Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural analyses of *Phycodnaviridae* and *Iridoviridae*

Alan A. Simpson,[a] Narayanasamy Nandhagopal,[a] James L. Van Etten[b] and Michael G. Rossmann[a]*

[a]Department of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, IN 47907-2054, USA, and [b]Department of Plant Pathology, University of Nebraska, Lincoln, NE 68583-0722 and Nebraska Center of Virology, University of Nebraska, Lincoln, NE 68588, USA

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

Received 29 August 2003
Accepted 7 October 2003

The *Phycodnaviridae*, *Iridoviridae* and related viruses, with diameters of 1500–2000 Å, are formed from large trigonal arrays of hexagonally close-packed capsomers forming the faces of icosahedra [Yan *et al.* (2000), *Nature Struct. Biol.* **7**, 101–103; Nandhagopal *et al.* (2002), *Proc. Natl Acad. Sci. USA*, **99**, 14758–14763]. Caspar and Klug predicted that such structures could be assembled from hexameric capsomers [Caspar & Klug (1962), *Cold Spring Harbor. Symp. Quant. Biol.* **27**, 1–24], as was subsequently found in numerous icosahedral viruses. During the course of evolution, some viruses, including the virus families mentioned above, replaced hexameric capsomers with pseudo-hexameric trimers by gene duplication. In large dsDNA icosahedral viruses, the capsomers are organized into 'pentasymmetrons' and 'trisymmetrons'. The interactions between the trimeric capsomers can be divided into three groups, one between similarly oriented trimers and two between oppositely oriented trimers (trimers related by an approximately sixfold rotation). The interactions within a trisymmetron belong to the first class, whereas those between trisymmetrons and within the pentasymmetron are of the other two types. Knowledge of these distances permits a more accurate fitting of the atomic structure of the capsomer into the cryo-electron microscopy (cryoEM) reconstruction of the whole virus. The adoption of pseudo-hexagonal capsomers places these viruses into a subset of the Caspar and Klug surface lattices.

1. The changing face of structural biology

It does not seem possible that *Acta Crystallographica Section D* started publication only ten years ago. Since then, so much has changed in the world of structural biology. Twenty years ago, there were only about 27 new structures deposited per year in the Protein Data Bank (PDB), but by 1993 the number had increased to 792 and the exponential increase had only just begun. Last year, there were not only 3381 new structures, but their complexity had increased from an average of 2047 non-H atoms per crystallographic asymmetric unit per deposition in 1983 to 2474 in 1993 and 5444 in 2002. This transformation has been accompanied by a radical change of attitudes. Whereas the public release of coordinates and structure factors was entirely dependent on the personal preference of each investigator in the 1970s and 1980s, today almost no journal will publish a new structure without the release of coordinates for deposition and some granting agencies are starting to insist on depositions prior to continuation of financial awards. Ten years ago the idea of compulsory depositions was a subject of hot debate, whereas today not only is deposition taken for granted, but priority is

often measured by the date of the PDB accession. Another major 'cultural' change is that most of the younger crystallographers are motivated by their interests in cell biology, enzymology, virology and other related subjects as opposed to physics, mathematics or perhaps chemistry. This has happened as a consequence of greatly improved equipment, such as incredibly powerful and cheap computers and generally available high-intensity tuneable X-ray sources provided by dedicated synchrotrons. This allows crystal structures to be solved in weeks without too much basic understanding of space groups, reciprocal lattices or Patterson functions. Whereas formerly a great deal of energy had to be spent on crystallographic technology, now that energy can be put into thinking about the structural results and their implications for biological function.

The International Union of Crystallography (IUCr) has undergone parallel changes. Time for biological topics at the international congresses was determined largely by the program of the previous meeting three years earlier. This system was unresponsive to the changing focus of crystallographic advances. It was not until Dorothy Hodgkin, as president of the Union, took a lead in the organization of the 1975 IUCr meeting in Amsterdam that biology started to become a significant component of international meetings. However, the publications of the Union lagged behind for quite a few more years, precipitating the creation of journals such as *Structure* and *Nature Structural Biology* as outlets for the rapidly growing community of crystallographers with biological objectives. Yet papers that were devoted to technical problems had no other place than the IUCr publications, where they were lost to those who were primarily biologists. Fortunately, Jenny P. Glusker and others were able to persuade the IUCr that a specialized biological crystallographic journal was essential, leading to the creation of *Acta Crystallographica Section D*. Today, '*Acta D*' is the one journal where justification for the publication of biological crystallographic papers can be based primarily on their crystallographic significance. Jenny has steered this new journal through its first decade with great skill, overcoming and correcting problems inherited from the parent IUCr journals.

Although the rate of deposition and complexity of new structures that are deposited with the PDB continues to increase, crystallography may be reaching its limits. For instance, the uniqueness of individual biological cells makes it impossible to form arrays of repeating units. Or, attempts at using rapid X-ray data-collection techniques for time-resolved 'stills' of dynamic processes have difficulties in retaining a crystal lattice and in maintaining synchronization between different unit cells. Thus, in recent years, a growing number of investigators have been exploring the use of cryo-electron microscopy (cryoEM) to study large molecular assemblies and assembly intermediates using higher resolution crystallographic structures of components to fit into lower resolution cryoEM densities of large molecular complexes (Stewart *et al.*, 1993; Cheng *et al.*, 1995; Baker & Johnson, 1996; Grimes *et al.*, 1997; Kolatkar *et al.*, 1999; Volkmann & Hanein, 1999; Wriggers *et al.*, 1999; Kikkawa *et al.*, 2000; Roseman, 2000; Rossmann, 2000; San Martin *et al.*, 2001; Zhou *et al.*, 2001, 2003; He *et al.*, 2003). Such studies have varied greatly both in technique and application. Here, we present one such study in which we analyze the structures of two ∼1900 Å diameter dsDNA viruses using cryoEM reconstructions of the whole viruses (Yan *et al.*, 2000) and the X-ray crystal structure of the heavily glycosylated major capsid protein of one of these viruses (Nandhagopal *et al.*, 2002).

2. Large icosahedral dsDNA viruses

The *Phycodnaviridae*, *Iridoviridae* and *Asfaviridae* (African swine fever-like viruses) are related groups of 1500–2000 Å

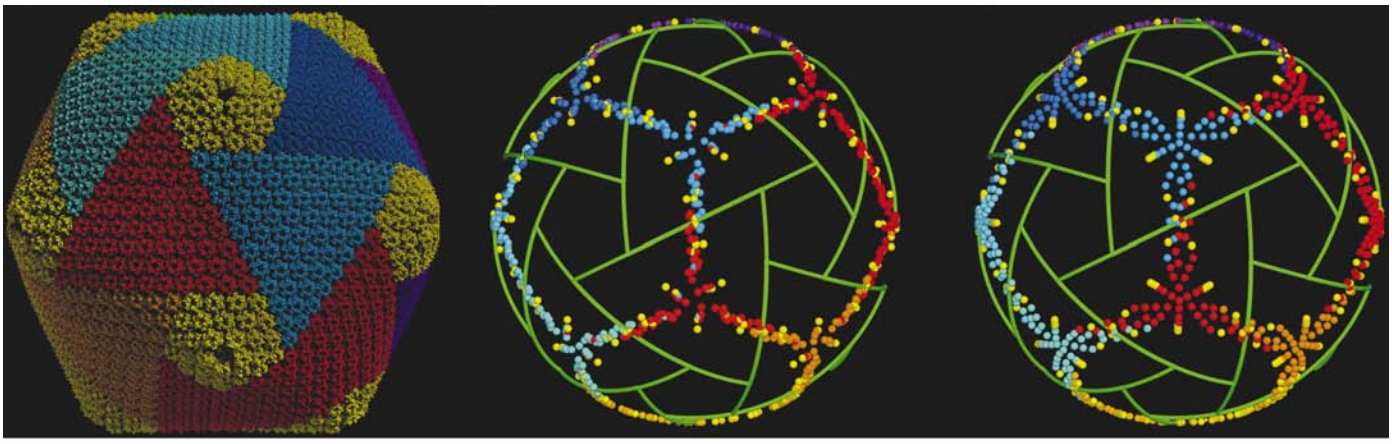

Figure 1
Virus structure. (*a*) Quasi-atomic model of PBCV-1 based on fitting the crystal structure of the Vp54 trimeric capsomer into the cryoEM density reconstruction. The pentasymmetrons are colored yellow, whereas the trisymmetrons are variously colored (adapted from Nandhagopal *et al.*, 2002). (*b* and *c*) The ends of vectors (spheres) representing the normals to the capsid surface, erected from the center of the icosahedron. The normals within the pentasymmetrons are colored yellow, whereas those in each trisymmetron are designated with colors corresponding to the color usage in (*a*). The outline of the icosahedral virus and the boundaries of the pentasymmetrons and trisymmetrons are shown in green. Note that the surface normals lie along great circles joining the threefold axes. (*b*) The distribution of normals for the independent fits of the capsomer atomic structure into the PBCV-1 cryoEM density; (*c*) the distribution of normals for CIV.

diameter icosahedral double-stranded DNA (dsDNA) viruses that infect algae, arthropods and vertebrates, respectively. The *Poxviridae* may also be related to this group, although they lack an icosahedral capsid (Iyer *et al.*, 2001). The bacterial virus PRD1 and adenoviruses are smaller (~750 Å diameter), but have similar capsid organizations (Yan *et al.*, 2000). *Paramecium bursaria Chlorella* virus type 1 (PBCV-1) is a member of the *Phycodnaviridae* and Chilo iridescent virus (CIV) is a representative of the *Iridoviridae*. Early studies using electron microscopy of negatively stained samples had shown that many of the *Iridoviridae* are icosahedral and can be broken down into triangular 'trisymmetrons', pentagonal 'pentasymmetrons' and possibly linear 'disymmetrons' (Wrigley, 1969; Stoltz, 1971), which are centered on the threefold, fivefold and twofold axes, respectively. A cryoEM study has shown that PBCV-1 and CIV have an outer glycoprotein shell composed almost entirely of 1680 (h = 7, k = 8) and 1460 (h = 7, k = 7) pseudo-hexagonal close-packed trimeric capsomers, respectively (Yan *et al.*, 2000). Their capsids are constructed from 20 trisymmetrons and 12 pentasymmetrons (Fig. 1*a*), but no disymmetrons.

The trisymmetrons are triangular two-dimensional assemblies of close-packed trimeric capsomers with approximate local $p3$ plane-group symmetry. In PBCV-1, each trisymmetron consists of 66 capsomers, with 11 capsomers along each edge centered on an icosahedral threefold axis. The icosahedral threefold axis is located between the three central capsomers. In contrast, the trisymmetrons of CIV consist of only 55 capsomers with ten along each edge, placing the icosahedral threefold axis in coincidence with the central capsomer. The trisymmetrons are distorted from planarity owing to the curvature produced by their wrap around the edges of the icosahedron (Fig. 1*a*). Adjacent trisymmetrons are related by icosahedral twofold axes, which create breaks in the $p3$ lattice that define the boundaries between trisymmetrons (Fig. 2). The orientations of the capsomers' threefold axes are neither radial (representing a uniform distribution on a spherical surface) nor are they parallel to the icosahedral threefold axes associated with each trisymmetron (representing a planar arrays of capsomers, which would form a flat-faced icosahedron). Instead, the vectors representing the directions of the capsomer threefold axes (which are also normals to the capsid surface), when erected from a common origin, tend to fall on great circles that contain the directions of the icosahedral threefold and twofold axes (Nandhagopal *et al.*, 2002) (Figs. 1*b* and 1*c*). The pentasymmetrons for both viruses consist of five triangular faces, each with six capsomers.

The structure of the PBCV-1 major capsid glycoprotein Vp54 (PDB code 1J5Q) has been determined by X-ray crystallography (Nandhagopal *et al.*, 2002). Each of the three monomers in the crystallized capsomer was found to consist of two sequential 'jelly-roll' domains, which give the trimeric capsomer pseudo-sixfold symmetry (Nandhagopal *et al.*, 2002). Similar domain organizations of viral capsid proteins

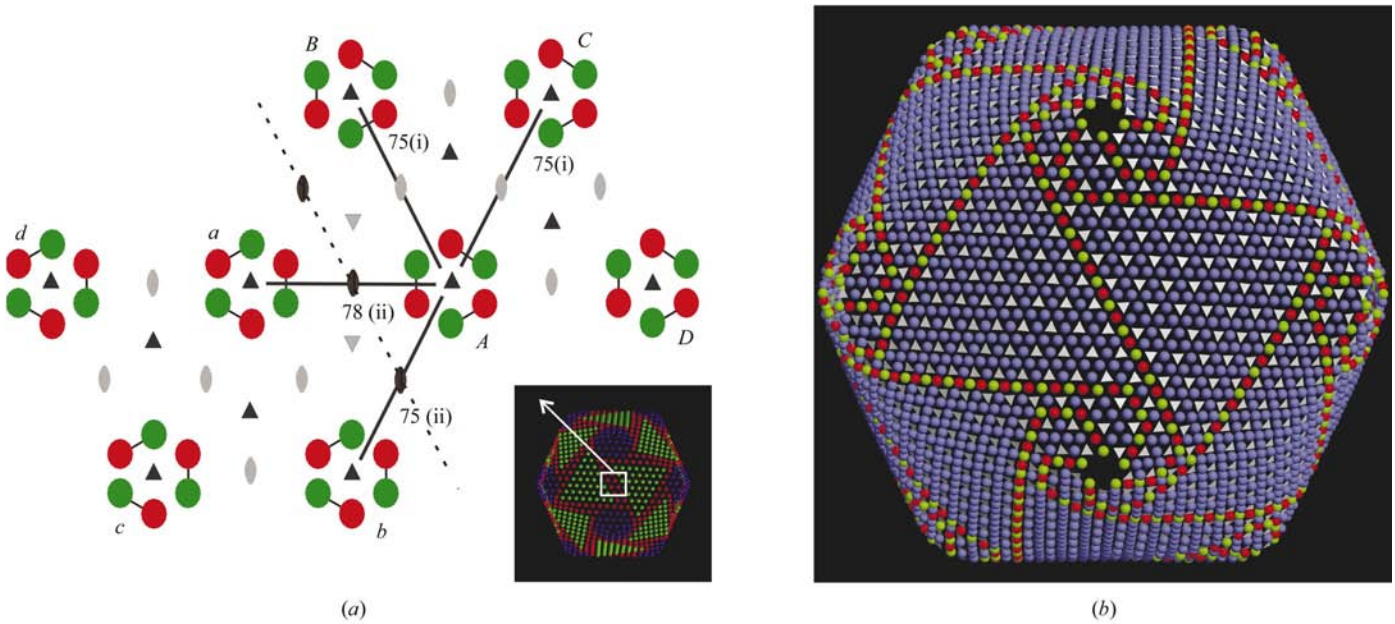

Figure 2
The $p3$ surface lattice. (*a*) Enlargement of the virus surface area (white square) seen in the inset. Each Vp54 monomer is represented diagrammatically by a green and a red circle corresponding to the two successive jelly-roll domains within the polypeptide. The domain organization within the trimeric capsomer generates a pseudo-hexameric structure. The capsomers are packed into a $p3$ lattice (symmetry elements in black) at the boundary (dashed line) between two adjacent trisymmetrons. Pseudosymmetry elements are in gray. Capsomers in the right-hand trisymmetron are labeled *A*, *B*, *C* and *D*, whereas the capsomers in the other trisymmetron are labeled *a*, *b*, *c* and *d*, with a central twofold axis relating *A* to *a etc.* The type of contacts between capsomers (see text) is shown as Roman numerals placed in parentheses along with their average distance. In (*b*) is shown the distribution of the three classes of inter-capsomer contacts in the capsid. The position of each capsomer is shown as a white triangle, the orientation and position of which reflects that of the fitted capsomer. A sphere is placed between each pair of capsomers. Blue contacts are between similarly oriented capsomers, whereas green and red spheres represent the long (~78 Å) and short (~75 Å) contacts between pseudo-sixfold related trimers. Note the alternating longer and shorter distances along the trisymmetron boundaries.

Table 1
Interactions between capsomers.

Values in parentheses are r.m.s. deviations from the mean.

	PBCV-1					CIV				
Contact type	No.†	Mean distance (Å)	Mean distance (restrained)‡ (Å)	Mean rotation§ (°)	Mean rotation§ (restrained) (°)	No.†	Mean distance (Å)	Mean distance (restrained)‡ (Å)	Mean rotation§ (°)	Mean rotation§ (restrained) (°)
Similarly oriented, type (i)										
All	64	75.1 (1.2)	75.6 (0.6)	7.9 (4.5)	5.3 (2.5)	54	77.2 (0.7)	77.2 (0.6)	6.5 (3.5)	5.7 (2.6)
In trisymmetron	55	75.0 (1.1)	75.6 (0.5)	7.0 (3.2)	4.7 (1.9)	45	77.2 (0.8)	77.3 (0.5)	5.7 (2.6)	5.1 (1.9)
In pentasymmetron	8	76.1 (1.1)	75.9 (0.8)	13.5 (7.0)	8.1 (3.1)	8	76.9 (0.5)	76.7 (0.8)	10.6 (4.1)	8.9 (3.4)
Oppositely oriented, type (ii)										
All	10	77.8 (1.5)	78.8 (1.4)	12.7 (8.6)	11.6 (5.4)	10	78.4 (1.1)	78.2 (3.3)	12.6 (9.1)	11.0 (8.0)
Between trisymmetrons	4	78.4 (1.2)	80.2 (0.2)	8.2 (2.9)	9.3 (1.9)	4	79.1 (0.3)	80.0 (0.3)	7.0 (1.1)	7.7 (0.5)
In pentasymmetron	3	76.7 (1.5)	77.1 (1.4)	21.7 (10.1)	17.5 (6.5)	3	76.9 (0.5)	74.6 (3.7)	22.9 (9.4)	17.9 (10.7)
Second type between oppositely oriented, type (iii)										
All	10	74.9 (1.2)	75.8 (0.6)	8.9 (4.9)	7.4 (3.8)	9	75.9 (1.3)	75.8 (1.2)	7.6 (5.8)	7.0 (5.8)
Between trisymmetrons	4	74.0 (0.5)	75.6 (0.5)	5.5 (2.7)	5.0 (1.7)	3	75.2 (0.2)	75.4 (0.2)	3.1 (0.8)	3.2 (1.0)
In pentasymmetron	3	75.9 (1.1)	76.1 (0.7)	14.6 (2.9)	13.0 (0.8)	3	77.2 (1.7)	76.6 (2.1)	15.0 (3.9)	14.2 (4.1)

† Number of interactions in an icosahedral asymmetric unit. ‡ Restrained refinement used a weight of $k = 0.4$ (see Table 2 and equation 1). § Rotations between 'oppositely oriented' pairs of capsomers have had the pseudo-sixfold rotation of one of the capsomers subtracted.

have been found for adenoviruses (Stewart *et al.*, 1993; San Martin *et al.*, 2001) and for the lipid-containing phage PRD1 (Bamford *et al.*, 1995). Hence, it is likely that these viral capsids evolved from a common ancestral protein with tandem jelly-roll domains that had already adopted pseudo-hexagonal symmetry.

A jelly-roll domain consists of an eight-stranded antiparallel β-barrel essentially composed of two opposing β-sheets. If the β-strands along the polypeptide strand are identified as *B*, *C*, *D*, ..., *I*, then the order of the strands in the two sheets is given by *BIDG* and *CHEF*. The surface of the capsomer can be divided into an outer highly glycosylated surface that forms the virus exterior, a positively charged internal surface that faces the acidic outer leaflet of the viral lipid membrane (Nandhagopal *et al.*, 2002) and an equatorial region that is mainly involved in making the contacts between the pseudo-hexameric capsomers (San Martin *et al.*, 2001). The surface of the equatorial region is primarily formed by the '*CHEF*' β-sheets of the six 'jelly-roll' structures in the trimeric capsomer, exposing alternatively the surface of the first and the second jelly-roll domains as represented by $[-(CHEF)^1-(CHEF)^2-]_3$ within each of the three monomers (where the superscripts refer to the domains within the monomer). Looking at the virus capsid from the outside, the sequence of β-sheets is anticlockwise within a capsomer.

3. Inter-capsomer interactions in PBCV-1 and CIV

The PBCV-1 capsomer crystal structure (Nandhagopal *et al.*, 2002) and the cryoEM image reconstruction of the entire capsid (Yan *et al.*, 2000) were combined to construct a quasi-atomic model of the viral capsid (Fig. 1*a*). The 28 capsomers in the icosahedral asymmetric unit of PBCV-1 (22 in a trisymmetron and six in a pentasymmetron) were fitted individually into a 28 Å resolution cryoEM density map using the programs *SITUS* (Wriggers *et al.*, 1999) and *EMfit* (Rossmann *et al.*, 2001). Although the capsomers are pseudo-hexagonal, the two domains were sufficiently different that the 60° rotations were clearly distinguishable. The results of the two different procedures were very similar and agreed well with expectations by placing the carbohydrate sites consistently on the exterior surface and the amino ends always in the viral interior (Nandhagopal *et al.*, 2002). All capsomers within a trisymmetron were oriented similarly with respect to rotation about the pseudo-sixfold axis. However, of the six independent capsomers in the face of a pentasymmetron, five were oriented one way and one was turned by ~60°. There were large differences in the amount of buried surface area between different capsomers (varying from 1240 to 100 Å^2), with the buried surface area being highly correlated with the inter-capsomer distance (73–80 Å, respectively).

The cryoEM reconstruction of CIV showed a very similar distribution of capsomers organized into trisymmetrons and pentasymmetrons as in PBCV-1 (Yan *et al.*, 2000). Since the capsomer of PBCV-1 shares ~20% sequence identity with CIV, with no large insertions or deletions, the PBCV-1 Vp54 capsid protein structure was used as an approximate model for that of CIV. However, the resulting fit of the various capsomers into the CIV cryoEM density did not consistently place the external side of Vp54 onto the exterior of the virus. Therefore, the averaged cryoEM density for the CIV capsomers was used as a model in place of the atomic structure of PBCV Vp54. For this purpose, the Vp54 model was used as a mask to define the density region corresponding to one capsomer and to define the position of the central threefold axis.

The contacts between capsomers could be separated into three types (Table 1; Fig. 2*a*): one between similarly oriented capsomers (type i) and two between those with different orientations (types ii and iii) depending on their position in the surface lattice (Fig. 2*a*). These classes of capsomer contacts were subdivided depending on whether they were within

Table 2
Effect of changing the relative weight, k, between density correlation and distance restraints (see equation 1).

	PBCV		CIV	
Scale factor (k)	Mean density correlation† C	R.m.s. distance error (Å)	Mean density correlation‡ C	R.m.s. distance error (Å)
0.0	0.730	2.7	0.960	2.4
0.1	0.725	1.8	0.953	1.7
0.2	0.719	1.5	0.946	1.5
0.4	0.708	1.3	0.913	1.3
0.8	0.693	1.0	0.913	1.0

† The correlation coefficient based on agreement with atomic model of PBCV Vp54. ‡ The correction coefficient based on agreement with averaged cryoEM density of CIV capsomers.

trisymmetrons, within pentasymmetrons or across their boundaries. All the capsomers within a trisymmetron are of the first type, whereas those between trisymmetrons are a mixture of the other two classes. Since each face of the pentasymmetron has five similarly oriented and one oppositely oriented capsomer in both PBCV and CIV, the pentasymmetron contains all three types of interaction. Owing to the curvature of the capsid surface, the inter-capsomer distance is dependent on the definition of the capsomer's origin along the threefold axis. The position of the capsomer origin was chosen so as to minimize the standard deviations for the inter-capsomer distances.

The distances (Table 1) are remarkably similar to the earlier determination by Wrigley (1969, 1970), who established the average inter-capsomer distance for *Sericesthis* iridescent virus to be 70.1 ± 2.3 Å, and to the distances measured from cryoEM reconstructions for PRD1 (71.8 Å). However, the type (iii) contacts are significantly shorter than the others for both PBCV-1 and CIV (Table 1). Considering that the area of contact between capsomers was shown to be roughly proportional to the distance between capsomers, the most stable entities are the trisymmetrons, whereas the interactions along their boundaries are weak in comparison. The possibility of weaker interactions between the trisymmetrons may also be supported by the larger rotation angles between adjacent capsomers (∼11° instead of ∼5°) found in the longer type of interaction. This is consistent with observations on various iridovirus capsids, which readily break apart into trisymmetrons and pentasymmetrons (Wrigley, 1969, 1970; Stoltz, 1971).

4. Minimizing error by assuming the observed systematic properties

It should be possible to improve the accuracy of the fitting procedure by imposing reasonable assumptions about the spatial and orientational relationships between capsomers. Although in general the translational and rotational transformations relating adjacent capsomers change smoothly across the capsid, they are discontinuous (see above) at the trisymmetron and pentasymmetron boundaries. In order to restrain the independent fitting of each capsomer, a set of soft distance restraints was applied which allowed for the different types of interactions. Variation from the average arises from differences in the local curvature of the capsid as well as random error. Thus, the error could be substantially reduced by fitting the capsomer structure into the cryoEM density while restraining the three types of inter-capsomer relationships to be equal to their averaged values. This was accomplished by independently refining the position (three coordinates) and orientation (three angles) for each capsomer by maximizing E, where

$$E = C - k\left(\frac{C_0}{R_0}\right)R, \tag{1}$$

and

$$R = \sum(d_{\rm exp} - d_{\rm obs})^2, \tag{2}$$

and C is the correlation between the cryoEM density and the model density of the trimeric capsomer at each C^{α} atom. The restraining term R was defined as a sum of the squared differences in distance between equivalent atoms in adjacent capsomers, $d_{\rm obs}$, and their expected values, $d_{\rm exp}$. The restraints had the effect of maintaining approximate inter-capsomer distances and keeping neighboring capsomers parallel. The scale factor k was used to adjust the contribution of the restraints relative to C. The normalizing coefficients C_0 and R_0 refer to the mean values of C and R when $k = 0$. After cycling over each of the capsomers independently, the procedure was repeated until convergence was attained. The resulting quasi-atomic structure was similar to the independent non-restrained fitting, except that the inter-capsomer distances had tighter distributions about the $d_{\rm exp}$ values and the magnitude of the rotations between neighboring capsomers was smaller (Table 1). As the scale factor k is increased, giving more weight to the distance restraints, there is a large reduction in the distance residuals accompanied by only a slight worsening of the density correlations (Table 2). Thus, taking the capsid as a whole, most of the distance deviations arise from non-systematic fluctuations rather than the smooth curvature of the capsid, which the restraints are unable to eliminate. Assuming no significant conformational differences between the PBCV-1 Vp54 major capsid protein in the crystal and in the virion, the repetitive character of the surface lattice has made it possible to determine the atomic positions in the viral capsids to an accuracy of maybe 2 Å (Table 1), although the cryoEM image reconstructions were limited to only 28 Å resolution.

5. Interactions between capsomers

Crick & Watson (1956, 1957) had predicted that because each identical protein subunit in a viral capsid would require an identical environment, the capsid would have the symmetry of a regular polyhedron. The largest number of asymmetric units in a regular polyhedron is 60 for an icosahedron. However, most viruses have more than 60 subunits. Caspar & Klug (1962) showed by studying the wrapping of two-dimensional

hexagonal grids onto the surface of an icosahedron that there could be quasi-equivalent environments in icosahedra with $60T$ subunits. Each grid point in the lattice would be associated with a hexamer of subunits, other than those at grid points defining the pentameric vertices. If h and k are integers defining the position of the fivefold vertices on a hexagonal lattice, then the number of monomeric units in the icosahedral asymmetric unit is given by $T = h^2 + hk + k^2$, distributed between $(T - 1)/6$ hexameric and one pentameric capsomer. Although many viruses have been shown to obey the rules of quasi-equivalence (Harrison *et al.*, 1978; Rossmann & Johnson, 1989; Liljas, 2002), the exactness of the equivalence is substantially compromised in real three-dimensional structures when examined at atomic resolution. The structures of $T = 3$ plant and insect viruses show regions of local structural disorder that enhance the equivalence of the ordered domains of the capsid proteins (Harrison *et al.*, 1978; Rossmann & Johnson, 1989). Greater deviations from the predicted triangulation lattices were found in the studies of polyoma viruses (Rayment *et al.*, 1982; Baker *et al.*, 1983) and SV40 (Baker *et al.*, 1988; Liddington *et al.*, 1991), in which the anticipated mixtures of hexameric and pentameric structures were found to be all pentamers, thus heightening the equivalence of interactions within the pentameric capsomers at the expense of reducing the equivalence of inter-pentamer contacts.

The structures of PBCV-1 ($h = 7$, $k = 8$, pseudo-$T = 169$; Yan *et al.*, 2000); CIV ($h = 7$, $k = 7$, pseudo-$T = 147$; Yan *et al.*, 2000); African swine fever virus (Carrascosa *et al.*, 1984; Andres *et al.*, 1997; Rouiller *et al.*, 1998); adenovirus ($h = 5$, $k = 0$, pseudo-$T = 25$; Stewart *et al.*, 1993; Bamford *et al.*, 1995) and PRD1 ($h = 5$, $k = 0$, pseudo-$T = 25$; Butcher *et al.*, 1995; San Martin *et al.*, 2001) not only have evolved unique proteins for their fivefold vertices, they have also undergone a second gene duplication and symmetry breaking, replacing the hexamers with pseudo-hexagonal trimers. Thus, the T number accounts for the number of jelly rolls in the icosahedral asymmetric unit and the number of monomers in the capsid will be $(T - 1) \times 30$, after subtracting the 60 jelly rolls that would be associated with the pentameric vertices.

The boundary that separates a trisymmetron from the adjacent structures is defined by the hexagonal grid coordinate $h/2$. Thus, if h is odd this line lies between capsomers, allowing adjacent trisymmetrons to make contact across the twofold axis; however, if h is even this line crosses the capsomer centers, necessitating the creation of a linear disymmetron, although it is not certain whether any viruses make use of these units. If both h and k are even, the central trimer of the disymmetron would lie exactly on the icosahedral twofold axis.

Table 3
Number, n, of quasi-sixfold capsomers along the edge of a trisymmetron.

Number of capsomers along the edge of a trisymmetron: $n = (h + 2k - 1)/2$. h and k are the integral coordinates on a hexagonal network denoting the position of pentameric vertices relative to such a vertex placed at the origin (Caspar & Klug, 1962). Note that h and k cannot both be even, otherwise a trimeric capsomer would have to coincide with an icosahedral twofold axis. Combinations of h and k that place the threefold center of a trisymmetron on a capsomer are in bold. Otherwise, the center is positioned between three adjacent capsomers. General notes: $T = h^2 + hk + k^2$ (h must be even). Number of pseudo-hexagonal capsomers in a virion: $10\,(T - 1)$. Number of capsomers in a trisymmetron: $[(h + 2k)^2 - 1]/8$. Number of capsomers along the edge of a pentasymmetron: $p = (h + 1)/2$. Number of capsomers in a pentasymmetron (excluding the penton): $[5(h^2 - 1)]/8$.

h	$p = (h + 1)/2$	$k = 0$	1	2	3	4	5	6	7	8
1	1	0	**1**†	2	3	**4**	5	6	**7**	8
3	2	**1**	2	3	**4**	5	6	**7**	8	9
5	3	2‡	3	**4**	5	6	**7**	8	9	**10**
7	4	3	**4**	5	6	**7**	8	9§	**10**¶	11††

† Picornaviruses. ‡ Adenoviruses and lipid-containing phage PRD1. § *Sericesthis* iridescent virus (tentative assignment). ¶ CIV. †† PBCV-1.

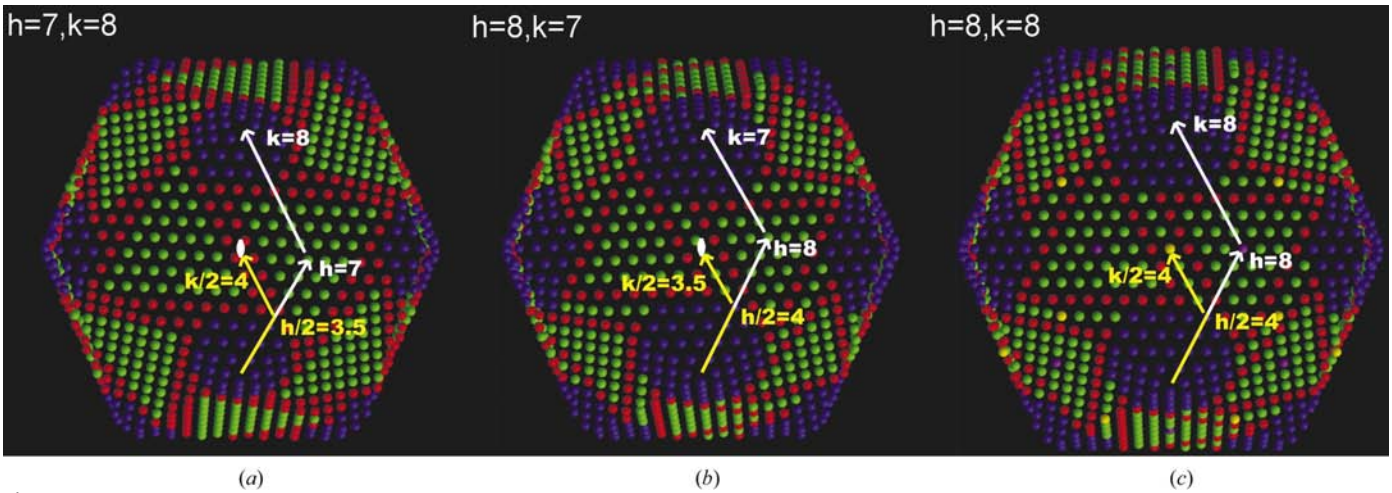

Figure 3
Possible arrangements of trisymmetrons and pentasymmetrons for different values of h and k. The pentasymmetrons are shown in purple and the trisymmetrons are bordered in red. The hexagonal surface lattice vectors which link adjacent fivefold vertices are shown in white and those linking fivefold and twofold axes are in yellow. The left figure shows the situation found in PBCV-1, where the capsid is formed solely from pentasymmetrons and trisymmetrons. In the center structure is an alternative arrangement for the same T number as PBCV-1, but which contains disymmetrons. Since k is even in this case, none of the capsomers are centered on the twofold axes. The right diagram shows the situation with both h and k even, where some capsomers lie on a twofold axis.

In this case, either the virus must abandon icosahedral symmetry or a special twofold (or sixfold) symmetric capsomer would be needed (Fig. 3). The modified Goldberg diagram (Goldberg, 1937) (Table 3) shows where viruses that have pseudo-hexameric as opposed to perfectly hexameric capsomers fit into the overall scheme.

If the capsomers were to have exact sixfold symmetry (as is assumed by Caspar & Klug, 1962), all contacts between neighboring capsomers would be similar, eliminating the natural cleavage lines which are created by the weaker interactions between trisymmetrons and pentasymmetrons. As it seems likely that these units are intermediate building blocks for viruses, such as PBCV and CIV, the use of such regular arrays as building blocks may be to allow the assembly of perfect icosahedra by permitting slight adjustments between close packed regular arrays. The use of trisymmetrons and pentasymmetrons in the assembly of large viruses may also control the size of the icosahedral faces. Thus, the pseudo-hexagonal motif, based on the consecutive repeat of the common jelly-roll motif, appears to provide a solution that is energetically favorable and therefore may explain the frequent occurrence of the jelly-roll motif in the structure of virus capsids.

We thank Xiaodong Yan and Timothy S. Baker for helpful discussions, as well as Cheryl Towell and Sharon Wilder for help in the preparation of the manuscript. The work was supported by NIH grants to MGR and JLVE.

References

Andres, G., Simon-Mateo, C. & Vinuela, E. (1997). *J. Virol.* **71**, 2331–2341.

Baker, T. S., Caspar, D. L. D. & Murakami, W. T. (1983). *Nature (London)*, **303**, 446–448.

Baker, T. S., Drak, J. & Bina, M. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 422–426.

Baker, T. S. & Johnson, J. E. (1996). *Curr. Opin. Struct. Biol.* **6**, 585–594.

Bamford, D. H., Caldentey, J. & Bamford, J. K. H. (1995). *Adv. Virus Res.* **45**, 281–319.

Butcher, S. J., Bamford, D. H. & Fuller, S. D. (1995). *EMBO J.* **14**, 6078–6086.

Carrascosa, J. L., Carazo, J. M., Carrascosa, A. L., Garcia, N., Santisteban, A. & Vinuela, E. (1984). *Virology*, **132**, 160–172.

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.

Crick, F. H. C. & Watson, J. D. (1956). *Nature (London)*, **177**, 473–475.

Crick, F. H. C. & Watson, J. D. (1957). *Ciba Foundation Symposium on The Nature of Viruses*, edited by G. E. W. Wolstenholme & E. C. P. Millar, pp. 5–13. Boston: Little, Brown & Co.

Goldberg, M. (1937). *Tohoku Math. J.* **43**, 104–108.

Grimes, J. M., Jakana, J., Ghosh, M., Basak, A. K., Roy, P., Chiu, W., Stuart, D. I. & Prasad, B. V. V. (1997). *Structure*, **5**, 885–893.

Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. (1978). *Nature (London)*, **276**, 368–373.

He, Y., Mueller, S., Chipman, P. R., Bator, C. M., Peng, X., Bowman, V. D., Mukhopadhyay, S., Wimmer, E., Kuhn, R. J. & Rossmann, M. G. (2003). *J. Virol.* **77**, 4827–4835.

Iyer, L. M., Aravind, L. & Koonin, E. V. (2001). *J. Virol.* **75**, 11720–11734.

Kikkawa, M., Okada, Y. & Hirokawa, N. (2000). *Cell*, **100**, 241–252.

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

Liddington, R. C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T. L. & Harrison, S. C. (1991). *Nature (London)*, **354**, 278–284.

Liljas, L. (2002). *Curr. Opin. Struct. Biol.* **6**, 151–156.

Nandhagopal, N., Simpson, A. A., Gurnon, J. R., Yan, X., Baker, T. S., Graves, M. V., Van Etten, J. L. & Rossmann, M. G. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 14758–14763.

Rayment, I., Baker, T. S., Caspar, D. L. D. & Murakami, W. T. (1982). *Nature (London)*, **295**, 110–115.

Roseman, A. M. (2000). *Acta Cryst.* D**56**, 1332–1340.

Rossmann, M. G. (2000). *Acta Cryst* D**56**, 1341–1349.

Rossmann, M. G., Bernal, R. & Pletnev, S. V. (2001). *J. Struct. Biol.* **136**, 190–200.

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

Rouiller, I., Brookes, S. M., Hyatt, A. D., Windsor, M. & Wileman, T. (1998). *J. Virol.* **72**, 2373–2387.

San Martin, C., Burnett, R. M., de Haas, F., Heinkel, R., Rutten, T., Fuller, S. D., Butcher, S. J. & Bamford, D. H. (2001). *Structure*, **9**, 917–930.

Stewart, P. L., Fuller, S. D. & Burnett, R. M. (1993). *EMBO J.* **12**, 2589–2599.

Stoltz, D. B. (1971). *J. Ultrastruct. Res.* **37**, 219–239.

Volkmann, N. & Hanein, D. (1999). *J. Struct. Biol.* **125**, 176–184.

Wriggers, W., Milligan, R. A. & McCammon, J. A. (1999). *J. Struct. Biol.* **125**, 185–189.

Wrigley, N. G. (1969). *J. Gen. Virol.* **5**, 123–134.

Wrigley, N. G. (1970). *J. Gen. Virol.* **6**, 169–173.

Yan, X., Olson, N. H., Van Etten, J. L., Bergoin, M., Rossmann, M. G. & Baker, T. S. (2000). *Nature Struct. Biol.* **7**, 101–103.

Zhou, Z. H., Baker, M. L., Jiang, W., Dougherty, M., Jakana, J., Dong, G., Lu, G. & Chiu, W. (2001). *Nature Struct. Biol.* **8**, 868–873.

Zhou, Z. H., Zhang, H., Jakana, J., Lu, X.-Y. & Zhang, J.-Q. (2003). *Structure*, **11**, 651–663.