

# Combining Electron Microscopic with X-Ray Crystallographic Structures

Michael G. Rossmann,<sup>1</sup> Ricardo Bernal, and Sergei V. Pletnev<sup>2</sup>

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

Received November 2, 2001, and in revised form December 26, 2001; published online April 10, 2002

**An algorithm has been developed for placing three-dimensional atomic structures into appropriately scaled cryoelectron microscopy maps. The first stage in this process is to conduct a three-dimensional angular search in which the center of gravity of an X-ray crystallographically determined structure is placed on a selected position in the cryoelectron microscopy map. The quality of the fit is measured by the sum of the density at each atomic position. The second stage is to refine the three angles and three translational parameters for the best (usually 25 to 100) fits. Useful criteria for this refinement include the sum of densities at atomic sites, the lack of atoms in negative or low density, the absence of atomic clashes between symmetry-related positions of the atomic structure, and the distances between identifiable features in the map and their positions on the fitted atomic structure. These refinements generally lead to a convergence of the originally chosen, top scoring fits to just a few (about 3 to 8) acceptable possibilities. Usually, the best remaining fit is clearly superior to any of the others.** © 2002 Elsevier Science (USA)

**Key Words:** cryoelectron microscopy; X-ray crystallography; virus structures; density-fitting algorithm.

## INTRODUCTION

Cryoelectron microscopy (cryoEM) image reconstructions of biological macromolecules or macromolecular assemblies are rapidly improving, both in quality and in the attainable resolution (Böttcher *et al.*, 1997; Conway *et al.*, 1997; Chiu *et al.*, 1999; Mancini *et al.*, 2000; Zhou *et al.*, 2000, 2001). This is a consequence of the development of field emission guns to produce more coherent electron beams, the

use of larger voltages to improve depth penetration, the development of liquid helium-cooled specimen holders, greater mechanical and magnetic stability, and computational advances (Kimura *et al.*, 1997; Thuman-Commike and Chiu, 2000; van Heel *et al.*, 2000). Whereas three-dimensional, single particle reconstructions were routinely limited to about 20 Å resolution, now many such reconstructions frequently extend to 15, 10, or even 7 Å resolution. This has made it possible to study objects that are unlikely to crystallize because of their complex or irregular shapes (e.g., tailed and fibered bacteriophages), because they exist as a mixture of conformational states such as might occur in a reaction mixture, because of inadequate purification techniques, or because of the transient nature of many biological phenomena, as is the case for virus assembly intermediates.

At this time, cryoEM techniques used to produce three-dimensional images from randomly oriented particles are usually inadequate to reach the near atomic resolution normally associated with X-ray crystallographic techniques. However, by combining cryoEM and X-ray results, it is possible to determine the structure of complex biological assemblies in atomic detail (Stewart *et al.*, 1993; Baker and Johnson, 1996; Grimes *et al.*, 1997; Volkman and Hanein, 1999; Roseman, 2000; Rossmann, 2000; Zhou *et al.*, 2001). Structures of component parts determined by crystallographic techniques can be fitted into the lower resolution cryoEM maps of larger assemblies. Although fitting of atomic models into cryoEM maps can be achieved by hand using standard graphical programs such as O (Jones *et al.*, 1991), the results are often unsatisfactory, as frequently there is considerable uncertainty whether there are alternative solutions and to what extent the best solution is better than any other solution. Furthermore, experience has shown that it is more difficult to fit a large atomic structure into a convoluted electron density map than it is to fit simple

<sup>1</sup> To whom correspondence and reprint requests should be addressed. Fax: 765-496-1189. E-mail: mgr@indiana.bio.purdue.edu.

<sup>2</sup> Present address: NCI-Frederick, MCL, P.O. Box B, Building 539, Frederick, MD 21702.

chemical units, such as peptides or aromatic rings, into easily recognizable densities. Here we describe procedures that make the art of fitting models into density into a more quantitative procedure and illustrate these techniques with various examples.

Comparison of a proposed structure with an observed cryoEM density can be performed either in reciprocal (Kolatkár *et al.*, 1999) or real (Cheng *et al.*, 1995; Grimes *et al.*, 1997; Volkman and Hain, 1999; Wriggers *et al.*, 1999; Kikkawa *et al.*, 2000; Roseman, 2000; Rossmann, 2000; Volkman *et al.*, 2001; Wriggers and Chacón, 2001) space. Real space has the benefit of being able to define more easily the volume within which the comparison is to be made, whereas reciprocal space has the advantage of utilizing a more readily definable search or refinement target. We have chosen to perform both the initial search and the final refinement in real space. The procedure described here also considers the impact of symmetry on the cryoEM density and the imposition of restraints provided by additional information, such as the location of glycosylation sites or knowledge about the function of specific amino acids. It is assumed that the model fitting operation will be made on a map that is the final product of the image reconstruction procedure, which includes such corrections of the raw images as are made necessary by the contrast transfer function. The algorithms described here have been incorporated into the program EMfit.

### CRYOEM MAP SCALING

Calibration of electron microscope magnification is not very accurate in general and can be off by as much as  $\pm 5\%$  in some cases. More certain results can be achieved by including objects of accurately known size in the specimen. However, that cannot always be done easily or such data might be unavailable. An alternative method is to compare the cryoEM map with a related X-ray crystallographically determined map (Hewat and Blaas, 1996) because the absolute size of the X-ray wavelength used in a crystallographic determination is usually known to within 1 part in 1000. In the absence of an accurate absolute standard of comparison, it is possible to compare two cryoEM maps directly in order to put these onto the same relative scale when calculating cryoEM difference maps (Pletnev *et al.*, 2001). However, these comparisons should be done only in regions of the map (defined by a mask) that have easily distinguishable features.

Let the density of the “experimental” cryoEM map be  $\rho_1(x_1, y_1, z_1)$  at the point  $(x_1, y_1, z_1)$  and the density in the standard map (X-ray crystallographic or any other map of reliable scale) be  $\rho_2(x_2, y_2, z_2)$  at the

point  $(x_2, y_2, z_2)$ , equivalent to  $(x_1, y_1, z_1)$  in map  $\rho_1$ . Then, the height of the density in the experimental map,  $\rho_1$ , can be put onto the scale of the standard map,  $\rho_2$ , by minimizing  $\sum[\rho_2 - (a + b\rho_1)]^2$ , where  $a$  and  $b$  are scaling constants and the sum is taken over all density grid points  $(x_2, y_2, z_2)$  within the defined mask. The factor  $a$  is required to “float” the density of  $\rho_1$  at the same arbitrary level as that of  $\rho_2$ . The coordinate systems in the two maps must be defined as having a common, superimposable origin. This could be, for instance, the center of the point group defining the 532 icosahedral symmetry in comparing maps of related isometric viruses.

The goodness-of-fit criteria for comparing different density maps can be defined as either a correlation coefficient, CC, where

$$CC = \left[ \frac{\sum(\langle \rho_1' \rangle - \rho_1') (\langle \rho_2 \rangle - \rho_2)}{\left[ \sum(\langle \rho_1' \rangle - \rho_1')^2 \sum(\langle \rho_2 \rangle - \rho_2)^2 \right]^{1/2}} \right] \quad (1a)$$

or as a residual,  $R$ , where

$$R = \sum |(\rho_1' - \rho_2)| / \sum |\rho_2|, \quad (1b)$$

and  $\langle \rho_1' \rangle$  and  $\langle \rho_2 \rangle$  are the mean of the densities in the scaled experimental and the standard maps. These criteria can be calculated for densities within the defined mask (or radial limits), while assuming successive increments of the pixel size in map 1, in order to search for the highest CC or lowest  $R$  factor (Rossmann, 2000). Because grid points in map 1 will not be coincident, in general, with grid points in map 2, the densities  $\rho_2(x_2, y_2, z_2)$  need to be interpolated in map 2. If  $\rho_2(x_2, y_2, z_2) = \rho_2(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$ , where  $(x_0, y_0, z_0)$  (with a density  $\rho_{000}$ ) is the nearest grid point position to  $(x_2, y_2, z_2)$  in map 2 and the other eight densities surrounding  $(x_2, y_2, z_2)$  are identified with appropriate subscripts, then, assuming approximately linear variation of density between grid points,

$$\begin{aligned} \rho_2(x_2, y_2, z_2) = & \rho_{000} \\ & + [(\rho_{100} - \rho_{000})\Delta x + (\rho_{010} - \rho_{000})\Delta y \\ & + (\rho_{001} - \rho_{000})\Delta z] \\ & + (\rho_{000} + \rho_{110} - \rho_{100} - \rho_{010})\Delta x\Delta y \\ & + (\rho_{000} + \rho_{011} - \rho_{010} - \rho_{001})\Delta y\Delta z \\ & + (\rho_{000} + \rho_{101} - \rho_{001} - \rho_{100})\Delta z\Delta x \\ & + [(\rho_{100} + \rho_{010} + \rho_{001} + \rho_{111} - \rho_{000} \\ & - \rho_{101} - \rho_{011} - \rho_{110})\Delta x\Delta y\Delta z]. \end{aligned} \quad (2)$$

TABLE I

Exploration of Pixel Size by Fitting the TBEV E Glycoprotein Dimer (Rey *et al.*, 1995) onto the Icosahedral Twofold Axis of a 24-Å Resolution Dengue Virus CryoEM Reconstruction (Kuhn *et al.*, 2001)

Pixel (Å)	<i>sumf</i> (%)	- <i>den</i> (No. of atoms)	$\theta_1$ (°)	<i>centx</i> (Å)	<i>centy</i> (Å)	<i>centz</i> (Å)	<i>sumf</i> for individual domains of E (%)		
							D1	D2	D3
2.7	50.0	47.0	29.0	0.0	0.0	216.5	50.0	54.6	34.7
2.8	51.5	40.0	27.2	0.0	0.0	224.0	50.3	55.3	41.9
2.9	53.0	29.0	24.5	0.0	0.0	232.5	48.6	56.5	49.4
3.0	47.6	62.0	25.5	0.0	0.0	236.5	49.6	47.1	47.8

Note. The microscope-calibrated pixel size was 2.68 Å. The allowed degrees of freedom were the rotation about the icosahedral twofold axis ( $\theta_1$ ) and the translation along the axis (*centz*). Although the best results were for a pixel size of 2.9 Å, this could be partly artifactual (see text). Definitions of *sumf*, -*den*, *centx*, *centy*, and *centz* are given in the text.

An alternative procedure can be used when an initial approximate fit of a known crystal structure into the EM density has already been established. The quality of the fit (see definition of *sumf* below) can then be explored as a function of pixel size (Table I). However, at low resolution this can result in an overestimation of the pixel size, as there will be a tendency to crowd the fitted atoms into the highest densities at the center of the fitted region by expanding the density relative to the atomic model.

#### DETERMINING THE ORIENTATIONS THAT GIVE THE BEST FITS OF AN ATOMIC MODEL INTO A CRYOEM MAP

Let the atomic positions of a model be given by ( $X, Y, Z$ ) or, in vector notation, by  $\mathbf{X}$ , relative to an orthogonal coordinate system. Let the center of gravity of the model be at  $\mathbf{S}$ . Let the rotation matrix required to place the model into the "reference" EM density be  $[E]$ , where the nine elements of this matrix are most conveniently expressed in terms of Eulerian angles (Rossmann and Blow, 1962). Then,

$$\mathbf{X}' = [E]\mathbf{X} + \mathbf{d}, \quad (3)$$

where  $\mathbf{X}'$  are the coordinates of the model atoms in the EM map and  $\mathbf{d}$  is a translation vector. Let  $\mathbf{S}'$  (with Cartesian components *centx*, *centy*, *centz*) be the approximate target position in the EM map for placement of the model's center of gravity. Then, it follows that

$$\mathbf{S}' = [E]\mathbf{S} + \mathbf{d}$$

and, hence,

$$\mathbf{d} = \mathbf{S}' - [E]\mathbf{S}. \quad (4)$$

In many cases, particularly in the study of icosahedral viruses, the cryoEM density has been generated with the assumption that the structure has specific symmetry properties (532 in the case of an icosahedron). The placement of an object at a specific reference position with an assumed orientation will then require the generation of other symmetry-related objects. By analogy with crystallography, symmetry that has been imposed in the EM reconstruction can be considered as being "crystallographic symmetry," whereas other symmetry operations (e.g., those resulting from quasi-symmetry in spherical virus structures) are valid only within a restricted volume of the total density, analogous to "noncrystallographic symmetry" (NCS). In fitting an atomic structure to the EM density, it will be necessary to test the fit of the structure at every "NCS"-related site, as these have independently determined densities. On the other hand, there is no point to fit the structure at "crystallographically" related sites, as these have identical density. The environment of every "NCS"-related structure can be completely different, although in the special case of "quasi" symmetry, as defined by Caspar and Klug (1962), the "NCS"-related sites will have similar environments. For convenience, the terms "crystallographic" and "NCS" symmetry are used in this paper to signify symmetry that is valid over the whole of the reconstructed image and symmetry that is valid only in a defined locality of the image. The term "quasi-equivalence" could be substituted for NCS symmetry for the special case whenever the symmetry of an icosahedral virus obeys the Caspar and Klug rules, but, as is well known, this is quite frequently violated.

Let the reference molecule or subunit be reproduced by  $L$  crystallographic and  $M$  NCS rotation operations about axes through the origin, given by  $[R_m]$  ( $m = 1, LM$ ) (Table II). Thus, using (3) and (4),

**TABLE II**

Generation of 532 Crystallographic and  $T = 4$  Noncrystallographic Symmetry (see Fig. 1) as Used in the Analysis of SINV (Pletnev *et al.*, 2001; W. Zhang *et al.*, unpublished results)

Symmetry elements			Comments	Operators generated
$\kappa$ (°)	$\psi$ (°)	$\phi$ (°)		
120.0	69.900	-90.000	Threefold crystallographic axis	(1), 2, 3
180.0	74.700	-81.200	Twofold NCS on operator 1	(1), 4
120.0	79.400	-72.500	Threefold NCS on operator 4	(4), 5, 6
72.0	90.000	-58.280	Fivefold crystallographic operator on 4	(4), 7, 8, 9, 10
72.0	90.000	-58.280	Fivefold crystallographic operator on 6	(6), 11, 12, 13, 14

*Note.* The polar angle ( $\kappa, \psi, \phi$ ) definitions are as given by Rossmann and Blow (1962). The twofold axes of the icosahedral 532 point group are along  $X, Y, Z$  (Fig. 1). The symmetry element operates on the operator given in parentheses. The polar angles  $\kappa, \psi$ , and  $\phi$  are input parameters to the EMfit program and, therefore, can be used to describe any combination of rotation axes.

$$\mathbf{X}_m = [R_m][E](\mathbf{X} - \mathbf{S}) + [R_m]\mathbf{S}' \quad (5)$$

The rotations implied by the matrices  $[R_m]$  can be defined by polar coordinates used as input to the EMfit program (Table II). Thus, the program can be used for any symmetry appropriate for the problem, including the fitting of an object to an asymmetric density map.

The quality of fit (*sumf*) can be measured as the average value of the density at all  $M$ -independent, NCS-related atomic positions using either all atoms or only the  $C_\alpha$  atoms of a protein. The position of an atom will not necessarily be at a grid point. Hence, the density will need to be interpolated by using Eq. (2). If there are  $N$  atoms in the model, then the normalized quality of fit of an atomic structure to a map is given by

$$sumf = \left\{ 100.0 \cdot \sum_M \sum_N [\rho(\mathbf{X}_{m,n})] \right\} [(M \cdot N \cdot \rho_{max})]^{1/2}, \quad (6)$$

where  $-den$  is the highest density in the map.

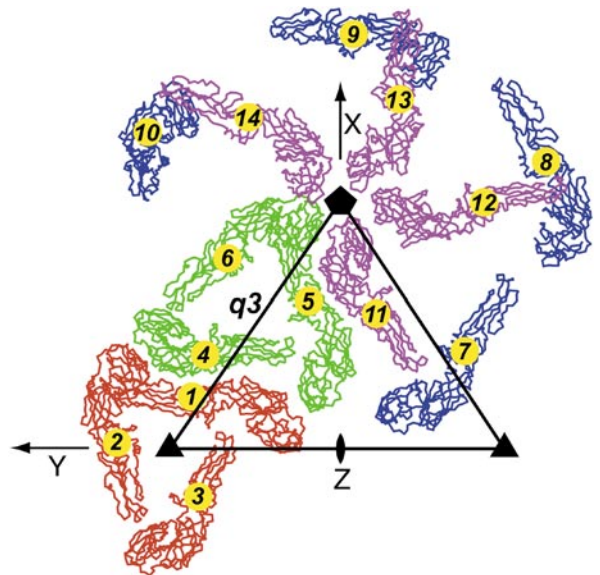
The sensitivity for finding a good fit can be enhanced by counting the number of atoms (*neg*) in each of the  $M$  molecules per NCS asymmetric unit, expressed as a percentage ( $-den$ ) of the total number of atoms ( $M \cdot N$ ) that occur in low or negative density, where

$$-den = \left( 100.0 \cdot \sum_M neg \right) [(M \cdot N)]^{1/2}. \quad (7)$$

In essence, this is putting more emphasis on keeping atoms out of low density than on insisting on fitting atoms to the highest possible density values. A good fit will be one that maximizes the sum of sampled

density and minimizes the number of atoms in low density.

In addition, the number of steric clashes should be minimized. Clashes can be defined as the number of atoms (*nclash*) in one subunit that come to within a specified distance ( $\Delta$ ) of any atom in a neighboring subunit. This can be expressed as a percentage (*clash*) of the number of atoms per subunit, where



**FIG. 1.**  $C_\alpha$  traces of TBEV E glycoprotein monomers, distributed by  $T = 4$  icosahedral symmetry. The molecules are numbered with the same symmetry operators as used in Table II. The triangle shows the outline of one icosahedral asymmetric unit. One NCS quasi-threefold symmetry element (q3) is shown, as are the fivefold, threefold, and twofold icosahedral symmetry elements around the periphery of the triangle. The  $(X, Y, Z)$  Cartesian coordinate system is also shown. The red molecules are related by an icosahedral threefold axis and the green molecules are related by a quasi-threefold axis. The red trimer is related to the green trimer by a quasi-twofold axis. The icosahedral fivefold axis generates the blue monomers from molecule 1 and the magenta monomers from molecule 6.

$$clash = \left( 100.0 \cdot \sum_M nclash \right) [(M \cdot N)]^{1/2}. \quad (8)$$

If the search structure contains all the main chain as well as side chain atoms, then the minimal approach distance should be set to about 3 Å or, if only C<sub>α</sub> atoms are available (each representing an amino acid), then an approach distance of about 7 Å is reasonable.

Because the time taken to determine the closest approach between two sets of atomic positions is proportional to the square of the total number of atoms, a less time-consuming, although less accurate, procedure, proportional to the number of atoms, was adopted for all but the final calculations. Two map representations (*a* and *b*) were used, each with the same three-dimensional array of grid points appropriate to the resolution of the cryoEM map. Initially, the value of each grid point in map *a* was set to 0. As a first step, grid points on map *b* were marked with a 1 if they were within the limiting radius, Δ, of any atom associated with a subunit that was not one of the *M* NCS-related subunits defining the “crystallographic” asymmetric unit. All other points in map *b* were set to 0. Then map *b* was added to map *a*. This was to assure finding contacts between atoms in the reference crystallographic asymmetric unit and atoms in neighboring crystallographic asymmetric units. In the next step, grid points in map *b* were marked if they were within Δ of any atom associated with the first of the *M* subunits within the crystallographic asymmetric unit, all other points being set to 0. Then map *b* was again added to map *a*. Thus, any grid point that now showed a 2 in map *a* indicated a clash between atoms outside the asymmetric unit and the first subunit within the asymmetric unit. This step was then repeated for each of the *M* subunits within the crystallographic asymmetric unit (for instance, there would be four independent subunits per asymmetric unit in a *T* = 4 icosahedral virus). In the final step, each atom of each subunit within the asymmetric unit was reexamined. If an atom center corresponded to a grid point that had a value greater than 1, then this atom had a steric clash with an atom either in another subunit within the crystallographic asymmetric unit or in a subunit outside the crystallographic asymmetric unit.

Other criteria can also help to establish the best fit by recognizing specific features in the cryoEM map. Examples are glycosylation sites on the atomic model whose positions in the cryoEM map had been established previously (Kolatkhar *et al.*, 1999; Pletnev *et al.*, 2001) and knowledge that a certain part of the atomic model (e.g., the receptor binding site) is

required to dock to a specific region of another structure (e.g., the receptor recognition site) whose location in the map is already known (He *et al.*, 2001). A set of restraining positions can be identified on the atomic model that corresponds to a set of target positions in the map. Preference can be given to those fits that produce the lowest rms or mean distances (“*avgdist*”) between restraining and target positions. Conversely, it may be required that regions in the map need to be avoided as they are known to be occupied by another structure (Pletnev *et al.*, 2001). A suitable criterion to control this situation is to count (*near*) the number of atoms in the atomic model that approach the occupied region to within a specified distance.

#### COMBINING THE VARIOUS CRITERIA FOR DETERMINATION OF THE BEST FIT

It is necessary to establish a single criterion of fit (*R*<sub>crit</sub>) from the diverse considerations that contribute to a reasonable fit of the atomic model to the cryoEM density. The various criteria mentioned above have a variety of different dimensions ranging from units proportional to electrons per atomic site (*sumf*), or lengths (*avgdist*), to being dimensionless (*neg*, *nclash*, *near*). It is possible, however, to put each criterion onto a comparable dimensionless scale by expressing each as a ratio of its value (*v*) with its standard deviation from its mean, σ(*v*), computed over the random set of orientations and positions encountered in the general three-dimensional orientation search (Table III). Then the average of these various criteria is given by

$$R_{\text{crit}} = \sum_i \omega_i s_i [(v_i - \langle v \rangle) / \sigma_i(v)] (\sum_i \omega_i)^{1/2}, \quad (9)$$

where the sums are taken over the *i* independent criteria, ω<sub>*i*</sub> are the weights to be placed on each, and s<sub>*i*</sub> is ±1 according to whether the difference needs to be maximized (s<sub>*i*</sub> = +1) or minimized (s<sub>*i*</sub> = -1) for the *i*th criterion. Normally, the weights would be unity, assuming that each has an equal importance in finding the best fit. But if, for instance, more weight is to be placed upon fitting the carbohydrate sites to their target positions in the map than on maximizing the fit of the model into the density, then the weight on *avgdist* should be increased to, say, 3.0 instead of 1.0 (Tables IIIb and IV).

#### THE SEARCH PROCESS

A complete three-dimensional search of all angles, using the relationship given by (5), can be computed to make sure that all likely fits of the model to the density have been found. This search can be per-

TABLE III

Selected Fits Found in a General Eulerian ( $\theta_1, \theta_2, \theta_3$ ) Search with Center,  $S'$ , at (21.4, 68.3, 286.5 Å) for the Reference Position of the TBEV E Monomer Fitted into the 12 Å CryoEM Map of SINV, the Latter Having  $T = 4$  Symmetry as Given in Table II (W. Zhang *et al.*, unpublished results)

Fit No.	$\theta_1$ (°)	$\theta_2$ (°)	$\theta_3$ (°)	a. Search results				
				<i>sumf</i> (%)	<i>clash</i> (No. of atoms)	<i>-den</i> (No. of atoms)	<i>avgdist</i> (Å)	<i>near</i> (No. of atoms)
1	0.00	45.00	195.00	38.58	2.0	181.0	23.37	0.0
2	15.00	45.00	180.00	38.17	3.0	170.0	22.46	0.0
3	345.00	45.00	210.00	37.57	7.0	198.0	23.82	0.0
4	345.00	60.00	210.00	37.26	1.0	222.0	29.09	0.0
5	45.00	60.00	150.00	37.13	8.0	172.0	24.02	0.0
10	30.00	45.00	150.00	36.07	13.0	134.0	21.62	0.0
13	345.00	60.00	195.00	35.53	0.0	209.0	27.74	0.0
14	330.00	60.00	225.00	35.48	4.0	235.0	28.70	0.0

Note.  $\theta_1, \theta_2, \theta_3$  are the Eulerian angles that define  $[E]$ . Definitions of *sumf*, *clash*, *-den*, *avgdist*, and *near* are given in the text.

b. Standard rms deviation from the mean ( $\sigma$ ) for the various fitting criteria among the top 25 fits

	<i>clash</i>			<i>-den</i>		<i>avgdist</i> (Å)	<i>near</i>	
	<i>sumf</i> (%)	(No. of atoms)	(%)	(No. of atoms)	(%)		(No. of atoms)	(%)
Average	35.52	11.28	0.71	203.76	12.90	24.35	0.00	0.00
$\sigma$	1.60	17.60	1.11	37.92	2.40	2.61	0.00	0.00
Weights	1.00	1.00	1.00	1.00	1.00	3.00	1.00	1.00

Note. Definitions of *sumf*, *clash*, *-den*, *avgdist*, and *near* are given in the text.

formed conveniently using Eulerian angles (Rossmann and Blow, 1962) covering all unique orientations by exploring the range  $0 = \theta_1 < 2\pi$ ,  $0 = \theta_2 = \pi$ ,  $0 = \theta_3 < 2\pi$ , saving the best (usually somewhere between 25 and 100) results for further refinement Table IIIa). Because of the considerable computing time required for a three-dimensional search, a good strategy is to use fairly large angular intervals for the initial calculation and then to refine subse-

quently the best retained fits. About  $10^\circ$  increments ( $\delta^\circ$ ) are suitable for cryoEM maps at 20 Å resolution (*res*) with an object that has a mean diameter (*dia*) of 40 Å, in accordance with the formula

$$\delta^\circ = (res \cdot 180) / (3 \cdot \pi \cdot dia). \quad (10)$$

A further enhancement of speed in performing the three-dimensional search can be obtained by reject-

TABLE IV  
Climb Results on the Top 25 Fit Results Shown in Table III

a. The top 25 results converge to only four separate results												
Fit No.	$R_{crit}$	<i>sumf</i> (%)	<i>clash</i> (%)	<i>-den</i> (%)	<i>avgdist</i> (Å)	<i>near</i> (%)	$\theta_1$ (°)	$\theta_2$ (°)	$\theta_3$ (°)	<i>centx</i> (Å)	<i>centy</i> (Å)	<i>centz</i> (Å)
10	0.810	37.8	2.2	10.1	20.5	0.0	30.2	41.8	163.5	24.4	65.8	285.5
14	0.264	36.3	3.7	11.9	21.2	0.0	326.8	43.5	219.0	23.4	63.3	284.5
25	-2.367	39.2	17.5	10.1	28.7	0.0	126.5	135.5	151.8	17.9	63.8	283.5

b. The  $R_{crit}$  factor is the weighted average of the different fit criteria expressed as  $s[(v - \langle v \rangle) / \sigma(v)]$ , as given in Eq. (10)

Fit No.	$R_{crit}$	<i>sumf</i> (%)	<i>clash</i> (%)	<i>-den</i> (%)	<i>avgdist</i> (Å)	<i>near</i> (%)
13	0.980	2.376	0.186	1.523	0.925	0.000
10	0.810	1.398	-1.348	1.180	1.481	0.000
14	0.264	0.476	-2.712	0.416	1.223	0.000
25	-2.367	2.315	-23.100	1.154	-1.665	0.000

Note. Definitions of  $R_{crit}$ , *sumf*, *clash*, *-den*, *avgdist*, *near*, *centx*, *centy*, and *centz* are given in the text.

**TABLE V**  
*Climb* Refinement of Fit No. 3 in Table III

The initial values of the six parameters were					
$\theta_1$	$\theta_2$	$\theta_3$	<i>centx</i>	<i>centy</i>	<i>centz</i>
345.0	45.0	210.0	21.4	68.3	286.5
$R_{\text{crit}}$ at					
Parameter	$x - \Delta x$	$x$	$x + \Delta x$	Best $x$	$\Delta x$
1	0.270	0.327	-1.344	345.00	15.00
2	-0.824	0.327	-0.608	45.00	15.00
3	-0.344	0.520	0.327	195.00	15.00
4	0.479	0.520	0.505	21.40	1.00
5	0.518	0.520	0.488	68.30	1.00
6	0.547	0.582	0.554	284.50	1.00
End of cycle 1					
1	0.473	0.818	-0.494	0.00	15.00
2	-0.421	0.818	-0.415	45.00	15.00
3	0.401	0.818	-0.951	195.00	15.00
4	0.781	0.818	0.794	22.40	1.00
5	0.814	0.818	0.799	70.30	1.00
6	0.776	0.818	0.792	283.50	1.00
End of cycle 4. Convergence achieved with current values of $\Delta x$ , therefore reduce increments.					
1	1.016	1.029	1.022	357.00	0.25
2	1.008	1.029	1.026	40.50	0.25
3	1.021	1.029	1.021	193.50	0.25
4	0.996	1.029	1.011	23.90	0.50
5	1.028	1.029	0.990	68.30	0.50
6	0.963	1.029	1.015	284.50	0.50
End of cycle 10. Convergence achieved with current values of $\Delta x$ . Smaller $\Delta x$ values are inappropriate for the map resolution. End of <i>climb</i> .					

*Note.* Each line of output following the section on initial values gives the  $R_{\text{crit}}$  values around the *climb*-determined maximum. The increment  $\Delta x$  is in degrees for parameters 1, 2, and 3 and in Å for parameters 4, 5, and 6. Note that the final  $R_{\text{crit}}$  value is a little different here than that reported for the final results in Table IV. That is because the calculation of clashes is done by the fast method (see text) for the *climb* refinement, whereas a more careful, but slower, procedure was used to tabulate the final results shown in Table IV.

ing all orientations in which any of the distances between corresponding restraining and target sites exceeds suitably chosen restraining distances. Another useful strategy is to utilize only the *sumf* estimation of the quality of fit (Eq. 6) during the initial search. The additional measures of fit given by Eq. (7) and (8), as well as the *avgdist* and *near* criteria described above, need only be computed during refinement of the best retained fits found by the crude three-dimensional search.

#### REFINEMENT OF THE BEST SEARCH RESULTS

The general search procedure is normally used only to find an approximate set of angles that give a reasonable fit of the search model when its center of gravity is placed on a designated position in the EM map, whereas a refinement procedure is required to more accurately define both the angular orientation and the position of the search model. Thus, the top (usually 25 to 100) best results from the general

search will need to be refined by either a least-squares procedure or a systematic or “*climb*” search (Rossmann, 2000). As it is difficult to define a target function for a least-squares procedure when the criterion of fit depends upon sampling density at atomic positions, a *climb* procedure has been developed. This procedure is useful for fitting one independent rigid body into density. A least-squares procedure would be preferable when fitting multiple domains into adjacent densities (Lescar *et al.*, 2001), as it would permit the simultaneous refinement of all the rotational and translational parameters, but a general least-squares algorithm for model fitting has not yet been completed for the EMfit program.

The *climb* procedure (Table V) consists of investigating one parameter (usually there are three rotational and three translational parameters) at a time using a set of preset increments. The first parameter is increased or decreased to attain the best fit measured by  $R_{\text{crit}}$ . This procedure is then repeated for

**TABLE VI**  
Various Checks on Quality of the Remaining Fits

a. *sumf* as a function of each of the three domains of the TBEV monomer and  $T = 4$  NCS fitted into the 12 Å resolution cryoEM density of SINV

Domain	NCS operator	<i>sumf</i> (%)	Number of atoms
1	1	35.2	122
1	4	33.0	122
1	5	34.6	122
1	6	32.1	122
2	1	41.8	180
2	4	42.9	180
2	5	43.2	180
2	6	44.5	180
3	1	38.7	93
3	4	38.3	93
3	5	40.7	93
3	6	39.9	93

*Note.* Results for the best (top) fit given in table IV. The *sumf* (the average height of density at all atoms in asymmetric unit) was computed over  $C_{\alpha}$  atoms only. The largest grid point density was set to 100. There is some evidence that the central domain 1 does not fit as well into the cryoEM density as domains 2 and 3 at the extremities of the long E monomer. However, each of the  $T = 4$  (Table II)-related monomers fits about equally well into the density.

b. Short contacts (<7.5 Å) between atoms in neighboring subunits for the best fit of the TBEV E monomer into the SINV cryoEM density (Table IV)

Neighboring subunit				Subunit in asymmetric unit				Distance (Å)
Atom	Residue	Number	Symmetry operator	Atom	Residue	Number	Symmetry operator	
$C_{\alpha}$	THR	310	5	$C_{\alpha}$	THR	303	1	7.1
$C_{\alpha}$	LYS	311	5	$C_{\alpha}$	THR	303	1	7.4
$C_{\alpha}$	THR	303	4	$C_{\alpha}$	LYS	311	2	6.2
$C_{\alpha}$	SER	168	11	$C_{\alpha}$	THR	56	5	7.2
$C_{\alpha}$	ALA	134	11	$C_{\alpha}$	ALA	195	5	7.3
$C_{\alpha}$	ALA	189	11	$C_{\alpha}$	ALA	195	5	6.9
$C_{\alpha}$	ALA	189	11	$C_{\alpha}$	GLN	196	5	6.5
$C_{\alpha}$	SER	168	11	$C_{\alpha}$	ASN	221	5	6.5

*Note.* None of the close contacts represent a serious clash.

c. Proportion of interpreted to uninterpreted grid densities

Height	-6	-4	-2	0	2	4	6	8	10
Ratio 1		38.9	11.6	5.1	2.7	1.8	1.8	0.9	2.3
Ratio 2	59.6	12.5	3.5	2.7	0.5	0.2	0.1	0.0	0.0

*Note.* Vacant space for fitting the TBEV E dimer into the dengue virus cryoEM map (Kuhn *et al.*, 2001). Dengue virus and TBEV are homologous flaviviruses. Thus, the known X-ray structure of the TBEV E dimer should be present in the icosahedrally averaged dengue map. Ratio 1 gives the number of vacant pixels in a shell between 230 and 260 Å divided by the number of pixels occupied by a dimer fitted onto the icosahedral axes (corresponding to one monomer per icosahedral asymmetric unit). The ratio is shown as a function of the height of the pixels in a range from -10 to +10 arbitrary height units. The fitted molecule was represented by spheres of 6 Å radius around each  $C_{\alpha}$  atom. The ratio will be large for low or negative density as these should be associated with few if any atoms. As the height of the density increases, ratio 1 approaches ~2.0, demonstrating that there would be space for two more monomers in each icosahedral asymmetric unit. Ratio 2 is the number of vacant pixels divided by the number of pixels occupied by the dimer fitted onto the icosahedral twofold axis plus the dimer fitted onto a quasi-twofold axis. This showed that essentially all positive density pixels had been interpreted.

the second parameter and so forth. When the last parameter has been adjusted, the search is returned to the first parameter until no further change has occurred in any of the parameters. All the parameter increments are then divided by 10, and the same procedure is repeated. The process is considered to have converged when the parameter increments have decreased to an insignificant value, say  $0.25^{\circ}$  in

angles and 0.5 Å in translations, when searching a 20 Å resolution map (see Eq. 10).

If reasonable increments were chosen for the angular increments (see Eq. 9 above), the top 25 to 100 or so best results are likely to cluster around the larger peaks of the search space. Hence, the refinement of the best results will converge to only a few different end points. These can often be reduced

**TABLE VII**  
Fitting the Same Structure to the Right- or Wrong-Handed Map

Fit No.	$R_{crit}$	$sumf$ (%)	$clash$ (%)	$-den$ (%)	$avgdist$ (Å)	$near$ (%)	$\theta_1$ (°)	$\theta_2$ (°)	$\theta_3$ (°)	$centx$ (Å)	$centy$ (Å)	$centz$ (Å)
a. Correct hand												
26	1.059	30.5	0.8	15.1	20.6	0.0	1.5	39.5	178.5	22.4	67.8	284.5
b. Wrong hand												
1	0.800	7.1	2.2	49.6	27.2	0.0	120.0	135.2	329.5	25.4	-64.3	286.5

*Note.* The TBEV E glycoprotein was fitted to a 9 Å resolution map of Semliki Forest virus (Mancini *et al.*, 2000). The fit was restrained with the glycosylation sites also used in fitting the TBEV E molecule to SINV (Table IV).

further because some, which may appear to have rather different angles, are in reality essentially the same as a consequence of degeneracy in the Eulerian angles. When the difference in orientation between any two refined end points is less than  $\delta^\circ$ , then the two results can be considered the same.

#### ADDITIONAL CRITERIA TO CHECK THE BEST STRUCTURAL CANDIDATES

The first check on the results after running an automatic fitting program should always be a viewing of the fitted coordinates in the cryoEM density using a computer graphics program, such as O (Jones *et al.*, 1991). This will require a little extra effort as it will be necessary to organize the map and coordinates in a suitable format. Thus, a quick, temporary, and not entirely satisfactory substitute is printed output of sections showing density height and the pixels covered by the fitted molecule.

There usually are a number of additional data that can verify the uniqueness of the structure. When fitting a molecule with compactly folded domains into cryoEM density, there should be some concern as to whether there is any change of the interdomain hinge angles in the cryoEM structure compared to those observed in the crystal structure of the molecule. Examples are immunoglobulin domains in cell surface molecules that can serve as viral receptors (Kolatkhar *et al.*, 1999; Belnap *et al.*, 2000; He *et al.*, 2000, 2001; Xiao *et al.*, 2001) or the three quite differently folded domains of the alpha-virus E1 and flavivirus E glycoproteins (Lescar *et al.*, 2001; Pletnev *et al.*, 2001). A check can be made on whether there is hinge movement by fitting the molecule into the cryoEM density as a rigid body and then checking on the fit of the atoms in each domain separately (Table VIa). Similarly, if the quality of fit for NCS-related molecules is substantially different, then the NCS symmetry operators can be adjusted to attain equally good fits for each of the NCS-related molecules (Table VIa) (Cheng *et al.*, 1995).

Another check can be the nature and seriousness of the molecular clashes between fitted molecules (Table VIb). This can also suggest the types of interactions that are required for the assembly of the complex, such as a virus. If there is a serious clash, then either the fitted molecule does not represent the structure in the complex or the fitting operation was in error.

A further check can be made to determine whether all the pixels associated with higher density have been interpreted once all components (e.g., viral protein subunits) of the complex (e.g., the whole virion) have been fitted into the density. The example given in Table VIc relates to the interpretation of the dengue virus 24 Å resolution map (Kuhn *et al.*, 2001). In this case, the exact triangulation number (Caspar and Klug, 1962) was uncertain and did not follow the quasi-symmetry rules. Fitting of one homologous tick-borne encephalitis virus (TBEV) E glycoprotein dimer onto an icosahedral twofold axis resulted in the interpretation of about one-third of all larger density pixels. The subsequent fitting of another dimer into a general position resulted in the interpretation of most of the remaining larger density pixels.

In general, the hand of a cryoEM reconstruction is unknown. An arbitrary decision is made in the early stages of the reconstruction when selecting the relative orientation of the initial particle orientations. However, it is possible to establish the hand experimentally by tilting the stage a known amount in a known direction. This technique was used to establish the hand of Ross River virus (RRV) (Cheng *et al.*, 1995). Subsequently, all other alphavirus reconstructions (Fuller *et al.*, 1995; Mancini *et al.*, 2000) presumably assumed the same hand. The hand of alphaviruses, such as RRV, is easily recognized on account of the highly asymmetric appearance of the 80 spike protrusions from the icosahedral virus. The original hand determination has now been verified by comparing the quality of fit to both the plus (correct) and minus (wrong) hands of the map (Table

VII). The results show that even for the best fit to the wrong-handed map, the average height of density at all  $C_{\alpha}$  atoms (*sumf*) was only 7.1% of the highest density in the map, compared to 30.5% for the correct hand.

The usual method for determining the resolution of a cryoEM map derived from a single particle reconstruction is to divide the original images into two groups and correlate the Fourier-inverted structure as a function of resolution. However, this method suffers from a variety of uncertainties. The quality of fit of a molecular component to different maps of the same or related objects could be another method of comparison. For instance, the TBEV E glycoprotein can be fitted to a Sindbis virus (SINV) map (nominal resolution of 12 Å) with a *sumf* value of 39.3% (Table IV), whereas fitting the same molecule to a homologous Semliki Forest virus (SFV) map (with a nominal resolution of 9 Å) gives a *sumf* value of 30.5% (Table VII). Although this might suggest that the “12 Å” map is “better” than the “9-Å” map, the NCS operations were optimized for SINV (Table II), making it possible that SINV and Semliki Forest viruses have slightly different quasi-symmetry operators. Furthermore, the two maps were placed onto equivalent density scales by setting the highest density to 100 in each map, which is a far from accurate way of comparing maps (see map scaling). Thus, attempts at comparing the quality of maps based upon fitting a known atomic structure must make sure that all appropriate factors have been properly considered. These results also demonstrate the well-known difficulty of estimating map resolution.

A convenient way of sequentially fitting different components to a map (Table VIc) can be achieved by setting those pixels to 0 that are within a given distance,  $\Delta$  (e.g., van der Waals radius), of a fitted atom. This produces a “difference” map showing only the density of the uninterpreted map. This technique was used for interpreting a 24 Å resolution map of dengue virus (Kuhn *et al.*, 2001) and for recognizing the shape and organization of the E2 glycoprotein of SINV (W. Zhang *et al.*, unpublished results), whose atomic resolution structure is unknown.

#### THE EMfit PROGRAM

The procedures described here have been incorporated into the EMfit program, written in Fortran. The program can be used with or without a graphical user interface, written in Perl, which also incorporates extensive help documentation. The EMfit program and its graphical user interface are available free from the authors. The program has been

tested on numerous problems, some of which have been referenced in this paper (see, for instance, He *et al.*, 2001; Kuhn *et al.*, 2001). One test considered the structure of coxsackievirus A21, a picornavirus. The atomic structure of a protomer assembly intermediate, consisting of the viral proteins VP1, VP2, VP3, and VP4, was fitted into a 20 Å resolution cryoEM map. The result showed that, by far, the best fit, based upon agreement with density and lack of clashes with surrounding protomers, was closely similar to the known X-ray structure. The  $C_{\alpha}$  atom positions determined by the EMfit program were all within 1 Å of the 3.0 Å resolution X-ray crystal structure.

We have greatly appreciated the encouragement we have received from Richard Kuhn, Suchetana Mukhopadhyay, Chuan Xiao, and Wei Zhang. We are grateful to Sharon Wilder for preparation of the manuscript. The work was supported by grants from the National Institutes of Health (AI 11219 and AI 45976) and the National Science Foundation (MCB-9986266). R.B. is the recipient of a Department of Education GAANN Training Fellowship.

#### REFERENCES

- 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., McDermott, B. M., Jr., Filman, D. J., Cheng, N., Trus, B. L., Zuccola, H. J., Racaniello, V. R., Hogle, J. M., and Steven, A. C. (2000) Three-dimensional structure of poliovirus receptor bound to poliovirus. *Proc. Natl. Acad. Sci. USA* **97**, 73–78.
- 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 (London)* **386**, 88–91.
- Caspar, D. L. D., and Klug, A. (1962) Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.
- Cheng, R. H., Kuhn, R. J., Olson, N. H., Rossmann, M. G., Choi, H. K., Smith, T. J., and Baker, T. S. (1995) Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell* **80**, 621–630.
- Chiu, W., McGough, A., Sherman, M. B., and Schmid, M. F. (1995) High-resolution electron cryomicroscopy of macromolecular assemblies. *Trends Cell Biol.* **9**, 154–159.
- 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 (London)* **386**, 91–94.
- Fuller, S. D., Berriman, J. A., Butcher, S. J., and Gowen, B. E. (1995) Low pH induces swiveling of the glycoprotein heterodimers in the Semliki Forest virus spike complex. *Cell* **81**, 715–725.
- 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 microscopy. *Structure* **5**, 885–893.
- He, Y., Bowman, V. D., Mueller, S., Bator, C. M., Bella, J., Peng, X., Baker, T. S., Wimmer, E., Kuhn, R. J., and Rossmann, M. G.

- (2000) Interaction of the poliovirus receptor with poliovirus. *Proc. Natl. Acad. Sci. USA* **97**, 79–84.
- He, Y., Chipman, P. R., Howitt, J., Bator, C. M., Whitt, M. A., Baker, T. S., Kuhn, R. J., Anderson, C. W., Freimuth, P., and Rossmann, M. G. (2001) Interaction of coxsackievirus B3 with the full-length coxsackievirus-adenovirus receptor. *Nat. Struct. Biol.* **8**, 874–878.
- Hewat, E. A., and Blaas, D. (1996) Structure of a neutralizing antibody bound bivalently to human rhinovirus 2. *EMBO J.* **15**, 1515–1523.
- 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.
- Kikkawa, M., Okada, Y., and Hirokawa, N. (2000) 15 Å resolution model of the monomeric kinesin motor, KIF1A. *Cell* **100**, 241–252.
- Kimura, Y., Vassilyev, D. G., Miyazawa, A., Kidera, A., Matsushima, M., Mitsuoka, K., Murata, K., Hirai, T., and Fujiyoshi, Y. (1997) Surface of bacteriorhodopsin revealed by high-resolution electron crystallography. *Nature (London)* **389**, 206–211.
- Kolatkar, P. R., Bella, J., Olson, N. H., Bator, C. M., Baker, T. S., and Rossmann, M. G. (1999) Structural studies of two rhinovirus serotypes complexed with fragments of their cellular receptor. *EMBO J.* **18**, 6249–6259.
- Kuhn, R. J., Zhang, W., Rossmann, M. G., Pletnev, S. V., Corver, J., Lenches, E., Jones, C. T., Mukhopadhyay, S., Chipman, P. R., Strauss, E. G., Baker, T. S., and Strauss, J. H. (2001) Structure of dengue virus: Implications for flavivirus organization, maturation, and fusion. *Cell* **108**, 717–725.
- Lescar, J., Roussel, A., Wein, M. W., Navaza, J., Fuller, S. D., Wengler, G., Wengler, G., and Rey, F. A. (2001) The fusion glycoprotein shell of Semliki Forest virus: An icosahedral assembly primed for fusogenic activation at endosomal pH. *Cell* **105**, 137–148.
- Mancini, E. J., Clarke, M., Gowen, B., Rutten, T., and Fuller, S. D. (2000) Cryo-electron microscopy reveals the functional anatomy of an enveloped virus, Semliki Forest virus. *Mol. Cell* **5**, 255–266.
- Pletnev, S. V., Zhang, W., Mukhopadhyay, S., Fisher, B. R., Hernandez, R., Brown, D. T., Baker, T. S., Rossmann, M. G., and Kuhn, R. J. (2001) Locations of carbohydrate sites on Sindbis virus glycoproteins show that E1 forms an icosahedral scaffold. *Cell* **105**, 127–136.
- Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C., and Harrison, S. C. (1995) The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature (London)* **375**, 291–298.
- Roseman, A. M. (2000) Docking structures of domains into maps from cryo-electron microscopy using local correlation. *Acta Crystallogr. Sect. D* **56**, 1332–1340.
- Rossmann, M. G. (2000) Fitting atomic models into electron microscopy maps. *Acta Crystallogr. Sect. D* **56**, 1341–1349.
- Rossmann, M. G., and Blow, D. M. (1962) The detection of subunits within the crystallographic asymmetric unit. *Acta Crystallogr.* **15**, 24–31.
- 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.
- Thuman-Commike, P. A., and Chiu, W. (2000) Reconstruction principles of icosahedral virus structure determination using electron cryomicroscopy. *Micron* **31**, 687–711.
- van Heel, M., Gowen, B., Matadeen, R., Orlova, E. V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M., and Patwardhan, A. (2000) Single-particle electron cryo-microscopy: Towards atomic resolution. *Q. Rev. Biophys.* **33**, 307–369.
- Volkman, N., Amann, K. J., Stoilova-McPhie, S., Egile, C., Winter, D. C., Hazelwood, L., Heuser, J. E., Li, R., Pollard, T. D., and Hanein, D. (2001) Structure of Arp2/3 complex in its activated state and in actin filament branch junctions. *Science* **293**, 2456–2459.
- Volkman, N., and Hanein, D. (1999) Quantitative fitting of atomic models into observed densities derived by electron microscopy. *J. Struct. Biol.* **125**, 176–184.
- Wriggers, W., and Chacón, P. (2001) Modeling tricks and fitting techniques for multiresolution structures. *Structure* **9**, 779–788.
- 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**, 185–189.
- Xiao, C., Bator, C. M., Bowman, V. D., Rieder, E., He, Y., Hébert, B., Bella, J., Baker, T. S., Wimmer, E., Kuhn, R. J., and Rossmann, M. G. (2001) Interaction of coxsackievirus A21 with its cellular receptor, ICAM-1. *J. Virol.* **75**, 2444–2451.
- Zhou, Z. H., Baker, M. L., Jiang, W., Dougherty, M., Jakana, J., Dong, G., Lu, G., and Chiu, W. (2001) Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus. *Nat. Struct. Biol.* **8**, 868–873.
- Zhou, Z. H., Dougherty, M., Jakana, J., He, J., Rixon, F. J., and Chiu, W. (2000) Seeing the herpesvirus capsid at 8.5 Å. *Science* **288**, 877–880.