

# Low-Resolution Density Maps from Atomic Models: How Stepping “Back” Can Be a Step “Forward”

David M. Belnap,<sup>1,2</sup> Abhinav Kumar,<sup>1,3</sup> Jon T. Folk, Thomas J. Smith, and Timothy S. Baker<sup>4</sup>

*Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907-1392*

Received November 14, 1998, and in revised form January 25, 1999

**Atomic-resolution structures have had a tremendous impact on modern biological science. Much useful information also has been gleaned by merging and correlating atomic-resolution structural details with lower-resolution (15–40 Å), three-dimensional (3D) reconstructions computed from images recorded with cryo-transmission electron microscopy (cryoTEM) procedures. One way to merge these structures involves reducing the resolution of an atomic model to a level comparable to a cryoTEM reconstruction. A low-resolution density map can be derived from an atomic-resolution structure by retrieving a set of atomic coordinates, editing the coordinate file, computing structure factors from the model coordinates, and computing the inverse Fourier transform of the structure factors. This method is a useful tool for structural studies primarily in combination with 3D cryoTEM reconstructions. It has been used to assess the quality of 3D reconstructions, to determine corrections for the phase-contrast transfer function of the transmission electron microscope, to calibrate the dimensions and handedness of 3D reconstructions, to produce difference maps, to model features in macromolecules or macromolecular complexes, and to generate models to initiate model-based determination of particle orientation and origin parameters for 3D reconstruction.** © 1999 Academic Press

**Key Words:** correlation of cryoelectron microscopy and X-ray crystallography; cryoelectron microscopy; modeling; scanning transmission electron microscopy; three-dimensional image reconstruction; truncation of resolution; X-ray crystallography.

## INTRODUCTION

The knowledge gleaned from atomic-resolution structures of biological macromolecules has had a tremendous impact on our discovery and understanding of basic life processes. Such structural information has led to new insights about how genes replicate, viruses assemble, and muscles contract, to name but a few examples. X-ray crystallography and nuclear magnetic resonance (NMR)<sup>5</sup> are the classic tools by which atomic structures are obtained. In a few favorable cases, two-dimensional crystalline samples have been analyzed by electron crystallographic techniques at atomic or near atomic detail in three dimensions (e.g., Henderson *et al.*, 1990; Kühlbrandt *et al.*, 1994; Nogales *et al.*, 1998). With few exceptions, three-dimensional (3D) structures of non-crystalline, randomly oriented particles determined by cryo-transmission electron microscopy (cryoTEM) and 3D image reconstruction have been determined at the resolution range of 15–30 Å. The pioneering work with hepatitis B virus (Böttcher *et al.*, 1997; Conway *et al.*, 1997) and bovine papillomavirus (Trus *et al.*, 1997), in which structural details beyond 10-Å resolution were obtained, has provided a key stimulus to achieve even higher, perhaps atomic, resolution with a wider variety of biological macromolecules using state-of-the-art specimen preparation, transmission electron microscopy, and 3D image reconstruction techniques.

Notwithstanding the importance of obtaining the highest resolution information possible, much relevant knowledge about the structure and function of biological macromolecules has been learned from studies in which complementary information from

<sup>1</sup> These authors contributed equally to this project.

<sup>2</sup> Current address: Building 6, Room B2-34, 6 Center Drive MSC 2717, National Institutes of Health, Bethesda, MD 20892-2717.

<sup>3</sup> Current address: Biomolecular Structure Center, University of Washington, Health Sciences Center, K-428, Box 357742, Seattle, WA 98195-7742.

<sup>4</sup> To whom correspondence should be addressed.

<sup>5</sup> Abbreviations used: 2D, two-dimensional; 3D, three-dimensional; CCMV, cowpea chlorotic mottle virus; CPMV, cowpea mosaic virus; CPV, canine parvovirus; cryoTEM, cryo-transmission electron microscopy; CTF, contrast transfer function; Fab, fragment antibody-binding; FHV, Flock House virus; FTP, file transfer protocol; HRV, human rhinovirus; NMR, nuclear magnetic resonance; PDB, Protein Data Bank.

both low and high resolution have been correlated (e.g., Baker and Johnson, 1996). For example, a “pseudo-atomic” model of the expanded (swollen) form of the plant virus, cowpea chlorotic mottle virus (CCMV), was obtained by fitting the atomic structure of the viral capsid protein subunits, as observed in crystals of the native (unexpanded) viral capsid, into a 28-Å resolution density map of the expanded CCMV as determined by cryoTEM and 3D reconstruction (Speir *et al.*, 1995). The resultant model showed how electrostatic repulsion could drive the expansion of the particle as the pH and metal-ion concentration changed. Of course, this and other similar examples clearly demonstrate that more precise definition and reliability of such models would accompany cryoTEM reconstructions determined at significantly higher resolutions.

Occasions arise when it is useful to examine structural information at *reduced* resolution, and in essence, step “backward” from high to low. For example, for comparative purposes, it is often beneficial to render crystallographic structures and cryoTEM reconstructions at the same resolution. These comparisons have helped inspire confidence in efforts to merge crystallographic and cryoTEM structures by showing that the two techniques produce nearly identical structures. In conjunction with cryoTEM images and reconstructions, the reduced resolution structures are a step “forward” in other ways as well by their usefulness (1) as models; (2) in calibrating size, density, handedness, position, or orientation; and (3) in making difference maps.

Here we outline a protocol for producing low-resolution density maps from atomic coordinates obtained by high-resolution analysis or from pseudo-atomic modeling. Though we have primarily used this procedure to generate density maps of icosahedral viruses, it has also been used to make a low-resolution map of a single polypeptide and could be used to make maps of any structure for which atomic coordinates are available or can be generated. Low-resolution density maps could be produced by any of a number of procedures similar to the one we describe. Indeed, we have concocted different protocols ourselves. Other groups have also developed similar procedures (e.g., Stewart *et al.*, 1991, 1993; Grimes *et al.*, 1997; Wriggers *et al.*, 1998). Our goal is to develop a user-friendly environment for producing density maps at any resolution from atomic model coordinates that can be useful in analysis of cryoTEM reconstructions. We desire to prevent others from expending great effort to “reinvent the wheel.” Unfortunately, the marriage of the mostly lab-based cryoTEM image reconstruction software with the more mature, universally accepted crystallographic software remains problematic. These problems range

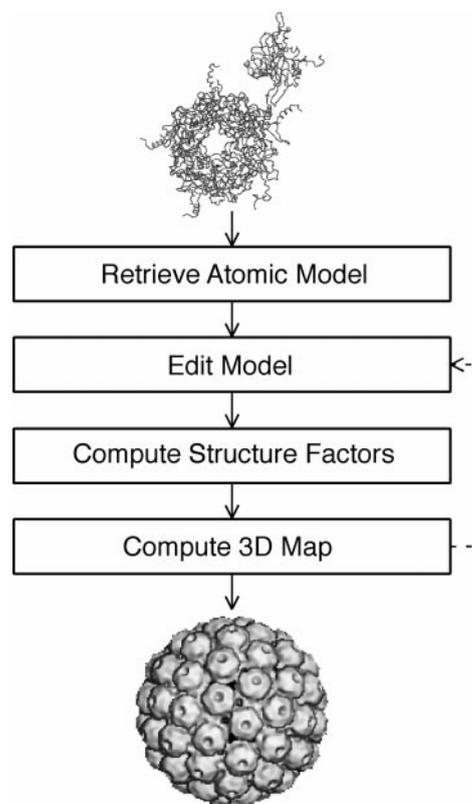
from an unfavorable mix of different computer operating systems, to difficulties in communicating between different data formats, and to the usual obstacles of training new users in the nuances of the different software. The work reported here is but a first attempt to streamline the process and make it available to the community. We anticipate that further improvements will be essential and should be incorporated as our experience grows.

## METHOD

The procedure we describe below consists of four main steps as illustrated in a flow diagram (Fig. 1). This includes (1) retrieving a high-resolution model, (2) editing the model, (3) computing structure factors, and (4) computing a density map. We developed two protocols—one uses XPLOR (Brünger, 1987) and the other CCP4 (Bailey, 1994)—for computation of the structure factors. Each protocol essentially follows the four-step outline although some substeps differ. Detailed instructions and sample command files are available at [http://bilbo.bio.purdue.edu/~baker/high\\_to\\_low/high\\_to\\_low.html](http://bilbo.bio.purdue.edu/~baker/high_to_low/high_to_low.html).

### Step 1: Retrieve High-Resolution Model

The first step is to obtain high-resolution coordinates from the Protein Data Bank (PDB) or perhaps from a magnanimous



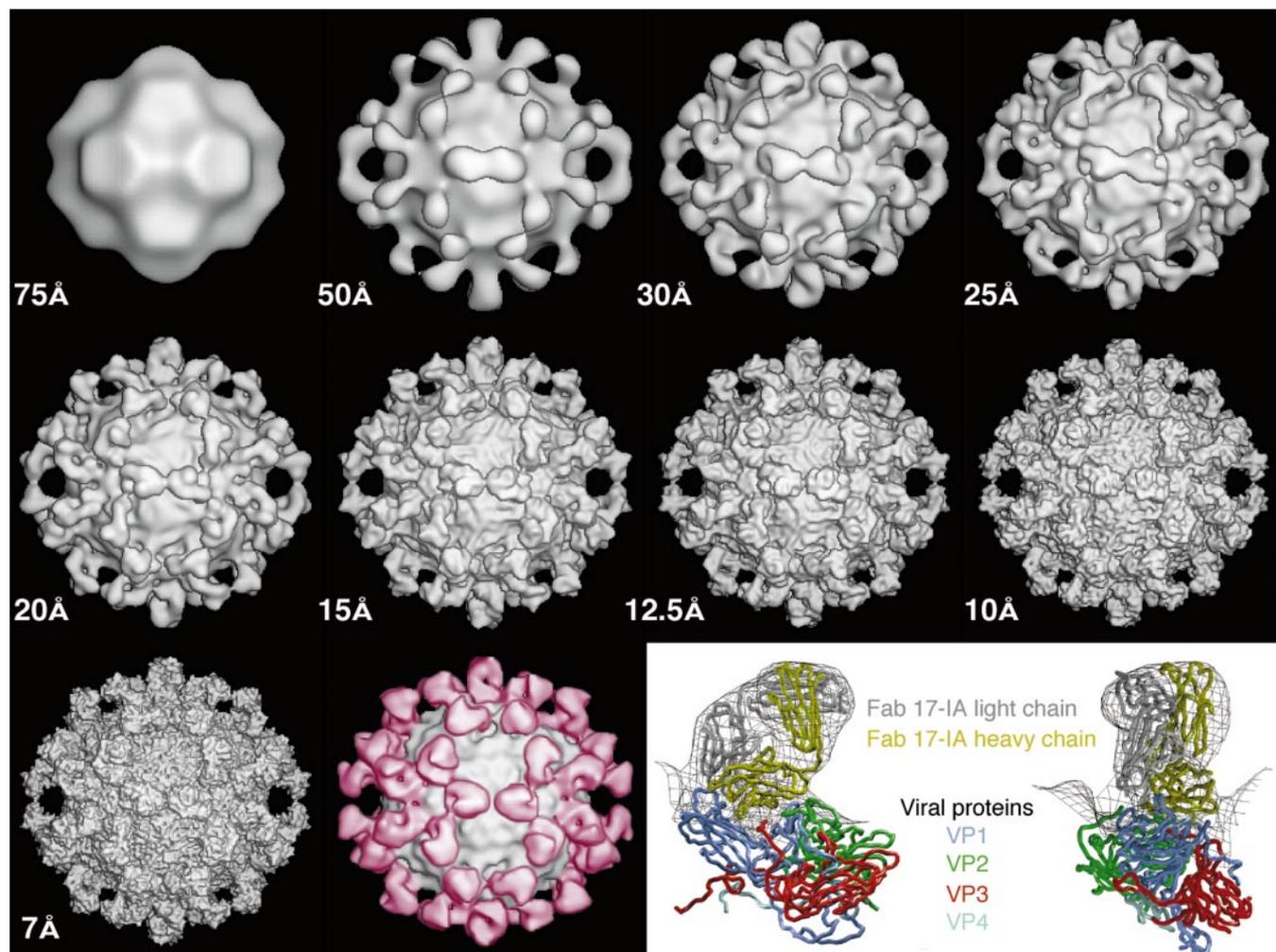
**FIG. 1.** Flow diagram outlining the steps used to compute a low-resolution density map from an atomic model. The example input and output are from the murine polyomavirus atomic structure (Stehle and Harrison, 1996). The wire diagram was prepared with the programs MOLSCRIPT (Kraulis, 1991) and Raster3D (Merritt and Bacon, 1997).

colleague. We used PDB-format text files, but other file formats could be used or converted to the PDB format. Files are generally retrieved with the use of a Web browser from the PDB Web site (<http://www.resb.org>) or with the use of a file transfer protocol (FTP) program from the PDB FTP site (<ftp.resb.org>).

Pseudo-atomic models of altered structures or macromolecular complexes can also be constructed and used to make low-resolution density maps. This involves reorienting as rigid bodies either one known structure to produce a model of an altered structure or two or more known structures to produce a model of a complex. In either case, the pseudo-atomic model is made with other available structural information such as a 3D image reconstruction. Care must be taken to avoid or minimize intermolecular clashes (i.e., overlaps). Rigid body motions should be employed as far as possible to prevent unwarranted distortions. The only exception is if rearrangements of domains within subunits are

known, for example, the bending of the Fab elbow region. Certain parts of the structure, for example, highly flexible loops, may be difficult to model with rigid body motions. In this case, it may be better to remove that particular piece of the structure and model only with the more rigid portions.

The pseudo-atomic model of swollen CCMV (Speir *et al.*, 1995) discussed above (see Introduction) is an example of an altered structure. A model of a complex was generated with human rhinovirus (HRV) and a Fab fragment from a neutralizing monoclonal antibody (Fig. 2; Smith *et al.*, 1993). In the HRV-Fab study, a 28-Å resolution cryoTEM image reconstruction provided the framework for constraining, to within a few angstroms, the fit of the atomic model of the Fab to the atomic model of the virus. Though discussion of manual or automated schemes for producing pseudo-atomic models is beyond the scope of this report, various automated protocols are discussed elsewhere (e.g., Wriggers *et al.*, 1998, 1999; Volkman and Hanein, 1999).



**FIG. 2.** Pseudo-atomic model of human rhinovirus (HRV) serotype 14 complexed with monoclonal antibody fragment Fab 17-IA (Smith *et al.*, 1996), displayed as shaded, solid-surface grayscale representations at resolutions progressing from 75 to 7 Å. The Fabs are invisible at the lowest resolution, but structural details become better defined at progressively higher resolutions. Thus, the popular shaded-surface rendering technique is best suited to revealing gross morphological features in structures at resolutions of about 15 Å or larger. The overall structure of the virus complex becomes increasingly difficult to decipher as the resolution (i.e., level of detail) improves. The cryoTEM reconstruction of the HRV-Fab complex (density map with Fabs colored pink) at ~28-Å resolution (Smith *et al.*, 1993) compares most favorably with the pseudo-atomic models rendered at 30- and 25-Å resolution. The panels at the bottom right give orthogonal views of the HRV14-Fab17-IA pseudo-atomic model (polypeptide chains displayed as "worms") fit into the wire mesh envelope (shows just one Fab density and adjacent viral surface) of the cryoTEM reconstruction (Che *et al.*, 1998).

### Step 2: Edit Model

The atomic coordinate file generally must be edited to prepare the data for the subsequent processing steps. Several actions may be necessary, such as placing, centering, and orienting the structure in a pseudo-unit cell; editing the coordinate file to remove extraneous text such as header information; and applying point-group symmetry. Also, it may be desirable to render the model simply as  $C_\alpha$  atoms or to include all atoms. (Though in general it is advisable to use all nonhydrogen atoms, for resolutions of 25 Å or lower, whether all atoms or just the  $C_\alpha$  atoms are used makes little difference. Computer memory limitations may necessitate use of only  $C_\alpha$  atoms.) We used the FORTRAN program DEALPDB (A. Kumar, unpublished; available at our Web site), simple text editing, or some in-house programs (available at our Web site) to perform the centering, rotating, and editing operations. Other software programs available for manipulating PDB files are MOLEMAN, which is part of the O program suite (Jones *et al.*, 1991); PDBSET, which is part of the CCP4 suite (Bailey, 1994); and EDPDB (Zhang and Matthews, 1995). A simple text editor also can be very useful in manipulating the coordinate file. Editing the coordinate file involves a trial-and-error process and, hence, may need to be repeated for multiple cycles of our procedure (Steps 2–4; Fig. 1) to achieve desired results.

In our protocol, the model needs to be positioned so that its center of mass coincides with the center of the unit cell (i.e., at fractional coordinates  $\frac{1}{2}$ ,  $\frac{1}{2}$ ,  $\frac{1}{2}$ ). However, this shift is either performed before (XPLOR) or after (CCP4) the structure factor calculation. (Our CCP4 protocol requires the model to be centered at the origin (0, 0, 0) for the structure factor calculation.) Other protocols may keep the model centered at the origin. The choice of centering is dictated by the programs used to compute and visualize the final 3D map. For X-ray crystallographic structures, the new unit cell can be different from what it was in the true crystal. A triclinic cell of any  $a$ ,  $b$ ,  $c$  length or  $\alpha$ ,  $\beta$ ,  $\gamma$  angle large enough to contain the model will suffice. We always chose  $a = b = c$  and  $\alpha = \beta = \gamma = 90^\circ$  for icosahedral viruses. The position of the model as specified in the original coordinate file needs to be determined empirically. For example, for some icosahedral viruses, the coordinate files center the virus atoms at the crystallographic unit cell origin, whereas others center them elsewhere in the unit cell. If the coordinate file specifies atoms for an entire model but its position in the cell is unknown, a quick way to obtain an estimate of the particle center of mass is to compute the mean  $x$ ,  $y$ , and  $z$  values. The coordinates may then be translated to center the model at the (0, 0, 0) or ( $\frac{1}{2}$ ,  $\frac{1}{2}$ ,  $\frac{1}{2}$ ) positions in the new unit cell.

The retrieved model may also not be in a desired or standard orientation, and hence, the coordinates must be adjusted by applying an orientation ("skew") matrix as necessary before structure factors are computed. Alternatively, after a 3D map is produced (Step 4), it could be rotated within the pseudo-unit cell.

The model coordinate file typically needs textual editing to prepare the data for subsequent programs. PDB files, for example, often contain extraneous information in addition to the atomic coordinates. With PDB files, we removed comment or "remark" lines, noncoordinate information (e.g., sequence and secondary structure), and non-amino-acid residues. Additional changes to the model file may be necessary to make it compatible with the input format for the program used in Step 3. For example, in our XPLOR protocol, unconventional atom designators such as "OXT" (for the terminal oxygen) are standardized ("O" in this example).

For oligomeric macromolecules, the coordinate file often only contains information about one or a few asymmetric units of the structure. The file would otherwise be up to 60 times larger, for example, for structures like the icosahedral viruses that possess 532 point-group (60-fold) symmetry. Hence, if symmetry must be applied to generate an entire molecule, this must be accomplished prior to calculation of structure factors. For our applications, we

applied only the necessary point-group symmetry to generate a single molecule centered in the new unit cell. We used a file that contains one matrix for each symmetry-related position. For icosahedral symmetry that meant 60 matrices. The model coordinates are multiplied by each matrix to generate a new set of coordinates in the symmetry-related position. Skew matrices and crystallographic symmetry operations could also be applied to generate additional molecules packed in a pseudo-crystal structure in any space group.

Note that it is important to distinguish between the matrices described here and the matrices that may be found in the coordinate file. For example, a PDB file often comes with noncrystallographic symmetry matrices and crystal symmetry matrices of the unit cell in which the protein was originally crystallized. The noncrystallographic matrices can generate an asymmetric unit that may or may not contain a biologically complete molecule. To generate the full molecular array one also must apply the coordinate file's crystal symmetry matrices. In our procedure, we only used a simple triclinic ( $P1$ ) cell, so all matrices are noncrystallographic (except the identity matrix) and the complete molecule is generated by them.

### Step 3: Compute Structure Factors

Once the model coordinate file is edited to satisfaction, decisions must be made about (1) what resolution limit should be imposed and (2) what temperature factor should be applied. Then the structure factors are calculated.

The resolution of the final density map is set by limiting the number of structure factors computed. All crystallographic software provides this flexibility. For most analyses we computed all structure factors from lower-resolution limits of  $F_{000}$  (XPLOR) or 1000 Å (CCP4) out to limits in the range 20–30 Å. Occasionally it is also useful to generate structure factors at higher or lower resolution (see Fig. 2).

A "B" or temperature factor defines the extent of atomic thermal motion in a crystal structure. It is applied to the structure factors to help minimize truncation errors ("Fourier ripples") that are generated when a sharp resolution limit is imposed. The B factor dampens the abrupt drop in the amplitudes of the structure factors near the resolution limit and thereby helps to eliminate the unwanted Fourier ripples in the calculated density map (Step 4). We found that a B value of 2000 Å<sup>2</sup> is appropriate for generating maps at 20- to 30-Å resolution. Other groups have used 500 Å<sup>2</sup> (Grimes *et al.*, 1997) and 6793 Å<sup>2</sup> (Stewart *et al.*, 1991). These values are much higher than the typical B values of 10–30 Å<sup>2</sup> for well-refined, high-resolution crystal structure determinations (Drenth, 1994).

Once the symmetry, resolution, and temperature factor values are set, the structure factors can be computed. We used the programs XPLOR (Brünger, 1987) or SFALL from the CCP4 program suite (Bailey, 1994).

### Step 4: Compute 3D Density Map

The final step in generating a density map involves inversion of the structure factors (Fourier synthesis). This can be accomplished by most crystallographic software packages, including XPLOR and CCP4, or by other similar programs. We have developed and used EMSF3DBT, a FORTRAN program that has evolved over many years and has benefited our image reconstruction work on both single particles and crystalline specimens. This program computes a Fourier synthesis of the structure factors calculated in Step 3 (Fig. 1) and outputs a 3D density map consisting of cubic voxels of user-specified dimension. The map is then examined with an X-windows-based, interactive graphics program, RobEM, that is written in "C" and FORTRAN and runs on VMS or UNIX workstations ([http://bilbo.bio.purdue.edu/~workshop/help\\_robem/](http://bilbo.bio.purdue.edu/~workshop/help_robem/)). If the map does not appear to be correct (e.g., wrong centering or

orientation), then the need for additional editing or other manipulations (Steps 2–4; Fig. 1) can be assessed.

## DISCUSSION

We have outlined a procedure for computing a low-resolution 3D density map from a set of high-resolution, atomic coordinates. This method of “stepping back” in resolution has been used by ourselves and others to “move forward” the analysis of primarily cryoTEM density maps (e.g., Baker and Cheng, 1996; Beuron *et al.*, 1998; Cheng, 1992; Cheng *et al.*, 1994; Grimes *et al.*, 1997; Stewart *et al.*, 1991, 1993; Trus *et al.*, 1997; Wikoff *et al.*, 1994; Wriggers *et al.*, 1998). Specifically, the computation of low-resolution density from atomic models has aided structural investigations in assessing the quality of 3D reconstructions, making accurate corrections to micrographs to compensate for effects of the contrast transfer function of the electron microscope, calibrating the dimensions or handedness of cryoTEM reconstructions, computing accurate difference maps between X-ray and EM-derived structures, modeling features in a macromolecule or macromolecular complex, and producing models for use in determining particle centers and orientations as needed to compute 3D reconstructions.

In some of the following examples a lower-resolution reconstruction was computed after an atomic structure of the same macromolecule was determined. One might wonder, if atomic resolution structures are available, why are lower-resolution reconstructions of the same structures needed? The answer is that there are gaps in the information gained by crystallography that can be filled in by combining cryoTEM and crystallographic structures. It is possible to combine these two types of data because experiments in which identical structures were determined by the two methods have shown that the two techniques yield compatible results. The four-step procedure we have outlined is one way this compatibility has been demonstrated.

### *Reliability of CryoTEM Reconstructions*

From the high-resolution coordinates, we produced and rendered several icosahedral virus structures at 20–40 Å for direct comparison with 3D reconstructions of the same viruses (Fig. 3). The compatibility between low-resolution structures determined from atomic coordinates and from cryoTEM images is excellent (Fig. 3). This high fidelity lends confidence in the reliability of the image reconstructions. This reliability is further supported by experiments in which atomic structures derived from X-ray crystallography are fit into relevant image reconstructions (Fig. 2; see Wriggers *et al.*, 1998, 1999; Volkman and Hanein, 1999). Several icosahedral viruses whose

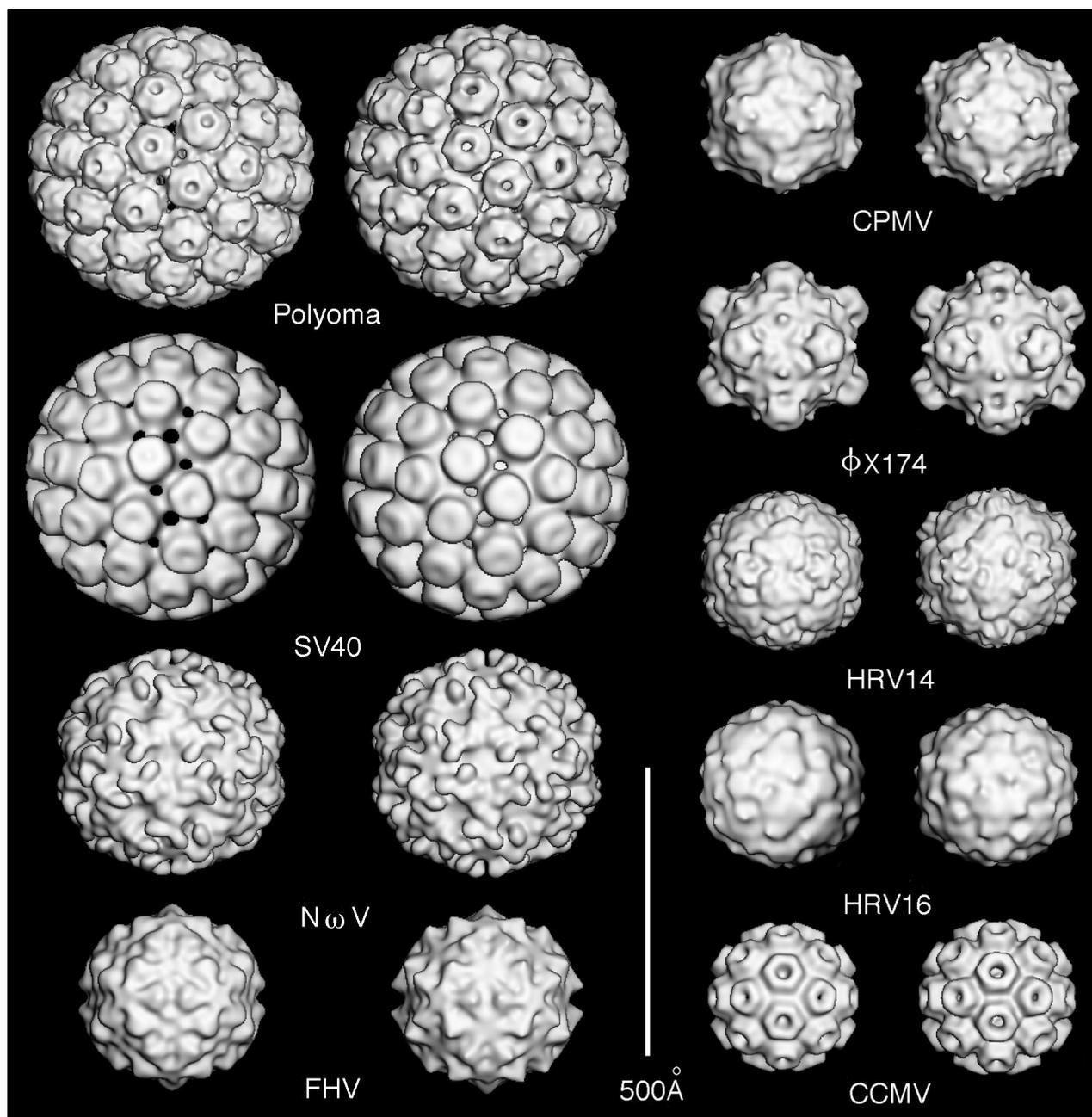
structures have been solved both by X-ray and cryoTEM methods have been compared in this way. For example, the high-resolution structures of the capsids of CCMV, Flock House virus (FHV), bacteriophage  $\Phi$ X174, and alfalfa mosaic virus were superimposed within the envelopes of the corresponding image reconstructions at 23 Å (Speir *et al.*, 1995, Fig. 8), 22 Å (Cheng *et al.*, 1994, Fig. 5), 26 Å (Ilag *et al.*, 1995, Fig. 5; Olson *et al.*, 1992), and 27 Å (Kumar *et al.*, 1997, Fig. 4) resolution, respectively. All four reconstructions exhibited excellent fidelity with the X-ray models. These same conclusions have been reached in investigations of many other types of specimens studied by cryoTEM image reconstruction (e.g., see Baker and Johnson, 1996). To our knowledge, there are no reported studies in which the relatively small structures derived from NMR have been compared with the relatively large cryoTEM image reconstructions.

By computing reduced-resolution maps at various resolution levels (e.g., Fig. 2), the described procedure can be used to assess the resolution of a cryoTEM reconstruction. Stewart *et al.* (1991) used the atomic coordinates of the adenovirus hexon protein to compute density maps of this trimeric protein at 15- to 60-Å resolution in 5-Å increments. They then compared each map to the hexon structure in their cryoTEM reconstruction. By visual inspection, the 35-Å hexon map most closely resembled the hexon structure revealed in the reconstruction. The 35-Å value agreed with a quantitative assessment of the reconstruction's resolution. A similar comparison was made between a HRV–Fab reconstruction and a reduced-resolution map (Fig. 2).

Many of the comparisons between cryoTEM and crystallographic structures mentioned above are visual and hence qualitative. Crystallographic density maps at reduced resolution can be compared quantitatively to cryoTEM density maps. Correlation coefficients (e.g., Cheng, 1992, Fig. 3; Cheng *et al.*, 1994, Fig. 10) and crystallographic R-factors provide two typical quantitative measures of compatibility.

### *Estimating the Contrast Transfer Function of CryoTEM Images*

Unstained, frozen-hydrated biological specimens generate very little amplitude contrast in electron images owing to the small difference between the densities of macromolecules and the vitrified buffer that surrounds the specimen. Electron images of suitable contrast are produced by underfocusing the microscope objective lens by up to several micrometers to generate sufficient phase contrast. This, along with the spherical aberration inherent in electromagnetic lenses, leads to electron images that are distorted in a manner mathematically described by the



**FIG. 3.** Comparisons of low-resolution density maps computed from atomic models (left member of each pair) with 3D reconstructions computed from cryoTEM images (right member of each pair) exhibit striking similarities. The atomic structures were computed at resolutions that gave best qualitative fits to the corresponding reconstructions. All the atomic models were obtained from X-ray crystallographic structure determinations, for which atomic coordinates are available. For these comparisons, no temperature factors were applied in generating the low-resolution density maps from the atomic structures. The resolutions at which the atomic models were rendered and appropriate citations for the crystallographic and cryoTEM work are as follows: murine polyomavirus (polyoma), 23 Å (Stehle and Harrison, 1996; Belnap *et al.*, unpublished results); simian virus 40 (SV40), 40 Å (Stehle *et al.*, 1996; Baker *et al.*, 1988); *Nudaurelia capensis*  $\omega$  virus (N $\omega$ V), 27 Å (Munshi *et al.*, 1996; Johnson *et al.*, 1994); Flock House virus (FHV), 28 Å (Fisher and Johnson, 1993; Cheng *et al.*, 1994); cowpea mosaic virus (CPMV), 27 Å (Stauffer *et al.*, 1987; Wang *et al.*, 1992); bacteriophage  $\phi$ X174 ( $\phi$ X174), 28 Å (McKenna *et al.*, 1992; Olson *et al.*, 1992); human rhinovirus serotype 14 (HRV14), 23 Å (Rossmann *et al.*, 1985; Olson *et al.*, unpublished results); human rhinovirus serotype 16 (HRV16), 35 Å (Hadfield *et al.*, 1997; Olson *et al.*, unpublished results); and cowpea chlorotic mottle virus (CCMV), 30 Å (Speir *et al.*, 1995).

contrast transfer function (CTF) (Erickson and Klug, 1971; Misell, 1978). X-ray diffraction patterns are not formed with lenses and hence are not distorted in the manner that electron images are. Consequently, X-ray structures provide an excellent standard for estimating the corrections needed to eliminate or at least minimize the effects of the CTF in image reconstructions. The ultimate goal in applying such CTF corrections to cryoTEM images or reconstructions is to produce an artifact-free reconstruction that faithfully represents the true density distribution within the structure being examined.

In one example (Stewart *et al.*, 1993), the atomic structure of adenovirus hexon protein was used to compute a density map at 25-Å resolution. Then, a CTF and an exponential-decay factor were applied to the reduced-resolution structure. Several different underfocus and decay-factor values were tried and the resulting map was compared visually to a 3D reconstruction of adenovirus computed to 25-Å resolution. The values that gave the best visual match were used to apply an appropriate inverse CTF to the reconstruction.

In another example, the atomic structure of FHV (Fisher and Johnson, 1993) was used to correct a 22-Å resolution image reconstruction of FHV (Cheng *et al.*, 1994). The correction function, which was applied in Fourier space (i.e., to structure factors), was generated on the basis of a least-squares fit between the Fourier amplitudes computed from the cryoTEM and X-ray structures. This procedure was first used to derive an empirical CTF-correction function that was applied to a 25-Å resolution reconstruction of cowpea mosaic virus (CPMV), a virus whose atomic structure was known (Cheng, 1992). The corrected FHV and CPMV reconstructions both revealed regions of locally ordered nucleic acid inside the capsids that could not be detected in the high-resolution X-ray structures.

#### *Calibrating Size or Handedness of Image Reconstructions*

The nominal magnification values displayed on most electron microscopes only approximate the true magnifications of the recorded micrographs. Accurate calibration of instrument magnification requires the use of reliable standards. Low-resolution density maps derived from X-ray experiments have served as excellent standards for scaling the size of 3D cryoTEM reconstructions. In fact, what is actually calibrated in this process is the dimension of the cubic voxels that constitute the density map. X-ray structures are accurate standards because the diffraction patterns from which they are derived are generated from X rays of known wavelengths. The 22.5-Å X-ray crystal structure of murine polyomavirus (Grif-

fith *et al.*, 1992) was used as a standard to calibrate a 29-Å cryoTEM reconstruction of the same virus, which in turn was used to determine the size of bovine papillomavirus (Belnap *et al.*, 1993). FHV (Cheng *et al.*, 1994), canine parvovirus (Wikoff *et al.*, 1994), and poliovirus (Belnap *et al.*, unpublished results) represent some examples of viruses whose image reconstructions have been calibrated against density maps of corresponding X-ray structures, computed to the resolution limits of the reconstructions.

The absolute handedness of a macromolecule is lost in an image from a transmission electron microscope because the image is a two-dimensional (2D) projection of the 3D specimen and identical 2D projections can be formed from either member of an enantiomeric pair. In contrast, the handedness of atomic-resolution structures is determined by simply making the structure consistent with right-handed  $\alpha$ -helices and L-amino acids. If an atomic structure is available and handedness is still present at the reduced resolution, then a lower-resolution structure derived from the atomic coordinates can be a handedness standard for a cryoTEM reconstruction and one avoids the necessity of a tilt experiment (for a review of the tilt experiment, see Belnap *et al.*, 1997). Such was the case for adenovirus, where the handedness of the hexon portion of the 3D reconstruction was matched to the handedness of the crystallographic hexon structure (Stewart *et al.*, 1991).

#### *Difference Maps*

Lower-resolution density maps derived from atomic coordinates can be used in conjunction with electron microscopic images or reconstructions to determine the location of density not modeled in the atomic-resolution structure. The location of such density can be found by fitting an atomic structure into a cryoTEM reconstruction or by subtracting the reduced-resolution density from the reconstruction. Three structural studies of the adenovirus capsid are superb examples of this (Furcinitti *et al.*, 1989; Stewart *et al.*, 1991, 1993). In conjunction with other biochemical studies, the atomic structure of adenovirus hexon was used to find the location of minor capsid proteins. A difference map was computed by subtracting a 25-Å map of the hexon from a 25-Å CTF-corrected reconstruction in one case (Stewart *et al.*, 1993) and the hexon structure was compared to a 35-Å reconstruction in another case (Stewart *et al.*, 1991). Furcinitti *et al.* (1989) low-pass-filtered a 6-Å crystal structure of the hexon to 15 Å and then projected it into a 2D image. This projected model image was then subtracted from 2D images of a portion of the adenovirus capsid. Interestingly, the 2D images were of negatively stained specimens and were recorded in a scanning transmission electron microscope. This

indicates that the procedure of computing lower-resolution maps from high-resolution structures is a technique applicable to more than just cryoTEM reconstructions.

Included in the density not observed by X-ray crystallography are flexible, disordered, or nonsymmetric portions of molecules such as the genomes of icosahedral viruses. However, density corresponding to such regions is often revealed at resolutions much lower ( $\gg 10\text{\AA}$ ) than those typically recorded during collection of crystallographic diffraction patterns. A low-resolution map computed from high-resolution coordinates provides a useful tool for computing difference maps to reveal the distribution of structural elements that are invisible in high-resolution structures. For example, density corresponding to the genomic RNA of FHV was revealed in a difference density map computed by subtracting a 22- $\text{\AA}$  resolution map of the known capsid structure (Fisher and Johnson, 1993) from a 22- $\text{\AA}$  cryoTEM reconstruction (Cheng *et al.*, 1994). The difference map showed regions of highly ordered density just beneath the inner wall of the capsid. The presence of this density, identified as duplex RNA in the X-ray structure but which was omitted from the atomic model used to produce the low resolution map, clearly established the validity of this approach (Cheng *et al.*, 1994).

#### *Modeling Macromolecules and Macromolecular Complexes*

In cryoTEM structural studies of macromolecular complexes, some components may be known to atomic detail and others may not. Also, a 3D cryoTEM reconstruction of the complex may be unavailable or may be difficult to obtain. In such instances, use of a low-resolution density map computed from the known structure may greatly assist in interpretation of the complex and may obviate the need for additional cryoTEM and 3D reconstruction work. These conditions occurred in a study of ClpAP (Beuron *et al.*, 1998). ClpAP is a chaperone-linked, ATP-dependent protease from *Escherichia coli*, whose quaternary structure is a hexameric ring of ClpA subunits axially aligned with a pair of heptameric rings of ClpP subunits (Kessel *et al.*, 1995). ClpA and ClpP associate to form a proteolytically active complex when they are mixed together. Though preliminary attempts to obtain a 3D reconstruction of the ClpAP complex were thwarted in part because the complex adopts preferred orientations in cryoTEM samples, an  $\sim 30\text{-\AA}$  resolution reconstruction of the ClpA structure was combined with a 30- $\text{\AA}$  map of the atomic ClpP structure to synthesize a low-resolution model of the ClpAP complex. The translational and rotational relationships of the sixfold ClpA and the sevenfold ClpP were determined by comparing projec-

tions of the 30- $\text{\AA}$  ClpA and ClpP structures to cryoTEM images of the ClpAP complex.

Studies of bluetongue virus, adenovirus, papillomavirus, and actin filaments provide additional examples for which reduced-resolution maps derived from atomic coordinates have been used in modeling macromolecules or macromolecular complexes. A 23- $\text{\AA}$  reconstruction of the bluetongue virus core was used as a template for modeling the atomic coordinates of the VP7 capsid protein (Grimes *et al.*, 1997). The orientation and position of the trimeric VP7 in the core were refined further by computing a 24- $\text{\AA}$  density map from the atomic coordinates and then using an automated fitting routine that fitted the 24- $\text{\AA}$  map into the 23- $\text{\AA}$  reconstruction. Two alternative crystal forms of VP7 were tested and resulted in fits with correlation coefficients of 0.9 and 0.67. The recently determined atomic-resolution structure of the core confirmed that the crystal form that best fit the image reconstruction was indeed the same form in the assembled core (Grimes *et al.*, 1998). In a manner similar to that used in the bluetongue virus study, a 25- $\text{\AA}$  map of the adenovirus hexon protein that was computed from the atomic coordinates was used as a probe to find the orientations and positions of hexons in the adenovirus capsid (Stewart *et al.*, 1993). In a third example, a 9- $\text{\AA}$  density map of murine polyomavirus capsid was computed from atomic coordinates and compared to a 9- $\text{\AA}$  reconstruction of the structurally related bovine papillomavirus (Trus *et al.*, 1997). In their development of an automated routine for merging atomic coordinates and lower-resolution 3D reconstructions, Wrigger *et al.* (1998) computed a reduced-resolution density map of actin filament from a pseudo-atomic model. They then used the map as a template to realign the atomic coordinates of actin monomers as a control for their procedure.

#### *Determination of Particle Orientations and Origins*

Accurate determination of each particle's orientation and origin (center) is the most critical step in computing reliable 3D reconstructions from images of randomly oriented particles. A low-resolution map derived from an atomic model or a high-resolution structure may be helpful in initiating model-based orientation and origin searches (e.g., Baker and Cheng, 1996) for an identical or a closely related structure. The 3D reconstruction of FHV (Cheng *et al.*, 1994) was derived in this manner by making use of the known FHV capsid structure. The atomic model of HRV serotype 14 was likewise used as a starting model to compute a 25- $\text{\AA}$  resolution 3D reconstruction of HRV type 16 (Baker and Cheng, 1996). Finally, a 3D reconstruction at 23- $\text{\AA}$  resolution of canine parvovirus (CPV) complexed with Fab

fragments from neutralizing, monoclonal antibody A3B10 was solved by making use of a pseudo-atomic model (Wikoff *et al.*, 1994). This model consisted of the crystal structure of CPV to which 60 copies of the known structure of Fab Kol were docked in a radial orientation near each of the 60 identical, presumed binding sites. In all three of these examples, features that were not present in the starting model were revealed in the resulting reconstructions, and hence, the final reconstructions were not strongly biased by the choice of starting model but instead were a faithful rendering of the details contained in the images (Baker and Cheng, 1996). However, one should always remain cautious of potential model-based bias in the reconstruction process when such starting models are used, and the consistency of the final result with the raw images should always be tested. Note also that (1) the sizes of the imaged particles and the model structure must be normalized and (2) the handedness of the model will be imposed on the resulting reconstruction.

We thank R. Cheng for performing initial studies at Purdue in which low-resolution structures were computed from atomic models; T. Dokland for suggesting the use of CCP4 and for help in setting up our CCP4 procedure; W. Zhang, T. Dokland, R. Cheng, Z. Che, and R. Ashmore for software; M. Skinner, R. Ashmore, and W. Zhang for programming assistance; B. Putnam for computer help; A. Steven for critically reviewing the manuscript; the reviewers of the paper for helpful suggestions; R. Ashmore for preparing computer programs for distribution on our Web site; and M. Sturgeon for setting up the Web site at [http://bilbo.bio.purdue.edu/~baker/high\\_to\\_low/high\\_to\\_low.html](http://bilbo.bio.purdue.edu/~baker/high_to_low/high_to_low.html) and for assistance in preparing the manuscript for submission. This work was supported by grants from the National Institutes of Health (GM33050) and National Science Foundation (MCB9527131) to T.S.B. and from the National Institutes of Health (GM10704) to T.J.S. J.T.F. was supported by an NIH Biophysics Training Grant (GM08296).

## REFERENCES

- Bailey, S. (1994). The CCP4 suite—Programs for protein crystallography, *Acta Crystallogr. Sect. D* **50**, 760–763.
- Baker, T. S., Drak, J., and Bina, M. (1988) Reconstruction of the three-dimensional structure of simian virus 40 and visualization of the chromatin core, *Proc. Natl. Acad. Sci. USA* **85**, 422–426.
- Baker, T. S., and Cheng, R. H. (1996) A model-based approach for determining orientations of biological macromolecules imaged by cryo-electron microscopy, *J. Struct. Biol.* **116**, 120–130.
- Baker, T. S., and Johnson, J. E. (1996) Low resolution meets high: Towards a resolution continuum from cells to atoms, *Curr. Opin. Struct. Biol.* **6**, 585–594.
- Belnap, D. M., Grochulski, W. D., Olson, N. H., and Baker, T. S. (1993) Use of radial density plots to calibrate image magnification for frozen-hydrated specimens, *Ultramicroscopy* **48**, 347–358.
- Belnap, D. M., Olson, N. H., and Baker, T. S. (1997) A method for establishing the handedness of biological macromolecules, *J. Struct. Biol.* **120**, 44–51.
- Beuron, F., Maurizi, M. R., Belnap, D. M., Kocsis, E., Booy, F. P., Kessel, M., and Steven, A. C. (1998) At sixes and sevens: Characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease, *J. Struct. Biol.* **123**, 248–259.
- Böttcher, B., Wynne, S. A., and Crowther, R. A. (1997) Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy, *Nature* **386**, 88–91.
- Brünger, A. T., Kuriyan, J., Karplus, M. (1987) Crystallographic R factor refinement by molecular dynamics, *Science* **235**, 458–460.
- Che, Z., Olson, N. H., Leippe, D., Lee, W.-M., Mosser, A. G., Ruckert, R. R., Baker, T. S., and Smith, T. J. (1998) Antibody-mediated neutralization of human rhinovirus 14 explored by means of cryo-electron microscopy and X-ray crystallography of virus–Fab complexes, *J. Virol.* **72**, 4610–4622.
- Cheng, R. H. (1992) Correlation of cryo-electron microscopic and X-ray data and compensation of the contrast transfer function, *In* Bailey, G. W., Bentley, J., and Small, J. A. (Eds), *Proceedings of the 50th Annual Meeting of the Electron Microscopy Society of America* pp. 996–997, San Francisco Press, San Francisco.
- Cheng, R. H., Reddy, V. S., Olson, N. H., Fisher, A. J., Baker, T. S., and Johnson, J. E. (1994) Functional implications of quasi-equivalence in a T = 3 icosahedral animal virus established by cryo-electron microscopy and X-ray crystallography, *Structure* **2**, 271–282.
- Conway, J. F., Cheng, N., Zlotnick, A., Wingfield, P. T., Stahl, S. J., and Steven, A. C. (1997) Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy, *Nature* **386**, 91–94.
- Drenth, J. (1994) *Principles of Protein X-ray Crystallography*, Springer-Verlag, New York.
- Erickson, H. P., and Klug, A. (1971) Measurement and compensation of defocusing and aberrations by Fourier processing of electron micrographs, *Philos. Trans. R. Soc. London B* **261**, 105–118.
- Fisher, A. J., and Johnson, J. E. (1993) Ordered duplex RNA controls capsid architecture in an icosahedral animal virus, *Nature* **361**, 176–179.
- Furcinitti, P. S., van Oostrum, J., and Burnett, R. M. (1989) Adenovirus polypeptide IX revealed as capsid cement by difference images from electron microscopy and crystallography, *EMBO J.* **8**, 3563–3570.
- Griffith, J. P., Griffith, D. L., Rayment, I., Murakami, W. T., and Caspar, D. L. D. (1992) Inside polyomavirus at 25-Å resolution, *Nature* **355**, 652–654.
- Grimes, J. M., Jakana, J., Ghosh, M., Basak, A. K., Roy, P., Chiu, W., Stuart, D. I., and Prasad, B. V. V. (1997) An atomic model of the outer layer of the bluetongue virus core derived from X-ray crystallography and electron cryomicroscopy, *Structure* **5**, 885–893.
- Grimes, J. M., Burroughs, J. N., Gouet, P., Diprose, J. M., Malby, R., Zientara, S., Mertens, P. P. C., and Stuart, D. I. (1998) The atomic structure of the bluetongue virus core, *Nature* **395**, 470–478.
- Hadfield, A. T., Lee, W.-M., Zhao, R., Oliveira, M. A., Minor, I., Rueckert, R., and Rossmann, M. G. (1997) The refined structure of human rhinovirus 16 at 2.15 Å resolution: Implications for the viral life cycle, *Structure* **5**, 427–441.
- Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E., and Downing, K. H. (1990) Model for the structure of bacteriorhodopsin based on high-resolution electron cryomicroscopy, *J. Mol. Biol.* **213**, 899–929.
- Ilag, L. L., Olson, N. H., Dokland, T., Music, C. L., Cheng, R. H., Bowen, Z., McKenna, R., Rossmann, M. G., Baker, T. S., and Incardona, N. L. (1995) DNA packaging intermediates of bacteriophage ΦX174, *Structure* **3**, 353–363.
- Johnson, J. E., Munshi, S., Liljas, L., Agrawal, D., Olson, N. H., Reddy, V., Fisher, A., McKinney, B., Schmidt, T., and Baker, T. S.

- (1994) Comparative studies of T = 3 and T = 4 icosahedral RNA insect viruses, *Arch. Virol.* **9**, 497–512.
- Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. Sect. A* **47**, 110–119.
- Kessel, M., Maurizi, M. R., Kim, B., Kocsis, E., Trus, B. L., Singh, S. K., and Steven, A. C. (1995) Homology in structural organization between *E. coli* ClpAP protease and the eukaryotic 26 S proteasome, *J. Mol. Biol.* **250**, 587–594.
- Kühlbrandt, W., Wang, D. N., and Fujiyoshi, Y. (1994) Atomic model of plant light-harvesting complex by electron crystallography, *Nature* **367**, 614–621.
- Kraulis, P. J. (1991) Molscript—A program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr.* **24**, 946–950.
- Kumar, A., Reddy, V. S., Yusibov, V., Chipman, P. R., Hata, Y., Fita, I., Fukuyama, K., Rossmann, M. G., Loesch-Fries, L. S., Baker, T. S., and Johnson, J. E. (1997) The structure of alfalfa mosaic virus capsid protein assembled as a T = 1 icosahedral particle at 4.0-Å resolution, *J. Virol.* **71**, 7911–7916.
- McKenna, R., Xia, D., Willingmann, P., Ilag, L. L., Krishnaswamy, S., Rossmann, M. G., Olson, N. H., Baker, T. S., and Incardona, N. L. (1992) Atomic structure of single-stranded DNA bacteriophage ΦX174 and its functional implications, *Nature* **355**, 137–143.
- Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic molecular graphics, *Methods Enzymol.* **277**, 505–524.
- Misell, D. L. (1978) Image analysis, enhancement, and interpretation, in Glauert, A. M. (Ed.), *Practical Methods in Electron Microscopy*, Vol. 7, Elsevier/North-Holland, Amsterdam.
- Munshi, S., Liljas, L., Cavarelli, J., Bomu, W., McKinney, B., Reddy, V., and Johnson, J. E. (1996) The 2.8 angstrom structure of a T = 4 animal virus and its implications for membrane translocation of RNA, *J. Mol. Biol.* **261**, 1–10.
- Nogales, E., Wolf, S. G., and Downing, K. H. (1998) Structure of the αβ tubulin dimer by electron crystallography, *Nature* **391**, 199–203.
- Olson, N. H., Baker, T. S., Willingmann, P., and Incardona, N. L. (1992) The three-dimensional structure of frozen-hydrated bacteriophage ΦX174, *J. Struct. Biol.* **108**, 168–175.
- Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B., and Vriend, G. (1985) Structure of a human common cold virus and functional relationship to other picornaviruses, *Nature* **317**, 145–153.
- Smith, T. J., Olson, N. H., Cheng, R. H., Liu, H., Chase, E. S., Lee, W. M., Leippe, D. M., Mosser, A. G., Rueckert, R. R., and Baker, T. S. (1993) Structure of human rhinovirus complexed with Fab fragments from a neutralizing antibody, *J. Virol.* **67**, 1148–1158.
- Smith, T. J., Chase, E. S., Schmidt, T. J., Olson, N. H., and Baker, T. S. (1996) Neutralizing antibody to human rhinovirus 14 penetrates the receptor-binding canyon, *Nature* **383**, 350–354.
- Speir, J. A., Munshi, S., Wang, G., Baker, T. S., and Johnson, J. E. (1995) Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy, *Structure* **3**, 63–78.
- Stauffer, C. V., Usha, R., Harrington, M., Schmidt, T., Hosur, M. V., and Johnson, J. E. (1987) The structure of cowpea mosaic virus at 3.5 Å resolution, in Morjas, D., Drenth, J., Strandberg, B., Suck, D., and Wilson, K. (Eds.), pp. 293–308, *Crystallography in Molecular Biology*, Plenum, New York.
- Stehle, T., and Harrison, S. C. (1996) Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments, *Structure* **4**, 183–194.
- Stehle, T., Gamblin, S. J., Yan, Y., and Harrison, S. C. (1996) The structure of simian virus 40 refined at 3.1 Å resolution, *Structure* **4**, 165–182.
- Stewart, P. L., Burnett, R. M., Cyrklaff, M., and Fuller, S. D. (1991) Image reconstruction reveals the complex molecular organization of adenovirus, *Cell* **67**, 145–154.
- Stewart, P. L., Fuller, S. D., and Burnett, R. M. (1993) Difference imaging of adenovirus: Bridging the resolution gap between X-ray crystallography and electron microscopy, *EMBO J.* **12**, 2589–2599.
- Trus, B. L., Roden, R. B. S., Greenstone, H. L., Vrhel, M., Schiller, J. T., and Booy, F. P. (1997) Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 Å resolution, *Nature Struct. Biol.* **4**, 413–420.
- Volkman, N., and Hanein, D. (1999) Quantitative fitting of atomic models into observed densities derived by electron microscopy, *J. Struct. Biol.* **125**, 176–184.
- Wang, G., Porta, C., Chen, Z., Baker, T. S., and Johnson, J. E. (1992) Identification of a Fab interaction footprint site on an icosahedral virus by cryoelectron microscopy and X-ray crystallography, *Nature* **355**, 275–278.
- Wikoff, W. R., Wang, G., Parrish, C. R., Cheng, R. H., Strassheim, M. L., Baker, T. S., and Rossmann, M. G. (1994) The structure of a neutralized virus: Canine parvovirus complexed with neutralizing antibody fragment, *Structure* **2**, 595–607.
- Wriggers, W., Milligan, R. A., Schulten, K., and McCammon, J. A. (1998) Self-organizing neural networks bridge the biomolecular resolution gap, *J. Mol. Biol.* **284**, 1247–1254.
- Wriggers, W., Milligan, R. A., and McCammon, J. A. (1999) Situs: A package for docking crystal structures into low-resolution maps from electron microscopy, *J. Struct. Biol.* **125**, 185–195.
- Zhang, X.-J., and Matthews, B. W. (1995) EDPDB: A multi-functional tool for protein structure analysis, *J. Appl. Crystallogr.* **28**, 624–630.