Fitting atomic models into electron-microscopy maps

Michael G. Rossmann

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

Correspondence e-mail:
mgr@indiana.bio.purdue.edu

Combining X-ray crystallographically determined atomic structures of component domains or subunits with cryo-electron microscopic three-dimensional images at around 22 Å resolution can produce structural information that is accurate to about 2.2 Å resolution. In an initial step, it is necessary to determine accurately the absolute scale and absolute hand of the cryo-electron microscopy map, the former of which can be off by up to 5%. It is also necessary to determine the relative height of density by using a suitable scaling function. Difference maps can identify, for instance, sites of glycosylation, the position of which helps to fit the component structures into the EM density maps. Examples are given from the analysis of alphaviruses, rhinovirus–receptor interactions and poliovirus–receptor interactions.

1. Introduction

Although there have been major advances in the use of single-particle cryo-electron microscopy (cryo-EM), current resolution limits are somewhere in the 6–10 Å range. It is probable that this limit will be pushed to perhaps 4 Å in favorable cases during the next decade. Two of the most impressive results to date are independent reconstructions of hepatitis B virus cores (Böttcher *et al.*, 1997; Conway *et al.*, 1997), where it was possible to make an interpretation of the organization of α-helices. The structure was confirmed subsequently by an X-ray crystallographic study (Wynne *et al.*, 1999). Another exceptional example is the analysis of herpes simplex virus cores at 8 Å resolution (W. Chiu, private communication). These and similar results provide confidence that cryo-EM structures can be taken seriously.

X-ray crystallographic structural results of biological macromolecules usually depend upon data between ∼25 Å and better than 3 Å resolution. With the aid of information on bond lengths and angles from higher resolution small-molecule data, it is then possible to determine with a fair degree of accuracy the relative positions of atoms, even in the absence of resolved atoms in the electron-density maps.

Combining cryo-EM and crystallographic data can take two forms.

(i) When the sample can be crystallized. The cryo-EM results can be used as an initial molecular-replacement phasing model. The principal difficulty here is the frequently poor overlap between the cryo-EM data (often extending to only 25 Å resolution) and the X-ray data (usually starting at around 30 Å resolution as a result of using unnecessarily large beamstops). If successful phasing can be initiated with the cryo-EM model, it can be used to find heavy atoms, as was the case in the analysis of a ribosome subunit (Ban *et al.*, 1999).

Alternatively, the initial EM phasing can be used to improve and extend crystallographic phases by non-crystallographic symmetry averaging (*e.g.* Speir *et al.*, 1995; Prasad *et al.*, 1999).

(ii) When the sample cannot be crystallized, but sub-components can be crystallized. This will be increasingly the case, as structural biology is starting to investigate larger complexes that are unlikely to fit easily into a crystal lattice (Kostyuchenko *et al.*, 1999), that are too unstable to maintain their structure over the long crystallization time period (Dokland *et al.*, 1997) or are undergoing rapid structural changes in response to functional stimuli. In this case, the crystallographic atomic structure of the component parts can be used to fit into the cryo-EM image in the same way as protein crystallographers depend upon the structure of peptides and specific side-chain conformations derived from small-molecule structures at higher resolution. Probably the first example of this procedure was the study of the interaction of an Fab molecule with a virion (Wang *et al.*, 1992).

Only the second type of problem will be discussed here. An example of what might be possible is our analysis of the complex machinery of the bacterial virus T4. This coliphage has a genome that consists of 169 kilobase pairs with about 272 open reading frames, although only about 40 of these proteins are required for viral assembly. The baseplate is the communication center during infection of the *Escherichia coli* host. We are studying the structures of T4 structural proteins, in particular those of the baseplate, in the anticipation of being able to fit these atomic resolution structures into EM data of the polymorphic baseplate and of the whole virus (Fig. 1).

The method of fitting can be performed either in real or reciprocal space. In general, real-space fitting is especially useful when it is necessary to verify the uniqueness of the final result by exploring all possible angles and orientations. Reciprocal space may be particularly useful as a refinement procedure once a reasonable initial fit has been determined.
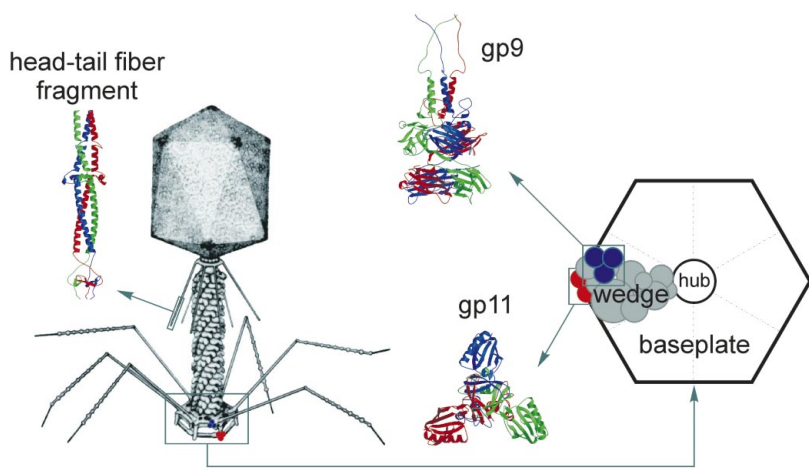
Three examples will be given.

(i) Real-space computerized fitting of the atomic structure of Ross River virus capsid protein into the cryo-EM structure of the whole virus.

(ii) Fitting and reciprocal-space refinement of the known intercellular adhesion molecule-1 (ICAM-1) into the cryo-EM structure of rhinoviruses complexed with its cellular receptor, ICAM-1.

(iii) Real-space based fitting of the homology-modeled poliovirus receptor (CD155) structure into the cryo-EM structure of a poliovirus complexed with CD155.

2. The structure of an alphavirus nucleocapsid core

Alphaviruses are a group of positive-strand RNA viruses that infect mammals and are transmitted by mosquitoes. Ross River virus (RRV), one of 27 known members of the group, infects humans, causing fever, rash and arthritis in peripheral joints. Prior to our investigation (Cheng *et al.*, 1995), it had been shown that alphaviruses had an outer group of 80 glycoprotein spikes arranged with $T = 4$ symmetry (Vogel *et al.*, 1986; Fuller, 1987). These spikes penetrate a lipid membrane derived from the cell's plasma membrane during budding and associate with an inner nucleocapsid core (Fig. 2). Earlier results had variously reported the structure of the inner core to be made of 180 ($T = 3$) or 240 ($T = 4$) protein subunits.

The carboxy-terminal domain (residues 114–264) of the core protein of Sindbis virus (SCP), another alphavirus, had been determined previously (Choi *et al.*, 1991; Tong *et al.*, 1992, 1993). The amino-terminal domain (residues 1–113) is mostly very basic and was expected to be associated internally with the genomic RNA, as has been frequently observed in plant viruses (Rossmann & Johnson, 1989). The cryo-EM structure of the virus (Fig. 3) showed that the core was organized as a series of pentamers and hexamers in a classical $T = 4$ arrangement (Caspar & Klug, 1962). The various X-ray structures of the core carboxy-terminal domain all contained a conserved dimer structure. However, there was no way in which the dimer could be utilized in fitting to the cryo-EM image (Fig. 4), as this would have left large parts of the atoms outside the cryo-EM density and large parts of the cryo-EM density uninterpreted. Thus, it was necessary to establish what was the best fit of the monomer subunit into the density while maintaining the $T = 4$ symmetry. An initial coarse three-dimensional orientational search was followed by a systematic 'climb' procedure to find the best fit using three orientational and three positional refinable parameters.

The quality of fit Q was defined as the sum of the electron density over all the predicted atomic sites (main chain and side chain) in the independent monomer. The density at each atom position was determined by an eight-point

Figure 1
Diagrammatic representation of the T4 bacteriophage, adapted from Eiserling & Black (1994), showing in color the currently known structures of its components. These are the head–tail fiber fragment (fibritin) (Tao *et al.*, 1997), gp9 (Kostyuchenko *et al.*, 1999) and gp11 (Leiman *et al.*, 2000).

interpolation within the RRV cryo-EM map. The number of steric clashes between neighboring monomers was assessed in terms of the number of grid points closest to each atom that were also closest to atoms in a neighboring monomer. No attempt was made to avoid clashes, but steric hindrance was checked numerically and graphically and found to be few for orientations that gave better fits of the structure to the density.
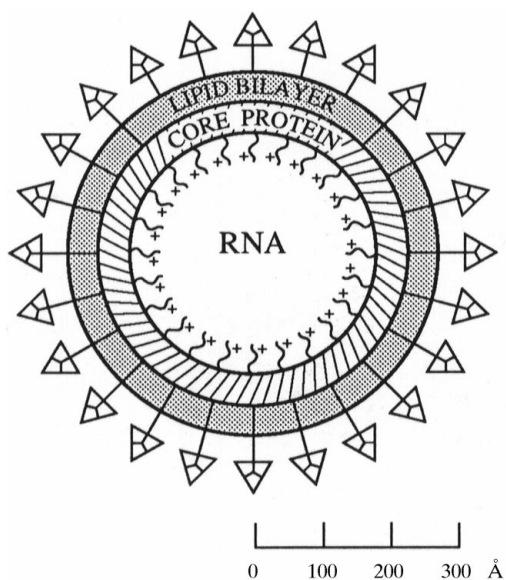

Figure 2
Diagrammatic representation of an alphavirus showing a cross-section of the Sindbis virus structure. The virion has 80 trimeric glycoprotein spikes arranged in a $T = 4$ surface lattice. The spikes penetrate a lipid bilayer and are anchored in the core protein capsid. The basic N-terminal segments of the SCP are associated with the Sindbis virus genomic RNA. [Reprinted by permission from Choi *et al.* (1991), *Nature (London)*, **354**, 37–43. Copyright (1991) Macmillan Magazines Ltd.]
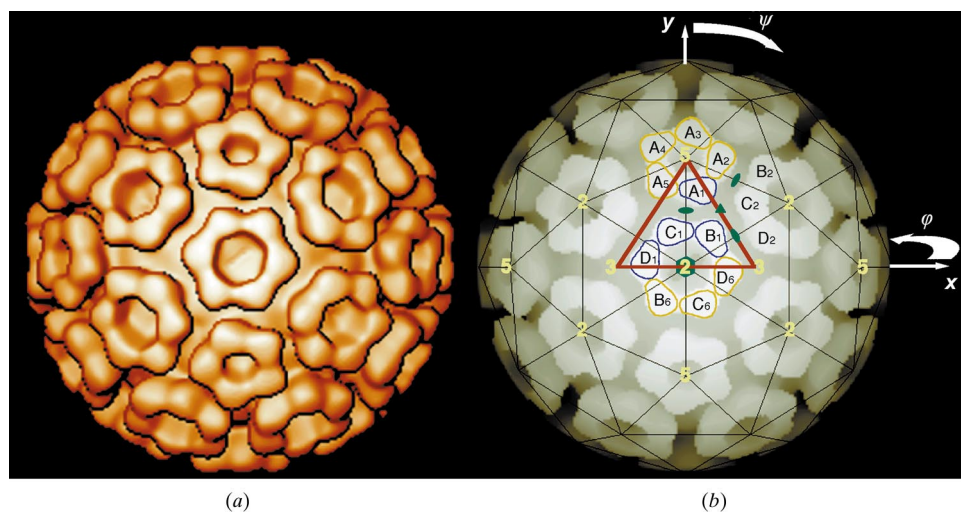
The C-terminal domain of the Sindbis virus capsid protein (SCP) monomer was rotated from its standard orientation by the Eulerian angles $(\theta_1, \theta_2, \theta_3)$. The center of gravity of the rotated monomer was then placed onto the approximate center (x, y, x) (Fig. 3) of one lump (A_1 in Fig. 2) of the RRV core electron density. Icosahedral symmetry-related monomers within a pentamer were generated to give A_2, A_3, A_4 and A_5 by means of the nearest fivefold icosahedral axis. The monomer at A_1 was rotated into the equivalent monomer position C_1 by means of the nearest quasi-twofold axis, assumed to be radial. Monomer C_1 was then reproduced to give C_1, D_1, B_6, C_6, D_6 and B_1 by virtue of the appropriate quasi-sixfold axis (Fig. 4; this quasi-sixfold axis is coincident with an icosahedral twofold axis). The position of the quasi-twofold axis between the pentamer and hexamer defined by the polar angles ψ, φ is not known precisely and therefore was systematically adjusted to obtain the best fit. In principle, it should be necessary to refine separately the orientational and positional parameters of the four quasi-equivalent monomers A_1, B_1, C_1 and D_1. However, to keep the number of refinable parameters to a minimum and hence to increase the significance of the resultant fit, the quasi-sixfold axis was treated as an exact sixfold axis. This gave six refinable parameters for the monomer at A_1 [three positional (x, y, z) and three orientational (θ_1, θ_2, θ_3)], alternated with the refinement of two parameters (ψ, φ) defining the quasi-twofold axis orientation, as well as the hand of the map relative to the known absolute hand of the monomer structure. The search for the best fit was pursued over all Eulerian angles ($0 \leq \theta_1 < 2\pi$; $0 \leq \theta_2 < \pi$; $0 \leq \theta_3 < 2\pi$) in increments of 30 or 20°, retaining the top 100 fits.

Each of the best fits was then subjected to an exploration by 'climbing' to the next best fit. In this, the first angle, θ_1, was incremented by a given amount, $\Delta\theta_1$. If the fit Q improved, θ_1 was further incremented stepwise until a maximum had been reached. If the fit became worse, θ_1 was systematically decreased stepwise until the maximum of Q was reached. After adjusting θ_1, θ_2 was adjusted, followed by θ_3 and then by each of the positional parameters x, y and z. The procedure was then repeated for each of the six parameters in turn until no better fit could be found. At this juncture, the step increments were divided by a factor of ten and the procedure was repeated again. Eventually, the best local fit was accepted when the angular increments had decreased from an initial 10° to a final 0.25° and the positional increments had decreased from an initial 5 Å to 0.5 Å.

(a)　　　　　　　　　　　　　(b)

Figure 3
Structure of the RRV nucleocapsid core. (a) Surface-shaded twofold view of the virion reconstruction after all densities outside a radius of 210 Å were removed. The pentamer and hexamer arrangements of the nucleocapsid subunits are visible. (b) Depth-cued representation of the map in (a) with superimposed $T = 4$ lattice. One icosahedral asymmetric unit lies within the triangle (red) bounded by two threefold axes and one fivefold axis. It includes four monomers (A_1, B_1, C_1 and D_1). Yellow numbers identify the icosahedral axes and green symbols label the pseudoaxes. The polar angles are defined as in Rossmann & Blow (1962). [Reprinted by permission from Cheng *et al.* (1995), *Cell*, **80**, 621–630. Copyright (1995) Cell Press.]

Table 1
Different ICAM-1 fragments used in the cryo-EM reconstructions.

Type	Residues	Domains	Mutations	Glycosylation	Reference to crystal structure
1	1–185	D1–D2	Wild type	Full complex	Kolatkar et al. (1999)
2	1–185	D1–D2	N103→Q N118→Q N156→Q	Reduced complex Site N175 only	Bella et al. (1998)
3	1–190	D1–D2	Wild type	Full, high mannose	Casasnovas et al. (1998)
4	1–453	D1–D5	Wild type	Full complex	Kolatkar et al. (1999)

Many of the final best fits were the same. That is, although the climb started at a different point, the best orientation and position were found to be the same, independent of the starting parameters. The original top 100 fits, obtained by the complete Eulerian angle search, reduced themselves down to about ten distinct best fits. Therefore, two other criteria of quality of fit were developed to differentiate the remaining best fits. A count was kept of atoms that were in steric collision between monomers (nclash) and of the number of atoms placed into negative density (−den). Although neither measurement was able to select clearly among the remaining best fits, a product of (nclash) × (−den) appeared to be the most sensitive quantity.

The best position of the quasi-twofold position was found to be at $\psi = 73$, $\varphi = -90.5°$. At this position, $\theta_1 = 102.5$, $\theta_2 = 8.0$, $\theta_3 = 22.0°$, $x = 9.4$, $y = 73.2$, $z = 166.5$ Å relative to the SC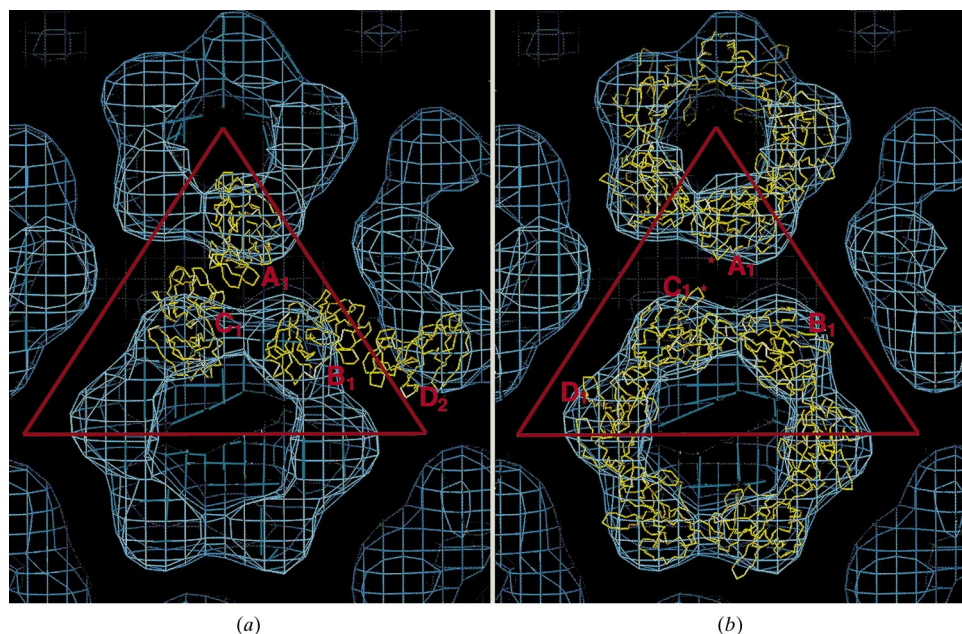P coordinates deposited with the Protein Data Bank (PDB code 2snv). These parameters produced a remarkably good fit to the cryo-EM density (Fig. 4), giving little steric interference between neighboring monomers. Subsequently, these results were confirmed by mutational analysis of the residues in the monomer interface (Tellinghuisen et al., 1999), by the proposed site of association with the transmembrane glycoprotein spikes (Lee et al., 1996; Skoging et al., 1996) and by the higher resolution cryo-EM analysis of the homologous Semliki Forest virus (Mancini et al., 2000).

3. Rhinoviruses complexed with their receptor

Human rhinoviruses (HRVs) are small icosahedral viruses with a diameter of ∼300 Å. They are composed of a protein shell that encapsidates a positive RNA strand of ∼7000 bases. The capsid is built from 60 copies of four viral proteins, VP1, VP2, VP3 and a small internal protein VP4 (Rossmann et al., 1985). About 90 of the known 100 or so serotypes utilize the cell-surface glycoprotein ICAM-1 or CD54 as receptor (Greve et al., 1989; Staunton et al., 1989). The structures of a number of HRVs have been determined crystallographically [HRV14 (Rossmann et al., 1985) and HRV16 (Oliveira et al., 1993), both belonging to the major group of HRVs]. The structure of ICAM-1 also has been determined crystallographically (Bella et al., 1998; Casasnovas et al., 1998). However, the structure of the complex of HRVs with ICAM-1 has been determined by cryo-EM only (Olson et al., 1993; Kolatkar et al., 1999). Indeed, it may be difficult to determine the complex crystallographically because binding of ICAM-1 with HRV is not only a recognition event but also initiates the virus break-up (uncoating) to release the RNA into the host cytoplasm. Thus, the cryo-EM experiments require careful timing and temperature control of the incubation of HRV with ICAM-1.

ICAM-1 belongs to the immunoglobulin (Ig) superfamily of cell-adhesion molecules, whose defining characteristic is an extracellular tandem of Ig-like domains (Fig. 4). ICAM-1 normally functions to promote intercellular adhesion and signaling in response to inflammation. The amino-terminal domain, D1, is responsible for binding to the normal cellular ligand as well as to the viral pathogen. Cryo-EM studies of HRV–ICAM-1 complexes have been performed mainly with a D1D2 fragment of ICAM-1. This fragment has four potential N-glycosylation sites. The original crystal structure was solved with

Figure 4
Structure of the RRV nucleocapsid core. (a) Placement of the crystallographic SCP dimer C$_1$ structure (yellow) into the observed electron density (blue). Neither the A$_1$C$_1$ nor the B$_1$D$_2$ dimers fit into the observed density. (b) In contrast with (a), the SCP monomer (yellow) fits into the RRV nucleocapsid density (blue) quite well. In agreement with cross-linking results, the Lys250 residues (denoted by asterisks) on adjacent capsomeres are ∼12 Å apart. [Reprinted by permission from Cheng et al. (1995), Cell, **80**, 621–630. Copyright (1995) Cell Press.]

three out of the four glycosylation sites deglycosylated (Bella *et al.*, 1998), but the crystal structure of a fully glycosylated form was also subsequently determined (Kolatkar *et al.*, 1999; Table 1). The various structures of ICAM-1 showed that the elbow angle between the two domains is flexible, but that the variation occurs mostly in one plane (Fig. 5). This restricted variation exists even though the crystal packing environments and glycosylation properties differ. The lack of a spacer region between the two domains, which permits several close interactions, appears to be the basis of the restricted flexibility.

One conspicuous feature in all the cryo-EM reconstructions of HRV–ICAM-1 complexes is the presence of protuberances or 'lumps' on the ICAM-1 spikes. Such lumps are approximately perpendicular to the long axis of the ICAM-1 molecules (Kolatkar *et al.*, 1999; Fig. 6; Table 1). Biochemical data had indicated that domain D1 contains the virus-binding surface. With this orientation of the ICAM-1 D1D2 fragment, all the lumps occur in the D2 domain, consistent with D1 having no potential glycosylation sites and D2 having four such sites. Hence, the lumps represent the carbohydrate moieties of the ICAM-1 fragments.

A difference cryo-EM density map computed from the HRV16-type 1 (fully glycosylated) and HRV16-type 2 (partially glycosylated) ICAM-1 reconstructions showed four areas of significant difference density (Fig. 7). Three of these sites were consistent with three of the density lumps putatively assigned to glycosylation sites. The remaining site, present in both reconstructions, disappears in the difference map and
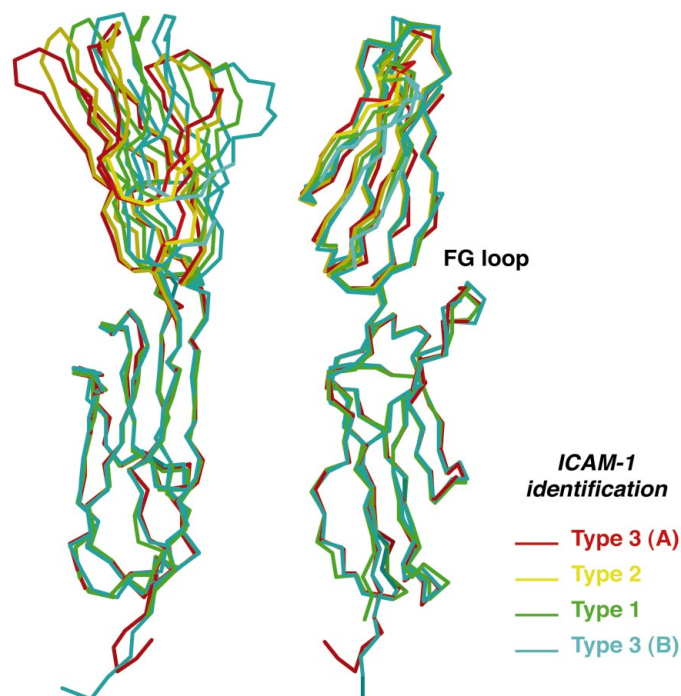
thus corresponds to the only common glycosylation site at Asn175. Thus, all four density lumps corresponded to carbohydrate moieties on D2. A small volume of difference density near the virus surface (Fig. 7a) is not compatible with a glycosylation site and might reflect a small change in orientation of ICAM-1 binding between the two complexes.

Appropriately glycosylated models of type 1 ICAM-1 with various elbow angles were fitted by eye into the cryo-EM electron-density maps. Each of the models of the HRV14-type 1 and HRV16-type 1 reconstructions were refined subsequently as rigid bodies in reciprocal space with respect to difference maps obtained by subtraction of the HRV contribution from the cryo-EM reconstructed density of the complexes (J. Bella & M. G. Rossmann, manuscript in preparation) using the program *X-PLOR* (Brünger, 1992). The refined fits of the ICAM-1 D1D2 models into the cryo-EM density maps were consistent with the predicted positions of glycoproteins. The reproducibility of the refinement process was accurate to better than 1.5 Å between equivalent atoms. Thus, the atomic interaction between virus and receptor can be accepted subject to the assumption that the crystallographically determined structures of the virus and of the receptor fragment are not altered substantially by the formation of the complex.
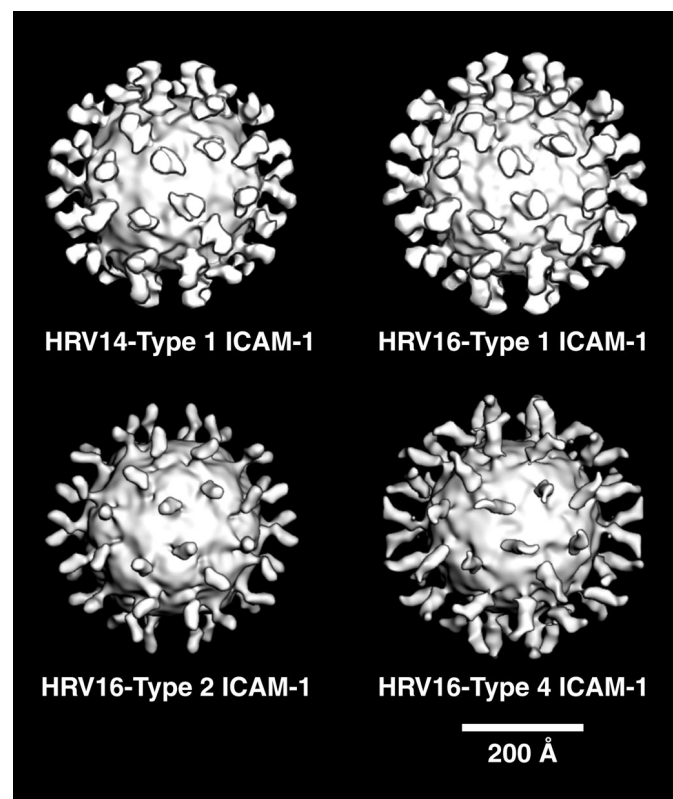

Figure 5
Variation of the elbow angle between domains D1 and D2 in the crystallographically determined structures of ICAM-1. The D2 domains have been aligned in two orthogonal views. [Reprinted by permission from Kolatkar *et al.* (1999), *EMBO J.* **18**, 6249–6259. Copyright (1999) European Molecular Biology Organization.]

Figure 6
Cryo-EM image reconstructions of complexes between HRV14 or HRV16 and different soluble fragments of ICAM-1 (see Table 1 for nomenclature identifying the ICAM-1 fragments). The occupancy of the type 2 and type 4 ICAM-1 fragments is lower, causing them to look thinner at the contour level appropriate for the virus density. [Reprinted by permission from Kolatkar *et al.* (1999), *EMBO J.* **18**, 6249–6259. Copyright (1999) European Molecular Biology Organization.]
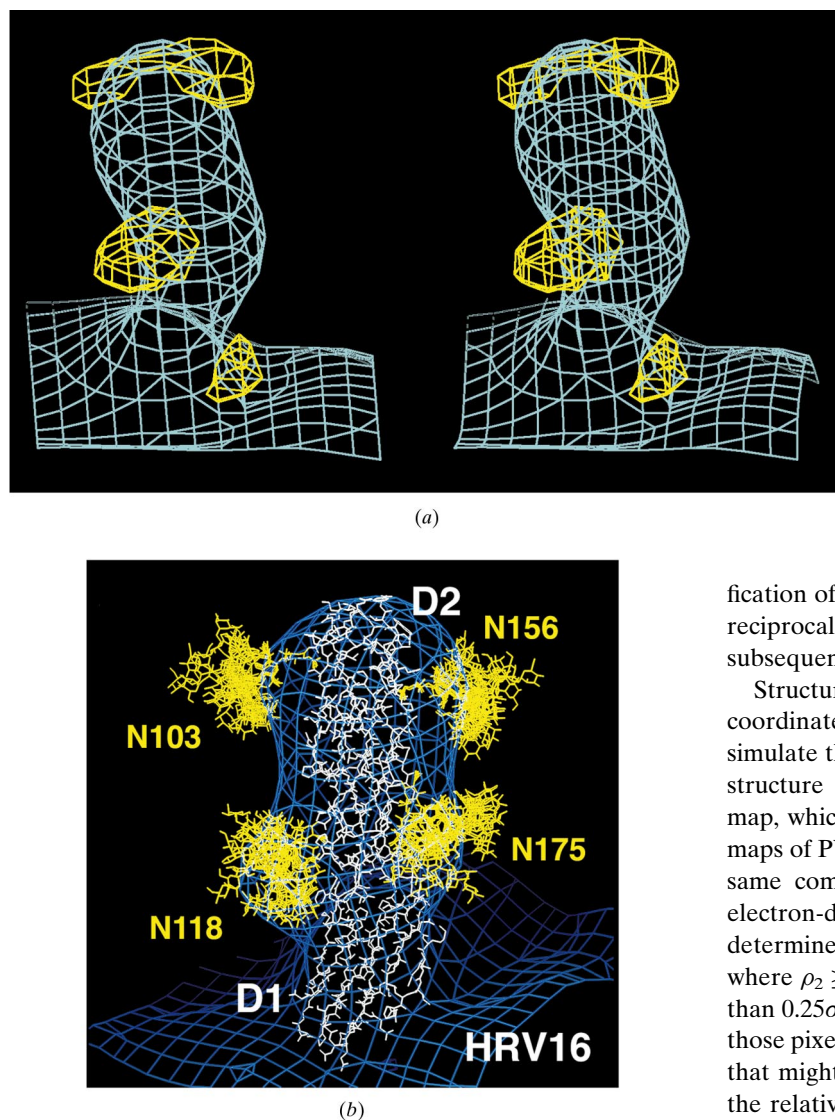
(a)

(b)

Figure 7
(*a*) Stereo diagram of a portion of the HRV16–type 2 ICAM-1 cryo-EM reconstruction corresponding to the density (light blue) for the two-domain ICAM-1 fragment. Superimposed in yellow is the difference density map between HRV16–type 1 (fully glycosylated) and HRV16–type 2 (mostly deglycosylated) reconstructions. The density in the HRV16–type 2 reconstruction has been suitably scaled to account for the lower occupancy of the ICAM-1 fragments. (*b*) Fitting of the refined type 1 ICAM-1 model into the cryo-EM reconstruction of the HRV16–type 1 ICAM-1 reconstruction. The protein is represented in white and the disordered carbohydrates are represented by an ensemble of conformations (yellow). [Reprinted by permission from Kolatkar *et al.* (1999), *EMBO J.* **18**, 6249–6259. Copyright (1999) European Molecular Biology Organization.]

4. Polioviruses complexed with their receptor

Poliovirus (PV) is the etiological agent of poliomyelitis, a human disease of the central nervous system. There are three known serotypes (PV1, PV2 and PV3), all of which recognize the same cellular poliovirus receptor (CD155; Mendelsohn *et al.*, 1989; Koike *et al.*, 1990; Wimmer *et al.*, 1994). Polioviruses, like human rhinoviruses, are members of the picornavirus family and have a similar structure (Hogle *et al.*, 1985).

CD155, like ICAM-1, is a membrane-anchored glycoprotein whose extracellular region has three Ig-like domains and whose glycosylated amino-terminal domain D1 contains the virus-recognition site (Fig. 4). Unlike ICAM-1, no crystal structure of CD155 has been determined yet, making it necessary to use homologous structures for modeling into the cryo-EM reconstruction of the PV1–CD155 complex (Fig. 8).

The X-ray crystal structure of PV1 (Hogle *et al.*, 1985; PDB code 2plv) was used as a reference for scaling the pixel size and density of the cryo-EM reconstructions (He *et al.*, 2000). This is a general problem encountered whenever it is necessary to combine cryo-EM data with crystallographic data owing to uncertainty ($\sim\pm 5\%$) of the degree of magnification of the EM image. Originally, this was accomplished in reciprocal space, but a sensitive real-space method was subsequently employed.

Structure factors were calculated from the X-ray coordinates of PV1 with a 'temperature' factor of 1000 Å^2 to simulate the cryo-EM reconstruction limit of about 22 Å. The structure factors were used to compute an electron-density map, which was taken as a standard for placing the cryo-EM maps of PV1 and of PV1 complexed with CD155 (ρ_1) onto the same common scale as the crystallographically determined electron-density map (ρ_2). The scale factors (a and b) were determined by minimizing $\sum[\rho_1 - (a + b\rho_2)]^2$ in the region where $\rho_2 \geq 0.25\sigma(\rho_2)$. The condition that ρ_2 should be greater than $0.25\sigma(\rho_2)$ assured that the comparison was made only for those pixels corresponding to the viral density, not those pixels that might represent the receptor or background in ρ_1. First, the relative pixel size was established for each cryo-EM map by comparing density in a shell between radii 132 and 144 Å (Table 2). Then, the unknown hand of the EM map was determined with respect to the known hand of the X-ray map (Table 3), although good evidence was also seen by visual inspection of the asymmetric canyon (He *et al.*, 2000). The results were then used for scaling the two cryo-EM maps together. The agreement between these two maps is best in a shell between a radius of 132 and 144 Å (Table 4). The presence of the receptor outside about 144 Å in one of the maps and of RNA inside about 132 Å in both of the maps interferes with the map equivalencing. The scaled maps were subtracted to give a difference map representing the density for the CD155 (He *et al.*, 2000).

In the absence of an atomic structure of CD155, it was necessary to find the closest available homologous models from the PDB. A *BLAST* search (Altschul *et al.*, 1990) with respect to the CD155 amino-acid sequence was conducted for each separate domain against all Ig superfamily entries in the PDB. The myelin membrane-adhesion molecule 'human protein zero' (PDB code 1neu) gave the best fit for domain D1, which corresponded to an Ig-like variable (V) domain. An Ig light-chain constant domain derived from an anti-idiotype

Fab–Fab complex (PDB code 1cic) was the best fit for domain D2 and corresponded to a constant Ig-like C2 domain. The insect immune protein (PDB code 1bih) was the best fit for domain D3.

There was considerable uncertainty about the choice of rotational orientation relative to the domain's long axis when fitting these domain structures by eye to the cryo-EM difference map. This problem was solved by noting that the hinge between the V and C domains in Ig superfamily multidomain structures such as CD2 and CD4 tends to bend to keep the same domain surfaces in contact. The elbow angle between domains was clearly evident in the cryo-EM difference map. The structure of the first two domains of CD2 (PDB code 1hnf) could be fitted readily into the corresponding difference map density with only a small adjustment of the elbow angle.

A least-squares procedure (Rao & Rossmann, 1973) was then used to superimpose the CD155 homologues of D1 and D2 onto the CD2 C^{α} atoms. Subsequently, D3 was fitted to the density by limiting the position of its amino end close to the carboxy end of D2.

Another difference map was calculated by subtracting the calculated contribution of the PV and CD155 atomic coordinates from the cryo-EM map of the complex. This difference map (Fig. 8d) identified the density corresponding to the glycosylation sites of CD155. Because of the increasing flexibility of the CD155 receptor at progressively larger distances from the virus surface, the scaling of the CD155 model density to the EM reconstruction was performed separately for each of the three Ig-like domains. (This separate scaling might have been avoided by using progressively increasing temperature
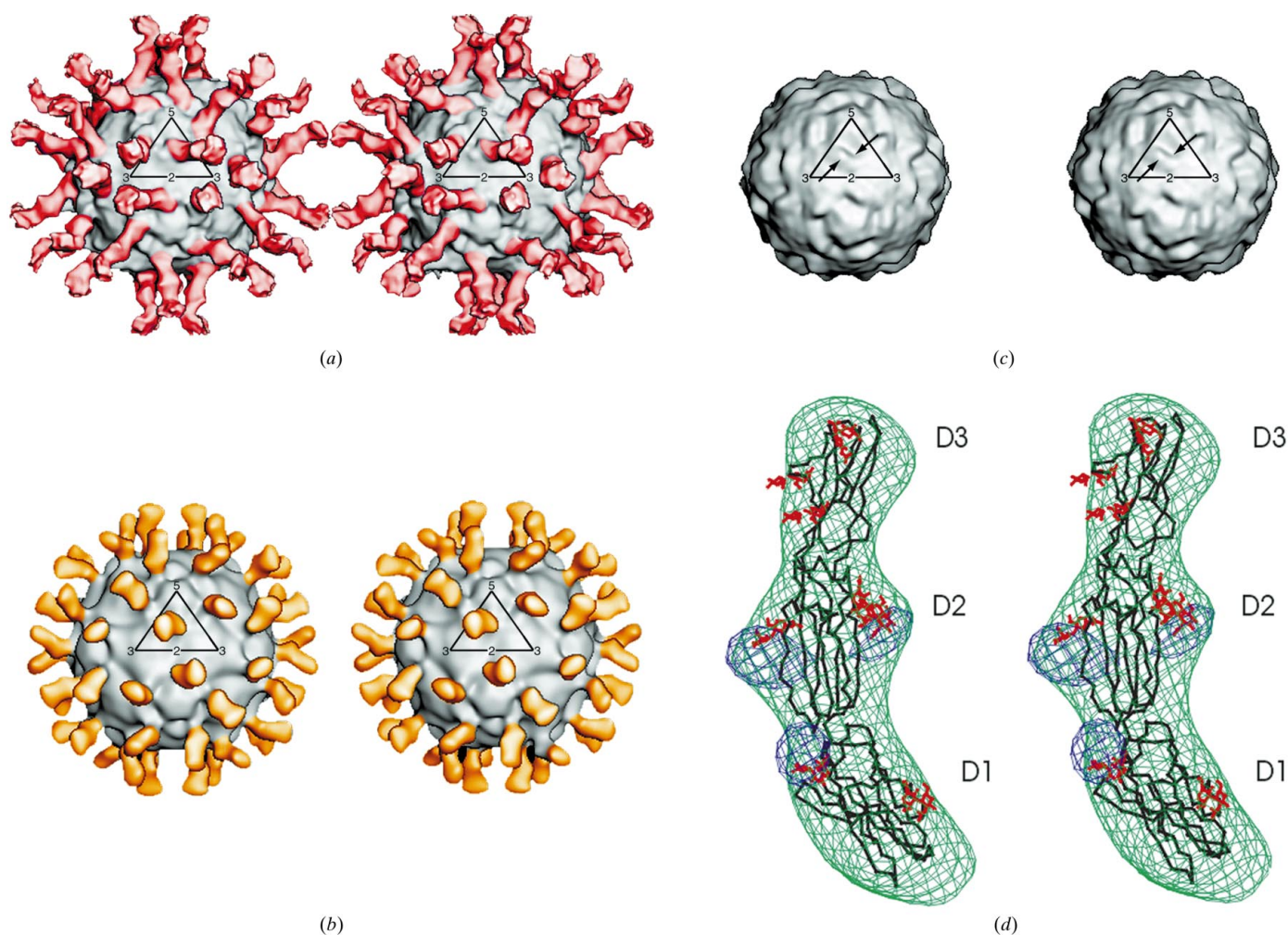

Figure 8
(a) Stereoview of the cryo-EM reconstruction showing the complex of PV1 with human CD155. The outline of one geometric icosahedral asymmetric unit is shown. Note that the receptor leans toward the southeast. (b) Stereoview of the cryo-EM reconstruction showing the complex of HRV16 with its ICAM-1 receptor. The outline of one geometric icosahedral asymmetric unit is shown. Note that the receptor leans towards the southwest. (c) Stereoview of a cryo-EM reconstruction of PV1, also showing the geometric icosahedral asymmetric unit. Note the asymmetric shape of the canyon with its most southerly point slightly east of center (arrow) and the smaller peak southwest of the canyon (arrow). These features establish the correct hand of the reconstructions in (a) and (b) and are consistent with the X-ray results, where the absolute hand is known. (d) Density (green) representing one CD155 molecule (black) fitted with the C^{α} backbone structure of the closest homologous structure found in the PDB for each of the three domains. Shown also is the difference map (blue) between the cryo-EM density and the unglycosylated CD155 model to show the sites of glycosylation. The potential glycosylation sites are shown on the CD155 backbone (red). [Reprinted by permission from He *et al.* (2000), *Proc. Natl Acad. Sci. USA*, **97**, 79–84. Copyright (2000) National Academy of Sciences, USA.]

Table 2
Determination of EM magnification scale.

Comparison of the PV1–CD155 cryo-EM map and of the PV1 cryo-EM map with the PV1 map based on crystallographic model coordinates. Density was compared in a shell between 132 and 144 Å radius. Numbers in bold establish the distance between pixels in the cryo-EM reconstructions, as determined by comparison with the X-ray results.

	Correlation coefficient†	
EM pixel size (Å)	(PV1–CD155) map	(PV1) map
3.300	0.7326	0.8660
3.325	0.7764	0.8994
3.350	0.8299	0.9283
3.375	0.8719	0.9490
3.400	0.8993	0.9595
3.425	**0.9105**	**0.9607**
3.450	0.9044	0.9527
3.475	0.8679	0.9277
3.500	0.8096	0.8878
3.525	0.7343	0.8312
3.550	0.6378	0.7471
3.575	0.5263	0.6302
3.600	0.4081	0.5024

† Correlation coefficient $= \sum(\langle\rho_1\rangle - \rho_1)(\langle\rho_2\rangle - \rho_2) / [\sum(\langle\rho_1\rangle - \rho_1)^2 \sum(\langle\rho_2\rangle - \rho_2)^2]^{1/2}$. Pixels with $\rho_2 < 0.25\sigma(\rho_2)$ were omitted.

Table 3
Determination of hand by comparison with the X-ray map.

Shown are the results for the best pixel size determined from density in a shell between 132 and 144 Å radii.

(*a*) Comparison of the PV1–CD155 EM map with the PV1 X-ray map (pixel size = 3.430 Å).

Hand	EM pixel size (Å)	Correlation coefficient†	Hand correlation coefficient‡
+1	3.425	0.910	0.664
−1	3.450	0.897	−0.734

(*b*) Comparison of the PV1 EM map with the PV1 X-ray map (pixel size = 3.430 Å).

Hand	EM pixel size (Å)	Correlation coefficient†	Hand correlation coefficient‡
+1	3.425	0.961	0.852
−1	3.450	0.921	−0.832

(*c*) Comparison of the PV1–CD155 EM map with the PV1 EM map (pixel size = 3.425 Å).

Hand	EM pixel size (Å)	Correlation coefficient†	Hand correlation coefficient‡
+1	3.425	0.968	0.560
−1	3.425	0.940	−0.561

† As defined in Table 2. ‡ Hand correlation coefficients $= \sum[(\rho_1^+ - \rho_1^-)(\rho_2^+ - \rho_2^-)] / [\sum(\rho_1^+ - \rho_1^-)^2]^{1/2}[\sum(\rho_2^+ - \rho_2^-)^2]^{1/2}$, where ρ_1 and ρ_2 are densities in maps 1 and 2, respectively, and ρ^+ and ρ^- are densities related by the mirror plane at (xyz) and $(x\bar{y}z)$, respectively.

factors along the length of the CD155 model in calculating the structure factors.) The results of the modeling were verified by observing that the difference densities attributed to carbohydrate moieties were juxtaposed to potential glycosylation sites in CD155 (Fig. 8*d*).

Table 4
Comparison of maps in successive radial shells.

(*a*) Comparison of the PV1–CD155 EM map with the PV1 X-ray map.

Shell radius (Å)	108–120	120–132	132–144	144–156
Number of pixels†	3918	22203	17914	3424
Correlation coefficient‡	0.1881	0.2662	0.9105	0.9860
Hand correlation coefficient§	−0.1282	0.0897	0.6642	0.0598

(*b*) Comparison of the PV1 EM map with the PV1 X-ray map.

Shell radius (Å)	108–120	120–132	132–144	144–156
Number of pixels†	2164	16183	12928	2370
Correlation coefficient‡	0.1360	0.0014	0.9607	0.9971
Hand correlation coefficient§	−0.3439	0.1040	0.8518	0.9742

(*c*) Comparison of the PV1–CD155 EM map with the PV1 EM map.

Shell radius (Å)	108–120	120–132	132–144	144–156
Number of pixels†	1530	16977	23474	2938
Correlation coefficient‡	0.8813	0.5728	0.9652	0.9912
Hand correlation coefficient§	0.2193	0.4730	0.5217	−0.1060

† Pixels selected as described in the text. ‡ As defined in Table 2. § As defined in Table 3.

5. Conclusions

The use of atomic resolution crystallographic data can greatly enhance the information available in lower resolution cryo-EM image reconstructions. There is a close parallel in the use of small-molecule high-resolution data that must be assumed for the interpretation of lower resolution macromolecular electron-density maps. Whereas the interpretation of protein and nucleic acid maps depends upon knowledge of bond lengths, bond angles and dihedral angles, the interpretation of cryo-EM maps depends upon the knowledge of whole domain structures and their conformational relationships to other domains. The process of fitting maps and refinement can be performed in either real or reciprocal space. Methods for real-space fitting, as described in this paper, have been developed in the program *EMfit*, written in Fortran. The source code is available free from the author.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B. & Steitz, T. A. (1999). *Nature (London)*, **400**, 841–847.

Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M. & Rossmann, M. G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4140–4145.

Böttcher, B., Wynne, S. A. & Crowther, R. A. (1997). *Nature (London)*, **386**, 88–91.

Brünger, A. T. (1992). *X-PLOR. Version* 3.1. *A System for X-ray Crystallography and NMR*. New Haven & London: Yale University Press.

Casasnovas, J. M., Stehle, T., Liu, J., Wang, J. & Springer, T. A. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4134–4139.

Caspar, D. L. D. & Klug, A. (1962). *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. & Baker, T. S. (1995). *Cell*, **80**, 621–630.

Choi, H. K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G. & Wengler, G. (1991). *Nature (London)*, **354**, 37–43.

Conway, J. F., Cheng, N., Zlotnick, A., Wingfield, P. T., Stahl, S. J. & Steven, A. C. (1997). *Nature (London)*, **386**, 91–94.

Dokland, T., McKenna, R., Ilag, L. L., Bowman, B. R., Incardona, N. L., Fane, B. A. & Rossmann, M. G. (1997). *Nature (London)*, **389**, 308–313.

Eiserling, F. A. & Black, L. W. (1994). *Molecular Biology of Bacteriophage T4*, edited by J. D. Karam, pp. 209–212. Washington, DC: American Society for Microbiology.

Fuller, S. D. (1987). *Cell*, **48**, 923–934.

Greve, J. M., Davis, G., Meyer, A. M., Forte, C. P., Yost, S. C., Marlor, C. W., Kamarck, M. E. & McClelland, A. (1989). *Cell*, **56**, 839–847.

He, Y., Bowman, V. D., Mueller, S., Bator, C. M., Bella, J., Peng, X., Baker, T. S., Wimmer, E., Kuhn, R. J. & Rossmann, M. G. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 79–84.

Hogle, J. M., Chow, M. & Filman, D. J. (1985). *Science*, **229**, 1358–1365.

Koike, S., Horie, H., Ise, I., Okitsu, A., Yoshida, M., Iizuka, N., Takeuchi, K., Takegami, T. & Nomoto, A. (1990). *EMBO J.* **9**, 3217–3224.

Kolatkar, P. R., Bella, J., Olson, N. H., Bator, C. M., Baker, T. S. & Rossmann, M. G. (1999). *EMBO J.* **18**, 6249–6259.

Kostyuchenko, V. A., Navruzbekov, G. A., Kurochkina, L. P., Strelkov, S. V., Mesyanzhinov, V. V. & Rossmann, M. G. (1999). *Structure*, **7**, 1213–1222.

Lee, S., Owen, K. E., Choi, H. K., Lee, H., Lu, G., Wengler, G., Brown, D. T., Rossmann, M. G. & Kuhn, R. J. (1996). *Structure*, **4**, 531–541.

Leiman, P. G., Kostyuchenko, V. A., Shneider, M. M., Kurochkina, L. P., Mesyanzhinov, V. V. & Rossmann, M. G. (2000). In the press.

Mancini, E. J., Clarke, M., Gowen, B., Rutten, T. & Fuller, S. D. (2000). *Mol. Cell*, **5**, 255–266.

Mendelsohn, C. L., Wimmer, E. & Racaniello, V. R. (1989). *Cell*, **56**, 855–865.

Oliveira, M. A., Zhao, R., Lee, W., Kremer, M. J., Minor, I., Rueckert, R. R., Diana, G. D., Pevear, D. C., Dutko, F. J., McKinlay, M. A. & Rossmann, M. G. (1993). *Structure*, **1**, 51–68.

Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S. & Rossmann, M. G. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 507–511.

Prasad, B. V. V., Hardy, M. E., Dokland, T., Bella, J., Rossmann, M. G. & Estes, M. K. (1999). *Science*, **286**, 287–290.

Rao, S. T. & Rossmann, M. G. (1973). *J. Mol. Biol.* **76**, 241–256.

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. & Vriend, G. (1985). *Nature (London)*, **317**, 145–153.

Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.

Rossmann, M. G. & Johnson, J. E. (1989). *Annu. Rev. Biochem.* **58**, 533–573.

Skoging, U., Vihinen, M., Nilsson, L. & Liljeström, P. (1996). *Structure*, **4**, 519–529.

Speir, J. A., Munshi, S., Wang, G., Baker, T. S. & Johnson, J. E. (1995). *Structure*, **3**, 63–78.

Staunton, D. E., Merluzzi, V. J., Rothlein, R., Barton, R., Marlin, S. D. & Springer, T. A. (1989). *Cell*, **56**, 849–853.

Tao, Y., Strelkov, S. V., Mesyanzhinov, V. V. & Rossmann, M. G. (1997). *Structure*, **5**, 789–798.

Tellinghuisen, T. L., Hamburger, A. E., Fisher, B. R., Ostendorp, R. & Kuhn, R. J. (1999). *J. Virol.* **73**, 5309–5319.

Tong, L., Choi, H. K., Minor, W. & Rossmann, M. G. (1992). *Acta Cryst.* A**48**, 430–442.

Tong, L., Wengler, G. & Rossmann, M. G. (1993). *J. Mol. Biol.* **230**, 228–247.

Vogel, R. H., Provencher, S. W., von Bonsdorff, C. H., Adrian, M. & Dubochet, J. (1986). *Nature (London)*, **320**, 533–535.

Wang, G., Porta, C., Chen, Z., Baker, T. S. & Johnson, J. E. (1992). *Nature (London)*, **355**, 275–278.

Wimmer, E., Harber, J. J., Bibb, J. A., Gromeier, M., Lu, H. H. & Bernhardt, G. (1994). *Cellular Receptors for Animal Viruses*, edited by E. Wimmer, pp. 101–127. Cold Spring Harbor Laboratory: Cold Spring Harbor Laboratory Press.

Wynne, S. A., Crowther, R. A. & Leslie, A. G. W. (1999). *Mol. Cell*, **3**, 771–780.