron microscopy in structural biology

Jonathan M. Grimes,^{a*}
Stephen D. Fuller^b and David I.
Stuart^{a,c}

^aStructural Biology Division, The Wellcome Trust Center for Human Genetics, Roosevelt Drive, Headington, Oxford OX3 7BN, England,

^bStructural Biology Programme, European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstrasse 1, 69117 Heidelberg, Germany, and ^cOxford Centre for Molecular Sciences, New Chemistry Building, South Parks Road, Oxford OX1 3QT, Oxford

Correspondence e-mail:
jonathan@strubi.ox.ac.uk

Dramatic improvements in experimental methods and computational techniques have revolutionized three-dimensional image reconstruction from electron micrographs (EM) of vitrified samples. Recent results include the first determination of a protein fold (for the core protein of the hepatitis B virus) by non-crystalline imaging techniques. These developments have generated interest within the crystallographic community and have led to a re-evaluation of the technique, particularly amongst those working in the field of virus structure or struggling with the phasing of large macromolecular assemblies. A simple discussion of the techniques of EM image reconstruction and its advantages and problems in terms familiar to crystallographers will hopefully allow an appreciation of the essential complementarity of the two techniques and the practical potentials for phasing applications.

Received 28 June 1999

Accepted 20 July 1999

1. Introduction: advantages of cryo-electron microscopy

The past few years have seen image reconstruction from single particles attain a certain degree of adolescent maturity, becoming a general tool in the armoury of the structural biologist rather than an esoteric rite practised by the high priests of electron microscopy. This has been a consequence of developments in techniques for handling delicate biological samples and for analysing much larger numbers of data as well as the development of improved microscopes. This paper is not an in-depth review of state-of-the-art techniques of cryo-electron microscopy (cryo-EM) reconstruction (see, for instance, Mancini *et al.*, 1997; Chiu & Schmid, 1997; Baker *et al.*, 1999). Instead, we try to draw to the attention of the protein crystallographer the general ideas behind image reconstruction and those technological improvements which have increased the detail in three-dimensional images derived from electron micrographs, in one case to almost 7 Å resolution (Böttcher *et al.*, 1997).

A fundamental advantage of cryo-EM compared with X-ray crystallography arises from the fact that cryo-EM usually works with images representing projections of the structure, rather than (as in the case of X-ray diffraction) measuring the amplitudes of the three-dimensional diffraction data. Cryo-EM therefore maintains the phase information. The use of unstained specimens motivates the taking of images at a defined underfocus to enhance the phase contrast in the micrograph. The effect of this enhancement is described quantitatively by the contrast-transfer function (CTF), which

must be corrected for in order to yield reliable densities in the three-dimensional reconstruction.

Cryo-EM reconstruction of single particles differs in terms of the preparation of the biological system because there is no requirement for the sample to be crystalline (indeed monodisperse particles are ideal). Because of this, there are several key areas where the technique has clear advantages compared with protein crystallography. In sample preparations where there is a structurally heterogeneous population of the biological species, the electron microscopist is able, in principle, to group objects with the same structure. This approach was exemplified in the structure determination of the protein fold of the core of hepatitis B virus, where there is a mixed population of structurally distinct particles, with capsids displaying either $T = 3$ or $T = 4$ icosahedral triangulation. From a single focal series of micrographs (micrographs taken of the same field but with different defocus levels), one group of experimenters was able to select 1618 $T = 4$ particles and 422 $T = 3$ particles which were reconstructed to 9 and 11 Å, respectively (Conway *et al.*, 1997).

Cryo-EM also allows access to structural data on large flexible macromolecular assemblies which are likely to be difficult or impossible to crystallize. By allowing the construction of detailed models for these assemblies as structures of sub-components become available from crystallographic or NMR studies, it provides the functional context to these results. Finally, there is the possibility of looking at dynamic processes such as membrane fusion (Fuller *et al.*, 1995), viral attachment to cellular receptors (Olson *et al.*, 1993), genome extrusion and virus disassembly. These studies can take advantage of the higher symmetry present for a portion of the structure to help guide the untangling of non-symmetric events (Rossmann & Tao, 1999).

There are deep similarities between the concepts and ideas of image reconstruction from two-dimensional micrographs and crystallography. Readers with a crystallographic bent will particularly enjoy the paper by Rossmann & Tao (1999), which draws very clear analogies between the two methods.

2. Sample preparation in cryo-electron microscopy

The three-dimensional reconstruction of the density of delicate biological systems requires that the samples under investigation are fully hydrated and as close to a physiological state as possible. Electrons interact more strongly with matter than X-rays and so the (thin) sample must be maintained at high vacuum in the microscope. One approach to the maintenance of a hydrated native environment within the vacuum is to solidify the specimen in order to reduce its vapour pressure. Conventional cooling would lead to the formation of ice crystals and the consequent destruction of the biological sample. To avoid this, very rapid cooling of a thin sample is used to obtain a vitrified specimen. Typically, samples are rapidly cooled by plunging them into ethane slush cooled by liquid nitrogen (Adrian *et al.*, 1984). Maintaining the specimen at liquid-nitrogen temperatures increases its longevity in the beam and allows the recording of high-resolution images.

Unfortunately, there is little difference in the mean scattering cross-section of water and biological material, and hence there is poor amplitude contrast in the micrographs. This is made worse by the need to take exposures at low electron doses to minimize radiation damage. Attempts to solve this contrast problem introduce further difficulties, as we will discuss below. Nevertheless, the developments of cryo-sample preparation (Dubochet *et al.*, 1988) have supported the parallel technical and computational improvements which have enabled image reconstruction to make such advances over the past 12 years (Henderson, 1995; Mancini *et al.*, 1997; Zemlin, 1994).

3. The interaction of electrons and X-rays with matter: scattering and image formation

The nature of the interaction of electrons and X-rays with biological matter is not the same (Henderson, 1995). The energy deposited by both inelastic and elastic collisions is larger for X-rays than electrons, so that radiation damage per scattering event is worse. In X-ray crystallography, this is overcome by a massive three-dimensional crystal (containing perhaps 10^{12} molecules) which is usually cooled to just above the boiling point of nitrogen. Cryo-cooling in protein crystallography enables far better quality data to be collected from a single crystal. Cooling and improved X-ray detectors have also allowed data to be collected from much smaller crystals than previously feasible. In contrast, in single-particle EM methods each particle used is considered individually and its position and orientation (five degrees of freedom in total) must be precisely defined. It can then be combined with data from many other particles, to perform *in silico* the averaging which comes naturally from the regularity of a crystal. Extracting information from individual particles involves a trade-off between improving the signal-to-noise ratio and preserving the specimen. The first requires longer exposure times, which obliterates the high-resolution detail. The second requires shorter exposure time in order to reduce radiation damage and results in a low-contrast image for which the determination of particle orientations and positions is more difficult. Because of the lack of amplitude contrast as discussed above, most of the contrast in the micrographs arises from phase contrast achieved by slightly defocusing the image. Finally, the strong interaction of electrons with matter causes multiple scattering, which severely limits the thickness of samples that can be used.

Improvements in the coherence of the electron beam are a major factor in the increase in resolution and quality of the images obtained by EM (Mancini *et al.*, 1997; Chiu & Schmid, 1997). Field-emission gun (FEG) microscopes and the use of higher voltages have enabled data to be collected with a better signal-to-noise ratio for the high-resolution information (Zemlin, 1994). Although the use of FEG microscopes, run at higher voltages, also reduces the number of multiple scattering events, the sample thickness is still limited to perhaps 2000 Å. Another issue is that the level of defocus will change through the sample, which limits the sample size. The use of higher

voltage electrons ameliorates this problem because the depth of field is increased.

In contrast, protein crystallography is a diffraction technique which does not rely on image formation by focusing and so the sample thickness is not a problem. In fact, within limits, the larger the crystal volume the better. The curvature of the Ewald sphere, which invalidates the assumption that the EM image corresponds to a plane central section through the Fourier transform, is explicitly dealt with in crystallography.

4. Particle reconstruction

The concept behind three-dimensional image reconstruction from cryo-electron micrographs is essentially the same as for image reconstruction from crystallographic diffraction; namely, that a sufficient set of different views are needed of an

object to derive a three-dimensional image of it (Crowther, 1971). In a crystal, the lattice imposes constraints on the information which can be obtained about the molecular transform, namely the various harmonics of the lattice spacings which constitute the reciprocal lattice. Rotation of the crystal then allows a full set of components to be gathered. In contrast, in cryo-EM the two-dimensional micrographs give, for each particle, a projection of one particular view of the object (corresponding to a central section in reciprocal space). EM image reconstruction then relies on filling in three-dimensional reciprocal space with different central sections corresponding to different views of the object. Once there is a sufficient sampling of reciprocal space, an electron-density map of the object is calculated by inverse Fourier transformation.

The determination of the orientation and position of the individual projected images from the two-dimensional micrographs is the most important step in this process and the one fraught with the most difficulty. The orientation of the particle will naturally define the direction of the central section observed, whereas the two degrees of translational freedom (within the plane of the micrograph) introduce a relative phase shift.

The problem in defining these five parameters arises from the poor signal-to-noise ratio in low-dose images. The earliest procedures for determining object orientation were reciprocal-space methods (Crowther, 1971). Although these methods were originally designed for dealing with particles which have internal symmetry, they have been extended to application with non-symmetric objects. The reciprocal-space methods are based upon the 'common-lines' approach (reviewed in Fuller *et al.*, 1996). These use, as a basis for determining the orientation of a particle, the knowledge that there are sets of lines in the central sections of the Fourier transform of a symmetric object related by this symmetry, along which the Fourier coefficients should be identical (see Fig. 1). In practice, this is never exactly true owing to the high noise levels inherent in the images. Particles with icosahedral symmetry, such as certain viruses, will have 37 pairs of such common lines whose directions, once detected, define the orientation of the set of icosahedral symmetry axes. Similarly, 60 pairs of cross common lines are defined by the relative

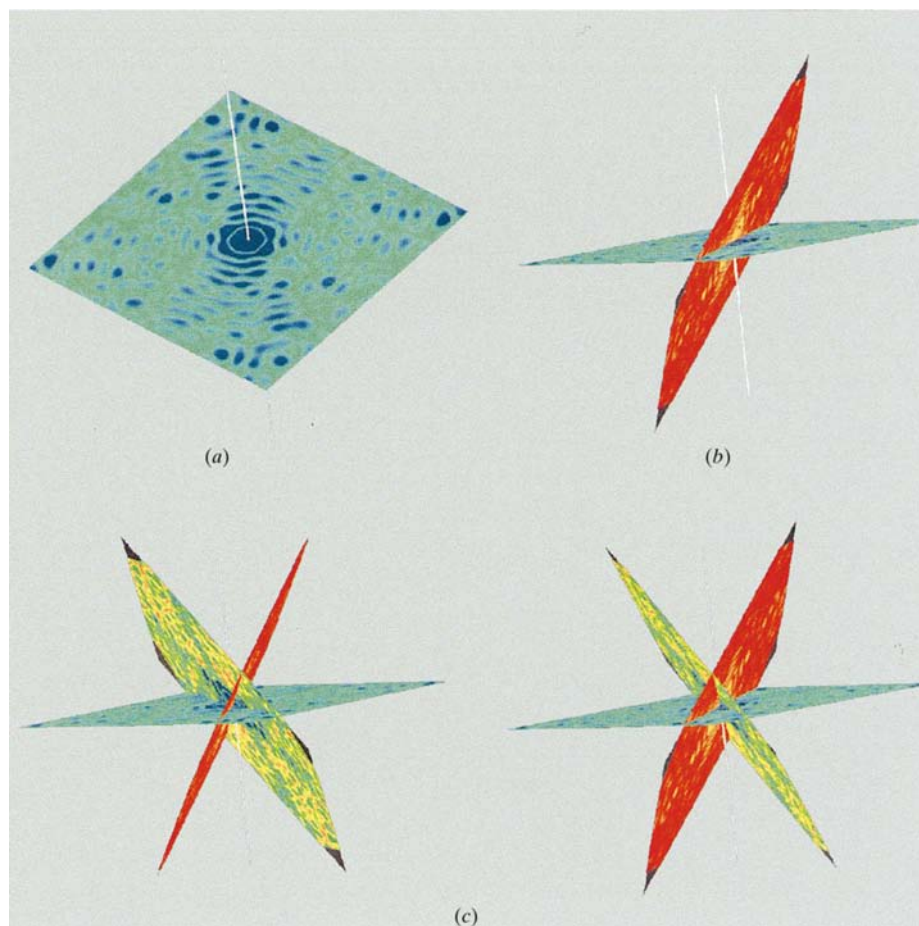


Figure 1

The origin of common lines in the transform of an icosahedrally symmetric object. The common lines arise from the application of a symmetry axis which is not coincident with the direction of view ($\theta = 89^\circ$, $\varphi = -1^\circ$, shown by the vertical white line). It is illustrated in the figure for a threefold axis ($\theta = 69.19^\circ$, $\varphi = 0^\circ$). (a) The transform of the projection is a central section through Fourier space. (b) The application of a threefold symmetry axis generates a new plane from the original. Application of the inverse generates a third plane. The intersections of these two symmetry-related planes with the original yield a pair of lines in the original image transform which must have identical values. These are seen in the stereo pair (c). The symmetry axis which created the pair of lines lies at the centre of the surface formed by the three planes. Each pair of symmetry elements and its inverse yield a pair of common lines. For an icosahedral object, this yields 37 pairs (12 from the fivefolds, 10 from the threefolds and 15 from the twofolds) in the transform of the projection.

orientations of two projections. This approach is computationally extremely elegant and indeed is the only approach which makes no initial assumption about the particle structure except its symmetry. Relating the transforms of different particles allows checking and refinement of the orientations. A corresponding approach has been implemented in real space using sinograms (van Heel, 1987).

Once an initial set of orientations has been established, a model can be calculated. This structure has a relatively high signal-to-noise ratio and allows the use of a model-based approach to orientation. Two implementations of this model-based orientation are commonly used. The first uses cross-common lines between the image transforms and transforms of a few projections of the model (Crowther *et al.*, 1994; Mancini *et al.*, 1997). The second utilizes a polar Fourier transform (pft) and is based upon classifying the original images against projections calculated in all possible directions from the model (Baker & Cheng, 1996). In both cases, this allows an improved model to be calculated and the process is iterated to convergence.

When combining data from several images, a common origin needs to be estimated. This origin, in the two dimensions within the plane of focus of the image only affects the phase coefficients of the transform of each image, as pointed out by Rossmann & Tao (1999). A shift in the position of the origin by $(\Delta u, \Delta v)$ simply changes the phases of two-dimensional Fourier coefficients by $2\pi(h\Delta u + k\Delta v)$, where h and k are the indices for a pseudo-lattice generated by placing a suitable unit cell around the particle. The problem of particle position is therefore easy to solve in principle. Difficulties arise where the orientation determination itself is dependent on the choice of origin, as will be the case with a virus containing proper symmetry elements which are used in the process of determining the orientation. This is embedded in both the common-lines methods and the pft methods, which require interpolation of the image transform into polar coordinates. As a result, an improperly centered image may fail to orient properly. Typically, image orientation and centering are refined iteratively to ameliorate the problem.

The quality of the final density of the reconstruction is ultimately dependent on the accuracy of the estimation of the orientation and position of the particle and the homogeneity of the preparation. For example, the maximum angular error allowed for a particle of diameter 500 Å with a maximum phase error of 36° at a resolution of 5 Å would be 0.2°, and this angular error will be linearly proportional to the particle size. The positional error is independent of particle size, but for a reconstruction to 5 Å would need to be no more than about 0.5 Å. Since these two sources of error are additive, these values would presumably have to be reduced by a factor of $2^{1/2}$ in practice.

There are great similarities between methods of deriving orientations from two-dimensional micrographs and crystallographic functions which are used to derive protein orientations in single crystals of macromolecules. Rossmann & Tao (1999) make the clear analogy between 'common lines' and the 'self-rotation function' in crystallography (Rossmann &

Blow, 1962), where the orientation of a symmetrical object is derived from the symmetry which is inherent in the structure factors. They also draw the further close analogy between the pft method, where comparisons are made against an homologous model, and the cross-rotation function (Rossmann & Blow, 1962), which searches orientation space with a known model and cross-correlates it with the X-ray data. In crystallography, however, the orientation is decoupled from the position of the molecule in most methods by discarding the phase information and working in Patterson space. We are not aware of such a method being applied in electron microscopy.

Crowther discussed the number of images required to calculate a reconstruction to a given resolution in his papers which introduced three-dimensional reconstruction (Crowther, 1971). The data from cryo-electron microscopy and experience with high-resolution structures from electron crystallography have prompted the revisiting of this issue. Henderson (1995) and Rossmann & Tao (1999) have considered the issue of the number of images required for a successful reconstruction at a particular resolution. Henderson (1995) came to the conclusion that the number of images needed to determine the atomic structure (in the absence of noise) is at least 10000, but is independent of the particle size. He points out that the quality of the data will be such that many more will be needed to produce useful maps. Rossmann & Tao (1999) disagree with the view that the number of images required is independent of particle size. It is clear that the amount of information needed to solve a particle of radius R is proportional to R^3 . However, the number of pixels of information in a two-dimensional micrograph only increases as R^2 . This argument, and that of Rossmann & Tao (1999), both suggest that the number of particles to be imaged to achieve a given resolution will increase with the diameter of the particle.

5. The contrast-transfer function

Micrographs are deliberately taken slightly away from focus, which introduces added contrast owing to relative phase shifts in the scattered electrons. Spherical aberration in the lens of the microscopes introduces further phase shifts. The net effect is that the resulting image corresponds to the convolution of the 'true' structure with a contrast-enhancing function; the Fourier transform of which is referred to as the contrast-transfer function (CTF).

In reciprocal space, the central section of the molecular transform is multiplied by a circularly symmetrical CTF which, as can be seen from Fig. 2, affects both the phases and the amplitudes of the Fourier spectrum. The phases are inverted in alternate resolution bands (the positions of these bands are dependent on the defocus level for the micrograph and the wavelength of the electrons). The amplitudes not only pass through zero at the positions where the phases invert and are weak at low angle, but are also dramatically attenuated at high scattering angle. Previously, these effects had been ignored and data simply discarded beyond the first minimum in the CTF. Typically, the relatively poor coherence of a conventional

electron source causes the amplitudes to be suppressed smoothly near the edge so that truncation ripples are very limited. Nevertheless, the sizes and relative densities of

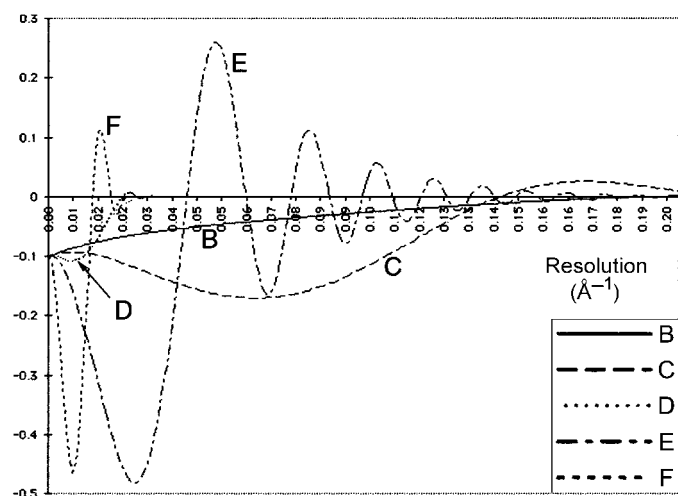


Figure 2 Contrast transfer functions are shown for five imaging conditions. All functions are shown for a microscope operating at 200 kV ($\lambda = 0.0253 \text{ \AA}$), a spherical aberration (C_s) of 2 mm and an amplitude contrast ratio of 0.1. The curves shown represent: B, an underfocus of 200 \AA and a source size of 0.01 \AA^{-1} ; C, an underfocus of 2000 \AA and a source size of 0.01 \AA^{-1} ; D, an underfocus of 20000 \AA and a source size of 0.05 \AA^{-1} ; E, an underfocus of 20000 \AA and a source size of 0.005 \AA^{-1} ; F, an underfocus of 20000 \AA and a source size of 0.005 \AA^{-1} . The coherence of the source decreases as the source size increases. The use of underfocus results in the amplitude contrast and the first peak of the phase contrast having the same sign.

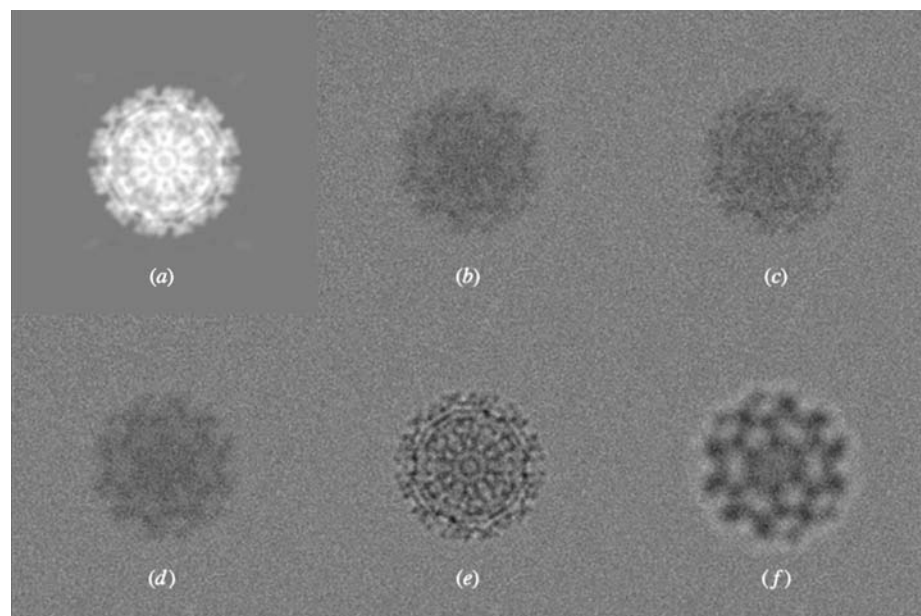


Figure 3 The effect of the contrast-transfer function on the projection of an enveloped virus is shown for Semliki Forest virus (Fuller *et al.*, 1995). (a) shows the projection of the three-dimensional density of the 700 \AA diameter virion. Note the lipid bilayer separating the internal capsid from the external spikes. (b)–(f) show the application of the transfer functions displayed in Fig. 2 to the projected density in (a). Gaussian-distributed noise with σ equal to half the maximum density in (a) was added to each image to simulate the effect of the low signal-to-noise ratio in a low-dose image. Note that the use of underfocus and amplitude contrast change the sign of the density.

features in the reconstruction are distorted by the effects of the CTF. These effects are shown in Fig. 3 on the projection of Semliki Forest virus, a surface representation of the three-dimensional image of which is shown in Fig. 4. Even when the CTF is corrected for, there are still data missing near the zeros in the CTF. This problem can be overcome by combining multiple micrographs at different defocus which have different CTFs, correcting for phase reversal and signal attenuation (methods differ since this approach is still in its infancy; Böttcher *et al.*, 1997; Zlotnick *et al.*, 1996; Böttcher & Crowther, 1996). Some recent reconstructions have used radial density profiles derived from X-ray results to facilitate making a precise correction of the CTF (Thumen-Commike *et al.*, 1999). Others have started to develop procedures to deconvolute the CTF using maximum-entropy methods (which have already had a profound impact in the field of macromolecular crystallography; Skogland *et al.*, 1996).

A simple method for compensating for various effects, including attenuation of high-resolution data because of the CTF, is by applying a negative B factor. In crystallography, this has been used in the artificial sharpening of the higher resolution features in electron-density maps (Grimes *et al.*, 1998; Stehle *et al.*, 1996). In a structure determination for hepatitis B core, an inverse B factor of 500 \AA was applied to the data used in the synthesis of the maps to determine the fold of the core protein (Böttcher *et al.*, 1997). This implicit attenuation of data would produce a reduction in the high-angle Fourier coefficients at a resolution of 5 \AA of more than a factor of 100, indicating the scale of the problem, even in the best analyses.

It is perhaps noteworthy that although cryo-EM is thought of as a method for obtaining good low-resolution data on macromolecular complexes, crystallography, given proper beam collimation and good backstop design, will allow very easily data collection to lower resolution than 100 \AA (see, for instance, Gouet *et al.*, 1999), whereas the CTF attenuates and inelastic scatter distorts such data in cryo-EM (see Fig. 2).

6. Defining criteria for correctness

With the increase in the numbers and the resolution of structures imaged by cryo-EM, there has been a realisation amongst the EM community that discerning criteria should be established to demonstrate the validity and accuracy of EM reconstructions. We believe that more use should be made of reciprocal-space comparisons of amplitude and phase, which provide well defined and accepted objective measures to establish the similarity between data sets and allow the

(sometimes dramatically different) resolution fall-off of different data sets to be evaluated.

At present, the usual method of assessing the quality of an EM reconstruction is to divide the data set into two and perform independent reconstructions. The quality of the reconstruction is monitored by assessing the similarity between phases derived from these two reconstructions. Once the number of images used in the reconstruction and the orientation and position of each particle have been defined by a cyclic refinement and rejection procedure, the two data sets are combined to produce the final reconstruction. This division into sets is a rather poor use of data. In contrast, we suggest that all, or nearly all, data should be used simultaneously. *R* factors, correlation coefficients and phase residuals should be given for structure factors derived from back-transformed maps and the Fourier coefficients derived from the CTF-corrected two-dimensional micrographs as objective criteria for assessing the quality of a reconstruction. This use of the Fourier components, calculated from the two-dimensional micrographs used, is the approach used by Rossmann & Tao (1999) in their paper on discriminating the correct solutions for the orientation of particles. Such a close analogy with crystallographic methods also immediately invites the use of the equivalent of the free *R* factor in crystallography, although in this case a vector *R* factor would be more appropriate (see Rossmann & Tao, 1999). However, initial attempts to use the free *R* factor in icosahedral reconstruction methods as a measure of the improvement in orientations during refinement have not been very successful (Fuller, unpublished data).

As an example of the use of reciprocal-space measures, in work on cytoplasmic polyhedrosis virus simple but illuminating common crystallographic measures were used to compare reconstructions of different particles (Hill *et al.*,

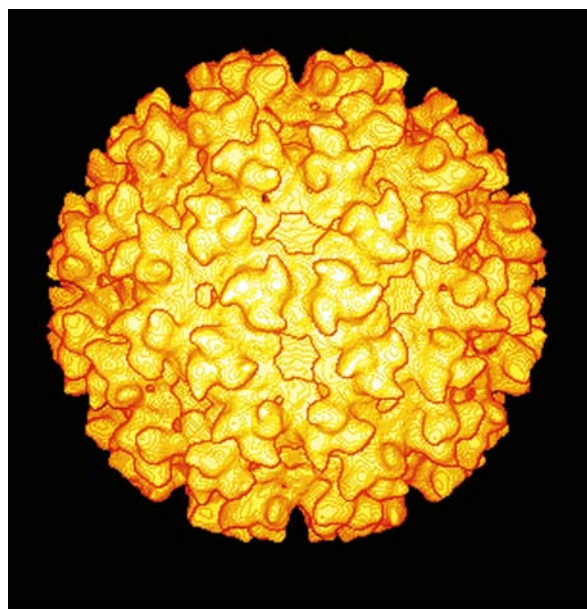


Figure 4

The surface representation of Semliki Forest virus, which is projected to yield the density shown in Fig. 3.

1999). CPV can be purified from viral inclusion bodies ('occluded' virions) or from the cell cytoplasm ('non-occluded' virions). Back-transformation of the two reconstructions allowed the scaling of the data sets using a scale factor and a *B* factor. The reconstruction for 'occluded' virions had a *B* factor of 3750 Å² compared with 'non-occluded' particles, giving an objective measure of the comparative data fall-off with resolution, and this clearly correlates with the relative sharpness of the electron-density maps. The *R* factor and phase difference between the two particles was 50% and 57° on all the data to 25 Å (65% and 69° between 26.7 and 25.0 Å), reflecting perhaps different distortions introduced by the CTF effect.

7. Phasing crystallographic structures.

The improved resolution and reliability of the three-dimensional reconstructions from cryo-EM data has seen an increasing use of cryo-EM data as initial phasing models for crystallographic problems. Three case studies will be given where envelopes of EM reconstructions have been used successfully as molecular-replacement models, two of them in combination with atomic models which allowed crystallographic phasing out to high resolution.

The DNA bacteriophage ΦX174 uses a scaffold protein in forming an assembly intermediate, known as a procapsid, during capsid formation. The mature ΦX174 virion contains 60 copies of the capsid, spike and DNA-associated protein (F, G and J, respectively). The procapsid has a further two scaffolding proteins, 240 copies of an external protein D and 60 copies of an internal protein B, which are both required for assembly. The structure of the 'closed' procapsid of ΦX174 was solved inadvertently from data collected from crystals grown from initially 'open' procapsids. During the crystallization, a maturation process took place mimicking the structural changes associated with DNA packaging. The molecular-replacement model was based upon the atomic model of the structure of the mature ΦX174 capsid, extensively modified and manually fitted into a 25 Å cryo-EM map of the open procapsid. In addition, for those parts corresponding to the scaffolding protein D, the spike protein G was used as a model since it has approximately the correct size. This phasing model proved to be very poor, since the crystals contained the closed procapsid, where the orientations and positions of F and G are very similar to those in mature virions. In addition, the reconstruction had the incorrect hand, which led to the F protein being rotated by about 90° from its correct orientation. To further compound the problems, only 120 copies rather than 240 copies of the D protein were modelled and the D protein has an α-helical fold compared with the G protein's essentially β-sheet structure (Dokland *et al.*, 1998). Nevertheless, this model sufficed as a starting point for real-space phase refinement and extension using non-crystallographic averaging which eventually led to an atomic model refined using data to 3.5 Å resolution (Dokland *et al.*, 1997). In retrospect, success probably owed more to the power of current methods for phase refinement than to the correctness of the starting phases.

The determination of the structure of the core particle of blue-tongue virus (BTV) used a starting model for the outer layer of the core particle based on the fitting of 260 trimers of the atomic model of VP7 into a nominally 24 Å resolution cryo-EM reconstruction (Grimes *et al.*, 1997). This combination of data provided phasing information to substantially higher resolution than that of the cryo-EM reconstruction alone (as judged by correlation coefficients). Unlike the structure solution for ΦX174 (where phases were extended and refined using real-space averaging techniques), for BTV the phases were extended to 12 Å and then 6 Å by rigid-body refinement of the model trimers with application of strict non-crystallographic constraints in reciprocal space. Real-space averaging of electron-density maps calculated at 6 Å resolution then produced excellent maps, allowing secondary structural elements for the capsid protein not used in the phasing model to be positioned. This hybrid procedure was then repeated, refining this crude model against all the data to 3.5 Å, giving phases at 3.5 Å which were refined by real-space averaging to yield a readily interpretable map (Grimes *et al.*, 1998).

The progress on the crystallography of the ribosome has also been greatly aided by reference to the cryo-EM reconstruction (Frank *et al.*, 1995). An envelope based on the cryo-EM reconstruction was successfully used as a molecular-replacement model to give the orientation and position of the ribosome within the crystal lattice (Ban *et al.*, 1998). The phases derived from the envelope alone were then sufficient to cross-phase derivatives using difference Fourier techniques. This use of difference Fourier maps also validated the molecular-replacement solution. This shows the utility of EM-based methods even in the absence of non-crystallographic symmetry.

8. Prospects for the future

Cryo-EM image reconstruction has already had a major impact on structural biology, in terms of our understanding of virus structure complexed with antibodies and receptors and the mechanisms of capsid assembly/disassembly (Olson *et al.*, 1993; Smith *et al.*, 1993). In addition, single-particle image reconstruction has allowed a first structural glimpse of low-symmetry macromolecular complexes such as the ribosome and GroES complexed with substrates (Chen *et al.*, 1994; Frank *et al.*, 1995; Stark *et al.*, 1995). When there are atomic models for protein subcomponents of these macromolecular complexes, this allows a 'divide-and-conquer' approach to generate atomic models for these complexes by quantitative fitting procedures of the models into the cryo-EM envelopes (Gilbert *et al.*, 1999; Grimes *et al.*, 1997).

In addition to providing reasonably accurate models for large flexible macromolecular assemblies, these cryo-EM reconstructions can also be used to phase crystallographic data, as the resolution of the cryo-EM data increases to provide more overlap with the X-ray data. With large amounts of structural redundancy (as occurs in virus crystals), the phasing can be extended to well beyond the resolution of the

cryo-EM source by application of non-crystallographic constraints in reciprocal-space refinement if there is atomic model information available or by non-crystallographic real-space averaging (Grimes *et al.*, 1998).

Even if there is no symmetry present and atomic models are not available, the envelopes derived from cryo-EM maps provide useful low-resolution phases (~15 Å) which can confirm heavy-atom positions in difference Fourier maps, as exemplified by the crystallographic analysis of the ribosome (Ban *et al.*, 1998). In addition, it is fairly common that crystals of large macromolecular complexes such as the ribosome suffer from a lack of isomorphism on cryo-cooling. This apparent problem could be exploited to provide a vehicle for phasing such complexes, because of the differential sampling of the molecular transform which occurs on changes in the unit cell (Esnouf *et al.*, 1998; Ren *et al.*, 1995). Thus, EM reconstructions could provide starting models for each of the crystal forms and would establish initial values for the rotation matrices and translation vectors which relate the molecules in the different forms prior to non-crystallographic averaging between them.

Beyond the use of image reconstruction from cryo-EM micrographs as simply a starting point for phasing low-resolution data, the techniques also allow dynamical problems such as virus fusion and disassembly to be addressed and structures where there is not a structurally homogeneous population (Fuller *et al.*, 1995). In addition, it is becoming increasingly clear that icosahedral viruses have lower symmetry components. It may be possible to use those parts of the capsid which have icosahedral symmetry to initially orient the particles (Rossmann & Tao, 1999). There is then the possibility that these orientations could then be used as a starting point for pulling out lower symmetry structural components by further alignment based on the lower symmetry. The future holds enormous promise.

References

- Adrian, M., Dubochet, J., Lepault, J. & McDowell, A. W. (1984). *Nature (London)*, **308**, 32–36.
- Baker, T. S. & Cheng, R. H. (1996). *J. Struct. Biol.* **116**, 120–130.
- Baker, T. S., Olson, N. H. & Fuller, S. D. (1999). In the press.
- Ban, N., Freyborn, B., Nissen, P., Penczek, P., Grassucci, R. A., Sweet, R., Frank, J., Moore, P. B. & Steitz, T. A. (1998). *Cell*, **93**, 1105–1115.
- Böttcher, B. & Crowther, R. A. (1996). *Structure*, **4**, 387–394.
- Böttcher, B., Wynne, S. A. & Crowther, R. A. (1997). *Nature (London)*, **386**, 88–91.
- Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., Clarke, A. R. & Saibil, H. R. (1994). *Nature (London)*, **371**, 261–264.
- Chiu, W. & Schmid, M. F. (1997). *Nature Struct. Biol.* **4**, 331–333.
- Conway, J. F., Cheng, N., Zlotnick, A., Wingfield, P. T., Stahl, S. J. & Steven, A. C. (1997). *Nature (London)*, **386**, 91–94.
- Crowther, R. A. (1971). *Philos. Trans. R. Soc. London Ser. B*, **261**, 221–230.
- Crowther, R. A., Kiselev, N. A., Böttcher, B., Berrimen, J. D., Borisova, G. P., Ose, V. & Pumpens, P. (1994). *Cell*, **77**, 943–950.
- Dokland, T., McKenna, R., Ilag, L. L., Bowman, B. R., Incardona, N. L., Fane, B. A. & Rossmann, M. G. (1997). *Nature (London)*, **389**, 308–313.

- Dokland, T., McKenna, R., Sherman, D. M., Bownman, B. R., Bean, W. F. & Rossmann, M. G. (1998). *Acta Cryst. D* **54**, 878–890.
- Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowell, A. W. & Schultz, P. (1988). *Quart. Rev. Biophys.* **21**, 129–228.
- Esnouf, R. M., Ren, J., Garman, E. F., Somers, D. O'N., Ross, C. K., Jones, E. Y., Stammers, D. K. & Stuart, D. I. (1998). *Acta Cryst. D* **54**, 938–954.
- Frank, J., Zhu, J., Penczek, P., Y., L., Srivastava, S., Verschoor, A., Radermacher, M., Grassucci, R., Lata, R. K. & Agrawal, R. K. (1995). *Nature (London)*, **276**, 441–444.
- Fuller, S., Butcher, S. J., Cheng, R. H. & Baker, T. S. (1996). *J. Struct. Biol.* **116**, 48–55.
- Fuller, S. D., Berrimean, J. A., Butcher, S. J. & Gowen, B. E. (1995). *Cell*, **81**, 715–725.
- Gilbert, R. J. C., Jiminez, J. L., Chen, S., Tickle, I. J., Rossjohn, J., Parker, M., Andrew, P. W. & Saibil, H. R. (1999). *Cell*, **97**, 657–666.
- Gouet, P., Diprose, J. M., Grimes, J. M., Malby, R., Burroughs, J. N., Zientara, S., Stuart, D. I. & Mertens, P. P. C. (1999). *Cell*, **97**, 481–490.
- Grimes, J. M., Burroughs, J. N., Gouet, P., Diprose, J. M., Malby, R., Mertens, P. P. C. & Stuart, D. I. (1998). *Nature (London)*, **395**, 470–478.
- Grimes, J. M., Jakana, J., Ghosh, M., Basak, A. K., Roy, P., Chiu, W., Stuart, D. I. & Prasad, B. V. V. (1997). *Structure*, **5**, 885–893.
- Heel, M. van (1987). *Ultramicroscopy*, **21**, 95–100.
- Henderson, R. (1995). *Quart. Rev. Biophys.* **28**, 171–193.
- Hill, C. L., Booth, T. F., Prasad, B. V. V., Grimes, J. M., Mertens, P. P. C., Sutton, G. C. & Stuart, D. I. (1999). *Nature Struct. Biol.* **6**, 565–568.
- Mancini, E. J., de Haas, F. & Fuller, S. D. (1997). *Structure*, **5**, 741–750.
- Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S. & Rossmann, M. G. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 507–11.
- Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C., Kirby, I., Keeling, J., Darby, G., Jones, Y. & Stuart, D. (1995). *Nature Struct. Biol.* **2**, 293–302.
- Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
- Rossmann, M. G. & Tao, Y. (1999). *J. Struct. Biol.* **125**, 196–208.
- Skogland, U., Ofverstedt, L. G., Burnett, R. M. & Bricogne, G. (1996). *J. Struct. Biol.* **117**, 173–188.
- Smith, T. J., Olson, N. H., Cheng, R. H., Chase, E. S. & Baker, T. S. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 7015–7018.
- Stark, H., Mueller, F., Orlova, E. V., Schatz, M., Dube, P., Erdemir, T., Zemlin, F., Brimacombe, R. & van Heel, M. (1995). *Structure*, **3**, 815–821.
- Stehle, T., Gamblin, S. J., Yan, Y. & Harrison, S. C. (1996). *Structure*, **4**, 165–182.
- Thuman-Commike, P. A., Tsuruta, H., Greene, B., Preveliege, P. E. Jr, King, J. & Chiu, W. (1999). *Biophys. J.* **76**(4), 2249–2261.
- Zemlin, F. (1994). *Micron*, **25**, 223–226.
- Zlotnick, A., Cheng, N., Conway, J. F., Booy, F. P., Steven, A. C., Stahl, S. J. & Wingfield, P. T. (1996). *Biochemistry*, **35**, 7412–7421.