The X-ray crystal structure of P3, the major coat protein of the lipid-containing bacteriophage PRD1, at 1.65 Å resolution

**Stacy D. Benson,[a] Jaana K. H.
Bamford,[b] Dennis H. Bamford[b]
and Roger M. Burnett[a]***

[a]The Wistar Institute, 3601 Spruce Street,
Philadelphia, PA 19104, USA, and [b]Department
of Biosciences and Institute of Biotechnology,
PO Box 56 (Viikinkaari 5), University of
Helsinki, FIN-00014, Helsinki, Finland

Correspondence e-mail:
burnett@wistar.upenn.edu

P3 has been imaged with X-ray crystallography to reveal a trimeric molecule with strikingly similar characteristics to hexon, the major coat protein of adenovirus. The structure of native P3 has now been extended to 1.65 Å resolution (R_{work} = 19.0% and R_{free} = 20.8%). The new high-resolution model shows that P3 forms crystals through hydrophobic patches solvated by 2-methyl-2,4-pentanediol molecules. It reveals details of how the molecule's high stability may be achieved through ordered solvent in addition to intra- and intersubunit interactions. Of particular importance is a 'puddle' at the top of the molecule containing a four-layer deep hydration shell that cross-links a complex structural feature formed by 'trimerization loops'. These loops also link subunits by extending over a neighbor to reach the third subunit in the trimer. As each subunit has two eight-stranded viral jelly rolls, the trimer has a pseudo-hexagonal shape to allow close packing in its 240 hexavalent capsid positions. Flexible regions in P3 facilitate these interactions within the capsid and with the underlying membrane. A selenomethionine P3 derivative, with which the structure was solved, has been refined to 2.2 Å resolution (R_{work} = 20.1% and R_{free} = 22.8%). The derivatized molecule is essentially unchanged, although synchrotron radiation has the curious effect of causing it to rotate about its threefold axis. P3 is a second example of a trimeric 'double-barrel' protein that forms a stable building block with optimal shape for constructing a large icosahedral viral capsid. A major difference is that hexon has long variable loops that distinguish different adenovirus species. The short loops in P3 and the severe constraints of its various interactions explain why the PRD1 family has highly conserved coat proteins.

1. Introduction

PRD1 is the prototypic example of the virus family Tectiviridae, which are bacteriophages characterized by their possession of a membrane beneath an icosahedral protein shell (Fig. 1). As internal membranes are unusual in bacteriophages, PRD1 has been a good model for studying membrane organization and biogenesis in a well defined bacterial system (Bamford et al., 1995). Five additional PRD1-related viruses have been isolated: PR3, PR4, PR5, PR772 and L17 (Bamford et al., 1981). Although these were found from locations around the world, the members of this family are very closely related and so the experimental findings for any one are applicable to all.

PRD1 was discovered when bacteriophages were sought that could infect bacteria with an incompatibility group P (IncP) plasmid (Olsen et al., 1974). The IncP plasmid often

contains genes conferring resistance to various antibiotics including ampicillin, kanamycin and tetracycline. It also encodes for a promiscuous conjugative DNA-transfer complex recognized by PRD1 (Grahn *et al.*, 1999; Kotilainen *et al.*, 1993). Thus, although PRD1 can infect a broad range of Gram-negative bacteria such as *Escherichia coli* and *Salmonella*, these must contain IncP, or the related IncN or IncW plasmids, to produce the DNA-transfer complex. There is some practical interest in bacteriophages that infect a wide range of Gram-negative bacteria because they may be useful as antibacterial agents (Barrow & Soothill, 1997; Merril *et al.*, 1996) to combat emerging antibiotic resistant strains.

The architecture of the PRD1 virion has been studied with cryo-electron microscopy. Image reconstructions to 28 Å resolution have been performed for both the complete virion and the empty capsid of *sus1*, a genome-packaging mutant (Butcher *et al.*, 1995). The *sus1* particles are complete virions apart from their missing genome (Mindich, Bamford, Goldthwaite *et al.*, 1982) and they are most likely to correspond to the procapsid stage of the bacteriophage. PRD1 is icosahedral but unlike other dsDNA bacteriophages (King & Chiu, 1997) it lacks a tail. In other bacteriophages, the tail is used as a conduit during infection to deliver the genome into the host bacterium. PRD1 compensates for the absence of this architectural feature by forming a tube with its internal membrane to produce a 'tail' with the same function (Bamford *et al.*, 1995).

Whereas other dsDNA bacteriophages display considerable reorganization upon maturation and genome packaging (King & Chiu, 1997), the PRD1 capsid does not expand or undergo a structural rearrangement of its proteins. The arrangements of

Figure 1
Schematic organization of bacteriophage PRD1. PRD1 has an icosahedral protein coat with facets made with 240 copies of the P3 trimer. The vertices contain the pentameric protein P31 in association with a spike composed of P5 and the host-recognition protein P2. The protein coat completely surrounds a lipid membrane, which closely follows the protein capsid and encloses a single copy of linear dsDNA with terminal proteins covalently attached to its 5′ ends.

the capsid proteins are identical in virions and empty particles, which both have the same diameter: 637 Å between opposite faces and 698 Å between opposite vertices (San Martín *et al.*, 2001). However, the organization of the PRD1 internal membrane does change upon maturation (Butcher *et al.*, 1995). The membrane is acquired from the host bacterium as assembly proceeds. Initially, a section of the bacterial plasma membrane accumulates up to 15 different PRD1-encoded membrane-associated proteins; bacterial proteins are excluded. This section then buds into the cytosol where it acts as a scaffold for the assembly of the capsid proteins into the lipid-containing procapsid (Bamford *et al.*, 1995; Butcher *et al.*, 1995).

The structural organization of the bacteriophage PRD1 capsid is surprisingly similar to that of the mammalian adenovirus, except for its inner membrane (Butcher *et al.*, 1995). The capsid is icosahedral, with 240 copies of the trimeric major coat protein P3 forming the 20 facets and the pentameric protein P31 occupying the vertices (Butcher *et al.*, 1995; Rydman *et al.*, 1999). P3 is encoded by gene *III* of the 14 925 bp genome of PRD1. The 394-residue P3 forms a very stable trimer (3 × 43 317 Da) with the assistance of the bacterial GroEL/ES chaperonins (Hänninen *et al.*, 1997). The P3 trimers, which account for ~80% of the protein content of the PRD1 virion (Davis *et al.*, 1982), are arranged on a pseudo $T = 25$ lattice in the capsid (Butcher *et al.*, 1995). This arrangement has only been observed in one other virus, adenovirus (Stewart *et al.*, 1991), with P3 equivalent to hexon and P31 to penton base. The vertex protein in each virus has an associated host-recognition protein, called spike or P2/P5 in PRD1 (Grahn *et al.*, 1999; Rydman *et al.*, 1999) and fiber in adenovirus (Stewart *et al.*, 1991). Other striking similarities occur in the genome and its replication. Both viruses have linear dsDNA genomes with inverted terminal repeats (Aleström *et al.*, 1982; Savilahti & Bamford, 1986) and 5′ covalently linked terminal proteins (Bamford & Mindich, 1984; Desiderio & Kelly, 1981). The terminal protein is used in priming the DNA replication through a similar 'sliding-back' mechanism to produce the 3′ nucleotide (Caldentey *et al.*, 1992; Kelly, 1984).

The finding that the overall topology of P3 is analogous to that of adenovirus hexon advances the similarities between the two viruses and points to their evolutionary relationship (Benson *et al.*, 1999). This similarity and its evolutionary significance has been further examined in two recent commentaries (Belnap & Steven, 2000; Hendrix, 1999). Both P3 and hexon are very stable trimeric proteins, with each subunit composed of a pair of jelly rolls (Athappilly *et al.*, 1994; Benson *et al.*, 1999). Both trimers have pseudo-hexagonal arrangements of their jelly rolls and a similar disposition of their loops, although the hexon polypeptide (967 residues for type 2 adenovirus) is much longer than that of P3 (394 residues).

The structure of P3 has now been refined to 1.65 Å resolution and this enhancement of detail has allowed a closer analysis of the components responsible for the stability of the trimer. Solvent atoms have been examined and their structural
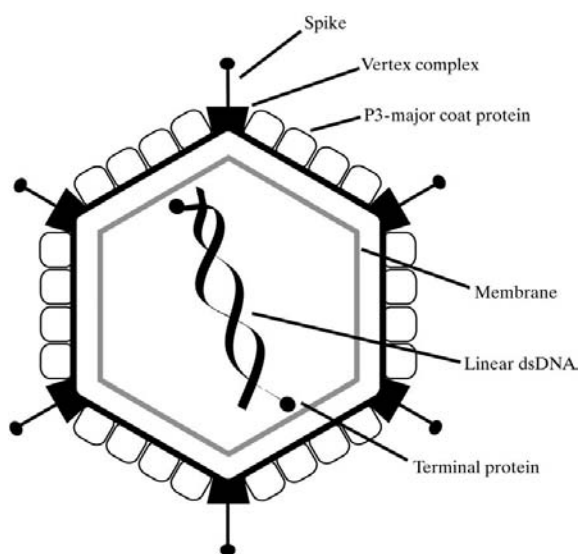
role in the trimer evaluated and an analysis of the locations of 2-methyl-2,4-pentanediol (MPD) molecules has pointed to their importance in crystallization. The refined seleno-methionine (Se-Met) P3 structure allows a direct comparison with native P3 to determine the effects of substituting selenium for sulfur. A recent structure for adenovirus type 5 hexon (Rux & Burnett, 2000) has provided some modifications to the original type 2 model. A comparison of the latest P3 and hexon structures reveals that they have even more elements in common and so further cements the remarkable relationship that is emerging between a bacteriophage and an animal virus. Certain features of the P3 structure, such as the N-terminal helix, can only be appreciated with reference to their environment in the PRD1 virion. This information is being provided by complementary electron-microscopy (EM) studies (San Martín *et al.*, 2001).

2. Materials and methods

2.1. Protein production and purification

As it has not been possible to produce recombinant P3, the protein was obtained from purified virus. Wild-type bacterio-phage PRD1 virions were grown in *S. enterica* serovar Typhimurium DS88 and purified in a sucrose gradient as described previously (Bamford & Bamford, 1991). Two different methods were used to disrupt the bacteriophage virions for P3 purification. Originally, virions were denatured with 2.0 M guanidine hydrochloride (Bamford & Mindich, 1982) and subsequently with high pH and heat (348 K for 15 min in 50 mM Tris–HCl pH 8.5; Caldentey *et al.*, 1993). Both methods disassembled the virions without breaking the P3 trimers into monomers. After disruption, the aggregated membrane and DNA material were removed by differential centrifugation (Sorwall T865 rotor, 32 000 rev min⁻¹, 3 h at 283 K) and the supernatant, containing the P3 trimers, was dialyzed against 10 mM Tris–HCl pH 8.5 for 17 h at 277 K. P3 was purified from the dialysate by ion-exchange chromatography (Stewart, Ghosh *et al.*, 1993). The purified P3 protein was concentrated to 5 mg ml⁻¹ in 10 mM Tris–HCl pH 8.5 and ∼300 mM sodium chloride for crystallization trials. Although protein from both purification methods gave crystals that were identical and could be used equally well for data collection, the heat-treated protein stock appeared cleaner when silver stained on SDS–PAGE gels. It also crystallized more readily, with fewer showers of small crystals.

Two methods were applied to obtain an Se-Met derivative of P3. In the first, the DB88 cells were subjected to nitroso-guanidine mutagenesis followed by penicillin selection and screening of methionine auxotrophs. Eight independent methionine mutants were picked and tested for their capacity to support PRD1 growth in M9 methionine medium. The best mutant had approximately the same capacity to produce viruses in the liquid culture as the wild-type bacterium. The virus was produced using this mutant in M9 medium containing 120 µg ml⁻¹ D,L-Se-Met (Sigma) and a tenfold multiplicity of infection. Because of low viral yield with

Se-Met, a second method was adopted that gave better results. Wild-type DB88 cells were grown to 1 × 10⁹ cells ml⁻¹ in M9 medium and infected with the same multiplicity. 20 min post-infection, 120 µg ml⁻¹ D,L-Se-Met was added and the growth was continued until lysis occurred. Se-Met P3 protein was purified from heat-disrupted virions and crystallized as for the native protein. MALDI-TOF mass-spectrometry analysis, using native P3 as a standard, showed >85% substitution for the five methionines of the P3 subunit.

2.2. Crystallization and data collection

The crystals were grown using the hanging-drop method in 24-well Linbro trays. The wells contained 1.0 ml of the crystallization solution and the drops were formed with 2.0 µl from the well and 2.0 µl of the protein solution. A previous crystallization condition (Stewart, Ghosh *et al.*, 1993) using aged polyethylene glycol (PEG) (4–6% PEG 4000, 25 mM sodium phosphate pH 6.0) produced only stacked thin plate-like crystals with fresh PEG solutions. These were unsuitable for diffraction studies, appeared infrequently and grew slowly. As attempts at artificially aging PEG were unsuccessful, new conditions were sought. An initial lead with Hampton Crystal Screen II (Cudney *et al.*, 1994) condition 10 (30% MPD, 0.1 M sodium acetate pH 4.6, 0.2 M sodium chloride) was refined to 35% MPD, 0.1 M sodium acetate pH 4.1–4.3 and 0.2 M NaCl. Large crystals appeared within a couple of days and grew to their final size, typically 0.50 × 0.45 × 0.40 mm, in a week. They were stable in the drop for up to six months with no apparent deterioration in their diffraction properties.

The initial diffraction data were collected with a Siemens X100A area detector and a Rigaku RU-200B rotating copper-anode X-ray generator, fitted with a double-mirror mono-chromator (Table 1). The crystal-to-detector distance was 120–150 mm, with a 2θ angle offset of between 0 and 25°. The MPD crystals diffracted well, with data extending to beyond 2.5 Å resolution, but were very X-ray sensitive unless frozen. As MPD was an excellent cryoprotectant, crystals could be mounted in a cryoloop directly from the hanging-drop solution, frozen under the nitrogen stream and data collected while the crystal was maintained at 100 K. The diffraction quality of the crystals no longer deteriorated during data collection and complete 90° scans could be collected from a single crystal.

The data were reduced with the program *XDS* (Kabsch, 1988). The unfrozen MPD crystals have the same space group (P2₁2₁2₁) as the crystals grown in PEG (Stewart, Ghosh *et al.*, 1993), but two of the unit-cell parameters are slightly smaller (a = 119.4, b = 124.9, c = 126.8 Å *cf.* a = 121.6, b = 123.2, c = 128.6 Å). The MPD crystal's unit cell shrinks even further at 100 K (a = 118.0, b = 121.3, c = 126.4 Å). There is one trimer in the asymmetric unit (3 × 43 317 Da) and the Matthews coefficient (Matthews, 1968), V_M, of 3.48 Å³ Da⁻¹ suggests a solvent content of 64.4% if a protein density of 1.34 g ml⁻¹ is assumed.

Self-rotation function calculations were performed on the native data over various resolution ranges using the program

research papers

Table 1
Data-collection statistics.

Numbers in parentheses correspond to the highest resolution shell.

Crystal†	λ (Å)	Source	Unique reflections	Multi-plicity‡	Complete-ness (%)	Resolution range (Å)	R_{sym}§ (%)
Native	0.9500	NSLS X12C	195631	4.6	94.2 (61.2)	87–1.65	4.7 (40.3)
Native$_{1cjd}$	0.9500	NSLS X12C	151224	5.2	99.8 (91.6)	35–1.85	4.6 (22.8)
Native$_{lab}$	1.5418	Laboratory	86433	6.4	97.3 (86.1)	2.5–2.2	19.4 (52.8)
Se-Met L1	0.9790	NSLS X12C	87526	5.0	98.5 (88.0)	35–2.2	5.7 (15.0)
Se-Met L2	0.9786	NSLS X12C	87548	5.1	98.5 (79.2)	35–2.2	6.4 (18.2)
Se-Met L3	0.9500	NSLS X12C	87564	4.9	98.5 (95.0)	35–2.2	6.5 (18.5)
Se-Met lab	1.5418	Laboratory	78177	3.9	91.6 (67.6)	28–2.2	12.7 (27.4)
Uranium1	1.5418	Laboratory	33672	2.3	86.9 (53.2)	25–2.9	5.9 (13.1)
Uranium2	1.5418	Laboratory	50581	1.6	83.6 (67.8)	25–2.8	5.7 (15.1)
Platinum	1.5418	Laboratory	55324	4.2	97.5 (92.8)	25–2.8	7.6 (20.1)
Gold	1.5418	Laboratory	22088	2.8	57.0 (46.8)	25–2.9	13.9 (52.8)
Mercury	1.5418	Laboratory	37985	2.9	62.8 (53.5)	25–2.8	8.2 (17.8)

† Crystal: space group $P2_12_12_1$; $a = 117.96$, $b = 121.30$, $c = 126.39$ Å; one trimer per asymmetric unit. Native, native data with frames refined from 87 to 1.65 Å; Native$_{1cjd}$, native data with frames refined over 35–1.85 Å and used to produce PDB code 1cjd (Benson *et al.*, 1999); Native$_{lab}$, native data collected on laboratory source; Uranium1, uranyl acetate derivative; Uranium2, uranyl nitrate derivative; Platinum, potassium tetrachloroplatinate (II) derivative; Gold, potassium dicyanoaurate (I) derivative; Mercury, mercuric acetate derivative. ‡ Average number of measurements for each unique reflection. § $R_{sym} = \sum |I_h - \langle I_h \rangle| / \sum I_h$, where I_h is the measured intensity of an individual reflection and $\langle I_h \rangle$ is the average intensity of that reflection.

X-PLOR (Brünger, 1992). In the resolution range 8–4 Å, a peak 5σ above background at $\kappa = 120°$ indicated a non-crystallographic threefold axis with polar angles $\varphi = 55.0$ and $\psi = 45.0°$. The molecular threefold axis thus lies almost along a body diagonal of the unit cell and the arrangement is therefore nearly that of the cubic space group $P2_13$, in which there would be one subunit per asymmetric unit.

2.3. Heavy-atom attempts and MIR phasing

Crystals were soaked in potential heavy-atom derivatives at concentrations of 0.1–10.0 mM for periods of a few minutes to over a week. Usually, the crystals dissolved, shattered or became non-isomorphous. The pH of the native buffer solution (4.1–4.3) was fairly low, but dropped to ~3.9 once equilibrium was established. This is low for favorable binding of most heavy atoms, but if the pH was raised the crystals dissolved. It is interesting that similar difficulties were experienced in the heavy-atom trials of adenovirus type 2 hexon (Burnett *et al.*, 1985). Diffraction data were collected with the Siemens area detector from the few crystals that withstood the soaking process. These included crystals soaked in uranyl acetate, uranyl nitrate, potassium tetrachloroplatinate (II), potassium dicyanoaurate (I) or mercuric acetate (Table 1). Data were also collected in-house on crystals of the Se-Met P3 derivative. Only the uranium derivatives could be solved through difference Patterson maps calculated using various programs from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) and *PHASES* (Furey & Swaminathan, 1997). Initially, two major sites were identified for the uranyl acetate derivative. These were related by the threefold non-crystallographic symmetry (NCS) known from the self-rotation analysis, but the expected third site was not present. When these two sites were refined with the

program *MLPHARE* (Otwinowski, 1991), four additional sites were discovered through difference Fourier maps. A group of three were related by the NCS, but the fourth was independent, though not on the molecular threefold axis.

The same two major uranium-binding sites found for uranyl acetate could also be identified in difference Patterson maps for a uranyl nitrate derivative. An additional seven sites were found using difference Fourier maps. Six sites were identical to those in the uranyl acetate derivative, but three new sites formed a group related by the molecular threefold. Heavy-atom binding sites in other derivatives could only be found using cross-difference Fourier maps calculated with the phases from the uranyl acetate derivative. The potassium tetrachloroplatinate (II) derivative provided seven sites, with one on the molecular threefold axis and two pairs of threefold-related sites. Six sites were found in the potassium dicyanoaurate (I) derivative and six in the mercuric acetate derivative, each with a pair of threefold-related sites. The platinum sites were distinct from those in the other derivatives. One of the pairs of threefold-related sites in the gold and mercury derivatives were identical, but the other three sites of each were unique. All the derivatives were refined with *MLPHARE* (Otwinowski, 1991) to provide multiple isomorphous replacement (MIR) phases to 2.9 Å resolution. The majority of the phasing power arose from the uranium derivatives (Table 2), but each derivative helped to diminish noise in the electron-density map. Phases calculated by SIR with uranium or by MIR revealed 12 of the predicted 15 selenium sites in the Se-Met P3 derivative from data collected on the laboratory X-ray source. The positions of the selenium sites were used to refine the location of the threefold NCS axis that was used in map averaging.

2.4. MAD data collection and phasing

Se-Met P3 crystallized under identical conditions as native P3, with the same unit-cell parameters. MAD data were collected at beamline X12C at the National Synchrotron Light Source (NSLS) from a single Se-Met P3 crystal maintained at 100 K. Data at three wavelengths around the absorption edge of selenium were collected to 2.2 Å resolution using a Brandeis CCD-based detector. The three wavelengths chosen were L1 ($\lambda = 0.9790$ Å) at the inflection point, L2 ($\lambda = 0.9786$ Å) at the absorption maximum and L3 ($\lambda = 0.9500$ Å) at a wavelength remote from the edge (Table 1). The synchrotron data were processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). While most of the 15 selenium sites

Table 2
Phasing statistics.

Numbers in parentheses are values in the highest resolution shell.

	No. of sites	Resolution (Å)	R_{Cullis}†	R_{Kraut} (iso)‡	R_{Kraut} (ano)§	Phasing power (iso)¶	Phasing power (ano)††
Se-Met L1	15	2.2	0.585	0.028	0.793	1.49 (1.21)	2.86 (1.96)
Se-Met L2	15	2.2	0.647	0.027	0.213	1.25 (1.12)	3.25 (2.31)
Se-Met L3	15	2.2			0.025		1.28 (0.74)
Uranium1	6	2.9	0.760			1.35 (0.97)	
Uranium2	9	2.9	0.760			1.36 (1.14)	
Platinum	7	2.9	0.970			0.38 (0.25)	
Gold	6	2.9	0.920			0.61 (0.37)	
Mercury	6	2.9	0.880			0.98 (0.89)	

† $R_{Cullis} = \sum |||F_{PH}(\text{obs})| - |F_P(\text{obs})|| - |F_H(\text{calc})||/\sum ||F_{PH}(\text{obs})| - |F_P(\text{obs})||$, where $F_{PH}(\text{obs})$ and $F_P(\text{obs})$ are the derivative and the native structure factors, respectively, and $F_H(\text{calc})$ is the calculated heavy-atom structure factor. The summation is over the centric reflections only. ‡ R_{Kraut} (iso) $= \sum ||F_{PH}(\text{obs})| - |F_{PH}(\text{calc})||/\sum |F_{PH}(\text{obs})|$, summed over all acentric reflections for the isomorphous case. § R_{Kraut} (ano) $= \sum ||F^+_{PH}(\text{obs})| - |F^+_{PH}(\text{calc})|| + ||F^-_{PH}(\text{obs})| - |F^-_{PH}(\text{calc})||/\sum ||F^+_{PH}(\text{obs})| - |F^-_{PH}(\text{obs})||$, summed over all acentric reflections for the anomalous case. ¶ Phasing power (iso) $= \sum |F_H|/\sum ||F_{PH}(\text{obs})| - |F_H(\text{calc})||$, for the isomorphous case. †† Phasing power (ano) $= \sum |F''_H|/\sum ||F^+_{PH}(\text{obs})| - |F^+_{PH}(\text{calc})|| + ||F^-_{PH}(\text{obs})| - |F^-_{PH}(\text{calc})||$, for the anomalous case.

appeared in cross-difference Fourier maps between L1 and L3, L1 and L2 or L2 and L3 using the uranyl acetate phases, all could easily be identified in cross-difference anomalous Fourier maps using L1 or L2. The selenium sites could not be found using only Patterson maps.

Native phases were calculated using the MAD method and L3 as the reference wavelength, as this gave the smallest anomalous signal from selenium and its data set was the most complete. The phasing of the MAD data (Table 2) was handled as an extension of MIR (Ramakrishnan & Biou, 1997) using the programs *PHASES* (Furey & Swaminathan, 1997) and *MLPHARE* (Otwinowski, 1991). Although the electron-density maps resulting from these programs were indistinguishable, the corresponding statistics were very different, as *MLPHARE* showed no improvement in phasing statistics upon refinement of the selenium positions. The MAD phases were good to the 2.2 Å resolution limit of the data set. Using phase combination to include the MIR data did not improve the electron density and so only MAD phases were used for map building and refinement. The initial electron density for the molecule was solvent flattened and threefold averaged with the program *DM* (Cowtan, 1994), resulting in a noticeable improvement in map quality.

2.5. Native P3 data collection at NSLS

Data from a frozen native P3 crystal were also collected at NSLS beamline X12C. As strong reflections initially were observed to the edge of the detector, its 2θ angle was offset by $5°$, whereupon diffraction was observed past 1.6 Å resolution. Initially, the data frames were processed at the synchrotron with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997), using the conservative resolution range of 35.0–1.85 Å suggested by the beamline software and were used to refine the native molecule (PDB code 1cjd; Benson *et al.*, 1999). Recently, the data frames were reprocessed so that all

recorded reflections were integrated. This gave the set of reflections from 87 to 1.65 Å resolution used in the current refinement (Table 1). The reflections beyond 1.65 Å were excluded, as they were weak and incomplete.

2.6. Se-Met P3 refinement

The MAD-phased electron-density map was calculated and an Se-Met P3 model was built with the baton feature of the program *O* (Jones *et al.*, 1991). *O* was used for all manual model building and for visualization of the molecule and electron-density maps. The first model covered 320 residues (81%) of one subunit of the trimer and it was only necessary to truncate the side chains of 14 residues to alanine because of poor density. The model was refined using simulated annealing for reflections in the resolution range 35.0–2.2 Å against the L3 set of the MAD data with a 2σ data cutoff (70 927 reflections) with the program *X-PLOR* (Brünger, 1992). A bulk-solvent correction (Jiang & Brünger, 1994) and strict NCS were imposed. After three rounds of refinement using strict NCS and manual model building, the other two subunits of the Se-Met P3 molecule were generated to produce an exact trimer and complete an asymmetric unit of the crystal. One round of refinement was performed with tight NCS restraints. The NCS restraints were then modified with a script that placed a modest weight of 419 kJ mol^{-1} Å^{-2} on the parts of the molecule that appeared to best obey the NCS. Side-chain atoms and areas of the model with more apparent flexibility were given a lower weight (21–209 kJ mol^{-1} Å^{-2}). Areas of the molecule that were difficult to fit were unrestrained. Water molecules were added, initially automatically with programs from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) and then manually within the *O* program. Manual model building using *O* followed cycles of refinement with *X-PLOR*. The Se-Met P3 model used to phase the native data contained 1094 residues and 429 waters with $R_{work} = 20.0\%$ and $R_{free} = 23.5\%$.

Subsequent rounds of refinement were performed with the *Crystallography and NMR System* (*CNS*) program (Brunger *et al.*, 1998). All data were included (82 172 reflections) in the simulated-annealing refinement with a maximum-likelihood target (Pannu & Read, 1996) and all NCS restraints were removed. The final model of Se-Met P3 at 2.2 Å resolution contains 1111 residues and 437 waters, with $R_{work} = 20.1\%$ and $R_{free} = 22.8\%$ (Table 3). Two residues are in disallowed regions: Asn17^I (modeled as Ala), the second residue at the N-terminus, and Leu281^{II}, which is in a turn in the CD2 loop and has its side-chain solvent exposed. Of the remaining residues, 85.5% are in the most favored regions and the rest in the additionally allowed regions. Residues with poor side-chain definition in the electron density were truncated to Ala. These include ten residues in subunit I (Arg16, Asn17, Gln18,

Table 3
Model statistics for native and Se-Met P3 structures.

	Native P3	Se-Met P3
Resolution (Å)	87–1.65	35.0–2.2
Total refined reflections	179469	82172
Reflections in free set	2072	978
R_{work}† (%)	19.0	20.1
R_{free}† (%)	20.8	22.8
R.m.s. bond deviation (Å)	0.009	0.007
R.m.s. angle deviation (°)	1.58	1.46
No. of protein atoms (in alternate conformations)	8608 (487)	8540
Average B factor (Å^2)	25.8	27.1
No. of water molecules	782	437
Average B factor (Å^2)	33.6	27.0
No. of MPD molecules	18	—
Average B factor (Å^2)	56.7	—
No. of Cl$^-$ atoms	2	—
Average B factor (Å^2)	26.2	—
No. of Na$^+$ atoms	6	—
Average B factor (Å^2)	28.0	—

† $R_{\mathrm{work}} = \sum |F_{\mathrm{obs}} - F_{\mathrm{calc}}| / \sum F_{\mathrm{obs}}$, where summation is over data used in the refinement; R_{free} includes only the 1.1% of data excluded from all refinements.

Gln19, Gln187, Asn312, Phe313, Gln354, Asn357 and Thr384), 14 residues in subunit II (Leu15, Arg16, Asn17, Gln18, Gln19, Gln33, Gln187, Gln244, Asn245, Asn312, Phe313, Gln354, Tyr355 and Thr384) and ten residues in subunit III (Leu15, Arg16, Asn17, Gln19, Gln244, Asn245, Asn312, Phe313, Gln354 and Glu385).

2.7. Native P3 refinement

Refinement of native P3 commenced with the Se-Met P3 model from *X-PLOR* with its waters removed. This refinement used reflections from 35.0 to 1.85 Å resolution with a 2σ data cutoff (121 093 reflections). The modified NCS restraint script was employed and manual model building followed each refinement. Water molecules were introduced as for the Se-Met model and the NCS restraints were removed. The 1.85 Å resolution native P3 model previously reported (PDB code 1cjd; Benson *et al.*, 1999) contained 1108 residues and 473 water molecules with $R_{\mathrm{work}} = 17.8\%$ and $R_{\mathrm{free}} = 20.6\%$.

After the native data frames were reprocessed, refinement continued with *CNS* (Brunger *et al.*, 1998) using all reflections (179 469) in the resolution range 87–1.65 Å and a maximum-likelihood target (Pannu & Read, 1996). After a few rounds of refinement and manual model building, more water molecules were added. Later, ions and MPD molecules were also added and refined. At this stage, some residues were built with alternate conformations. The current native P3 model contains 1115 residues, 61 with alternative conformations, 782 water molecules, two chloride ions, six sodium ions and 18 MPD molecules. The residuals are $R_{\mathrm{work}} = 19.0\%$ and $R_{\mathrm{free}} = 20.8\%$ (Table 3). Four residues are in disallowed regions. The first is Asn17$^{\mathrm{I}}$ at the N-terminus. The others cluster in the CD2 loop (Leu181$^{\mathrm{I}}$, Leu181$^{\mathrm{II}}$and Gln187$^{\mathrm{I}}$) where the leucine residues are solvent exposed and the glutamine residue is kinked, but their φ,ψ angles are just out of the allowed area. Of the remaining residues, 87.8% are in the most favored area and 12.2% in additionally allowed regions. Some residues that

displayed poor side-chain density were truncated to Ala. These include eight residues in subunit I (Leu15, Arg16, Gln244, Asn245, Asn312, Phe313, Leu353 and Thr384), ten residues in subunit II (Gln11, Gln12, Arg16, Gln187, Gln244, Asn245, Asn312, Phe313, Leu353 and Thr384) and seven residues in subunit III (Arg16, Gln244, Asn245, Asn280, Asn312, Phe313 and Thr384). Additionally, the 61 residues that display alternate conformations include Leu25, Gln26, Thr44, Arg52, Asn79, Gln97, Val99, Asn106, Met133, Glu227, Ser229, Ser255, Ser263, Thr269, Gln277, Lys322, Thr352, Leu353, Gln354, Tyr355, Gly356, Asn357, Val358 and Gln369 in subunit I, Leu25, Gln26, Thr44, Arg98, Val99, Ile100, Asn106, Gln107, Ser113, Met133, Asp149, Glu174, Glu212, Glu227, Ser229, Asn238, Ser255, Thr269, Ile288, Lys322, Lys365 and Gln369 in subunit II, and Leu25, Thr44, Asp56, Val99, Glu111, Met133, Asp149, Thr163, Lys193, Asn200, Glu212, Ser229, Gln239, Gln277 and Gln369 in subunit III.

3. Results and discussion

3.1. Crystallization and structure determination

Crystallization conditions that solved earlier reproducibility problems (Stewart, Ghosh *et al.*, 1993) were found using the
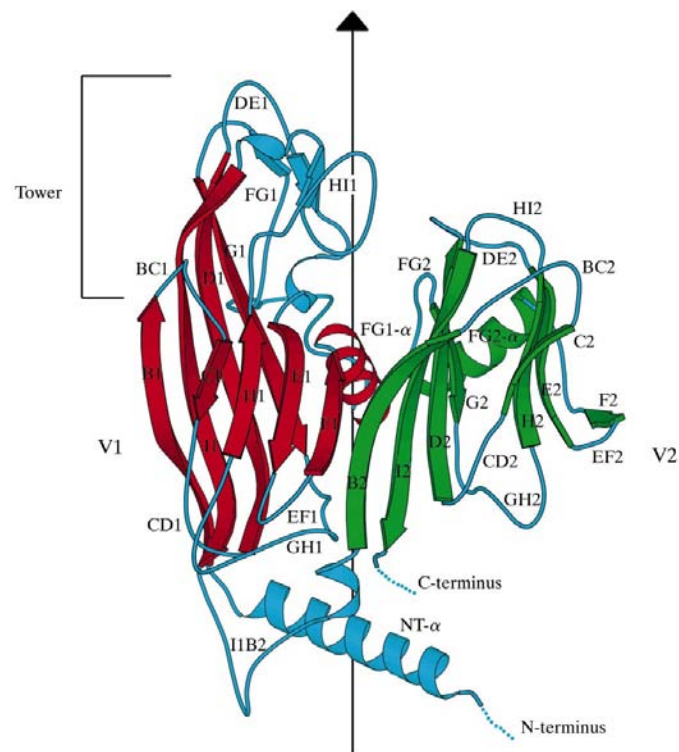

Figure 2
The P3 subunit showing its secondary structure. The β-strands (B–I) forming the two jelly rolls, V1 (red) and V2 (green), are labeled, as are the FG-α helices following each F strand. Also labeled are the N-terminal helix NT-α, the loops between strands of the jelly rolls, the tower region, and the N- and C-termini. In the virion, the tower lies to the outside and the base of the molecule is adjacent to the inner membrane. The line with the triangle is along the molecular threefold. The figure was produced with *MOLSCRIPT* (Kraulis, 1991).
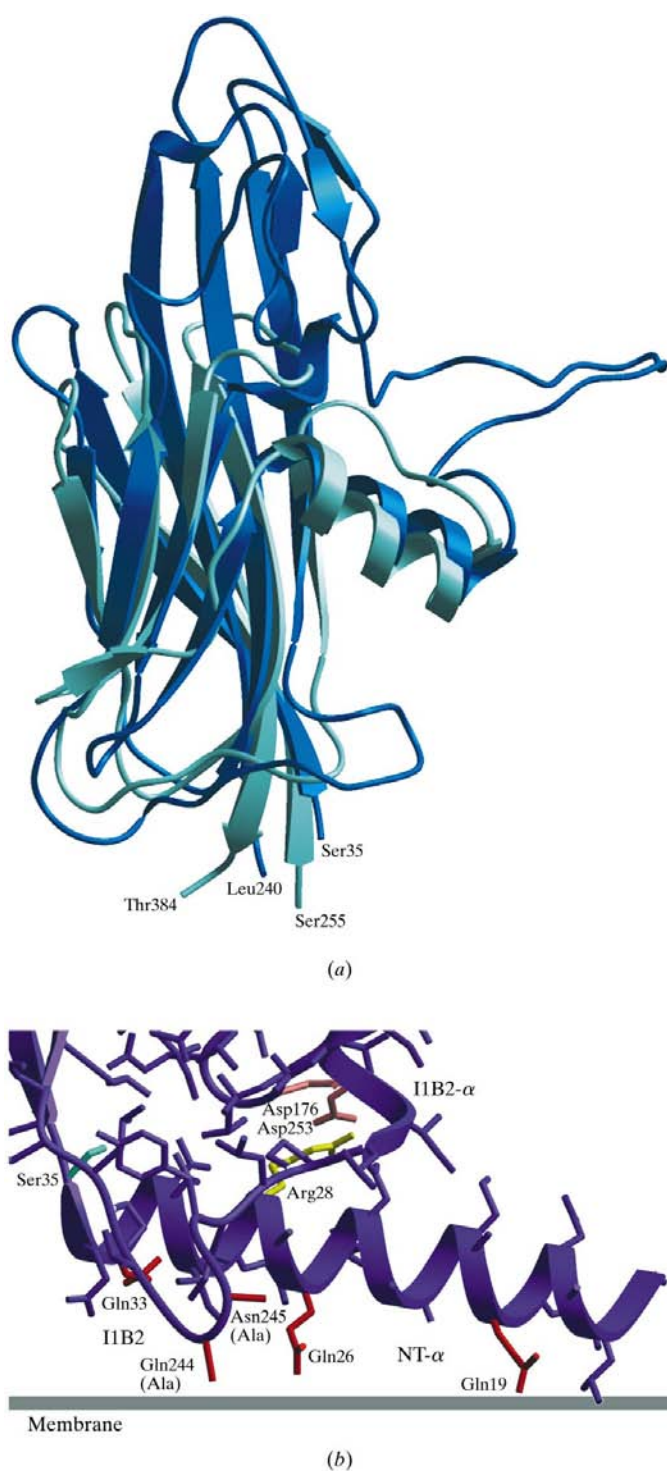
precipitant MPD, which produced large single crystals that diffracted well. These were X-ray sensitive, but could be flash-frozen without modification so that a complete data set at high resolution could be obtained from a single crystal.

Figure 3
Superposition of jelly rolls V1 and V2 in P3 and the molecular base region. (a) The overall folds of the jelly rolls are very similar, but V1 (dark blue) is taller than V2 (light blue). The alignment of the common sections of the strands and the α-helix is very similar, as is the separation between the two four-stranded sheets. (b) The N-terminal helix NT-α and loop I1B2 showing the hinge residue (Ser35, green) and the residues in red (Gln19, Gln26, Gln33, Gln244 and Asn245) that are capable of forming hydrogen bonds with the charged head groups of the lipids in the underlying membrane (244 and 245 are modeled as alanine). Arg28 (yellow) tethers NT-α to the base of the molecule by forming salt links with Asp176 and Asp253 (pink). The figure was produced with *SETOR* (Evans, 1993).
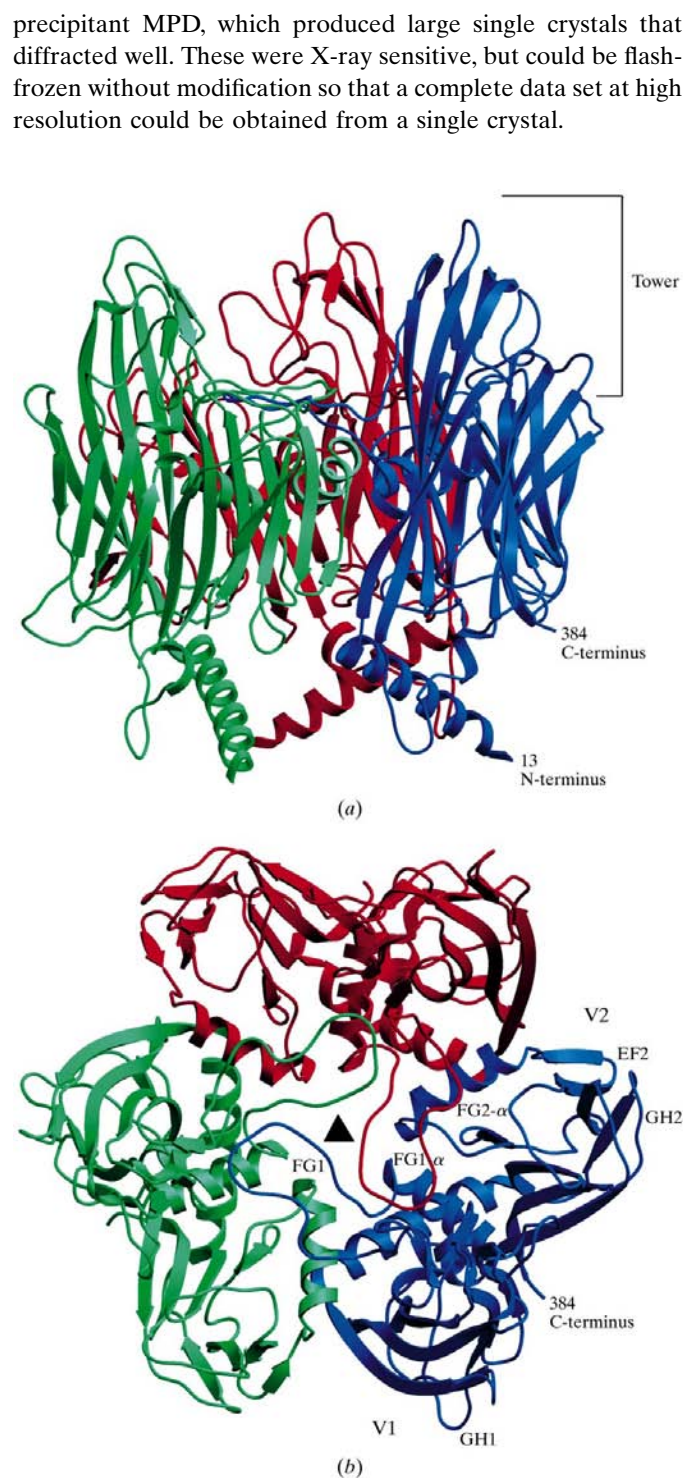

Figure 4
The P3 trimer. (a) A ribbon diagram of the trimer viewed almost parallel to the plane of the viral facet so that the towers that form the external surface of the capsid are facing up. The tower is marked, as are the N- and C-termini. (b) The trimer viewed from outside the virus along the molecular threefold (triangle). The trimer has a pseudo-hexagonal daisy shape with α-helices FG1-α and FG2-α lying between the jelly rolls (V1 and V2) like the spokes of a wheel. The trimerization loop from one subunit (FG1) interacts with its counterparts from the other two subunits to stabilize the trimer and cap the internal cavity. In the capsid, loops along the perimeter of the P3 trimer (GH1, EF2 and GH2) could interact with other P3 trimers or the vertex protein P31. The figure was produced with *SETOR* (Evans, 1993).

research papers

Table 4
Secondary-structural elements in P3.

Element	Residues	Region
NT-α	13–34	N-terminus
B1	36–47	V1
Loop BC1	48–53	Tower
C1	54–57	V1
Loop CD1	58–63	Molecular base
D1	64–78	V1
Loop DE1	79–95	Tower
DE1-β	85–87	Tower
DE1-α	91–94	Tower
E1	96–102	V1
Loop EF1	103–107	Molecular base
F1	108–113	V1
Loop FG1	114–158	Tower
FG1-α	114–125	V1
FG1-β	153–155	Tower
G1	159–169	V1
Loop GH1	170–190	Molecular base
H1	191–196	V1
Loop HI1	197–222	Tower
HI1-β	215–217	Tower
HI1-α	220–222	Tower
I1	223–238	V1
Loop I1B2	239–255	Molecular base
I1B2-α	251–254	Molecular base
B2	256–265	V2
Loop BC2	266–272	
C2	273–277	V2
Loop CD2	278–282	Molecular base
D2	283–293	V2
Loop DE2	294–302	
E2	303–310	V2
Loop EF2	311–312	Molecular base
F2	313–315	V2
Loop FG2	316–339	
FG2-α	321–332	V2
G2	340–344	V2
Loop GH2	345–357	Molecular base
H2	358–364	V2
Loop HI2	365–370	
I2	371–382	V2

After some difficulties in finding heavy-atom derivatives, a crystal soaked for 5 h in 0.1 mM uranyl acetate gave derivative data at 2.9 Å resolution (Table 1). Only two major sites were visible in the Harker sections of the difference Patterson map, although a self-rotation function analysis of the native data had predicted that a trimer should occupy the asymmetric unit of the crystal. Therefore, three heavy-atom sites would have been expected. A uranyl nitrate derivative (Tables 1 and 2) gave similar results. Cross-difference Fourier maps using uranyl acetate phases revealed heavy-atom sites for other derivatives that could not be solved from their difference Patterson maps: potassium tetrachloroplatinate (II), potassium dicyanoaurate (I) and mercuric acetate (Tables 1 and 2). Using different combinations of the derivatives, MIR-phased electron-density maps were calculated, solvent flattened and threefold averaged. They were noisy and the main-chain density was discontinuous, but portions of the model could be traced.

This approach was superseded by the production of an Se-Met P3 derivative from PRD1 grown in bacteria using a minimal medium that was pulsed with Se-Met at a late stage of infection (see §2). Crystals of Se-Met P3, which were isomorphous to native P3, were used for a MAD experiment (Table 1). The 15 possible selenium sites of the Se-Met P3 trimer could not easily be identified from difference Patterson maps, but single isomorphous replacement (SIR) phases from the uranyl acetate derivative immediately revealed all of them through both isomorphous and anomalous difference Fourier maps. The initial MAD-phased map was superb, with the majority of the phasing power arising from the anomalous component of the data (Table 2), and >80% of the chain for one subunit could be easily built to provide a starting model. The model was refined with *X-PLOR* (Brünger, 1992) and later with *CNS* (Brunger *et al.*, 1998) against data at the remote wavelength, L3 (λ = 0.9500 Å), of the MAD experiment. The final Se-Met P3 model at 2.2 Å resolution has R_{work} = 20.1% and R_{free} = 22.8% (Table 3).

Synchrotron data collected for a native P3 crystal were initially processed over the conservative resolution range of 35.0–1.85 Å and used with *X-PLOR* (Brünger, 1992) to refine the molecule (PDB code 1cjd; Benson *et al.*, 1999). This process started with an Se-Met P3 trimer model at a late stage of refinement and revealed a curious result. The native and Se-Met P3 molecules have a relative rotation of ~1.0° about their molecular threefold axes, even though their unit cells are essentially identical.

As the synchrotron diffraction data were strong and complete to 1.85 Å resolution and reflections were observed beyond that point, it became evident that the native crystal diffracted to higher resolution than was first apparent. When the data frames were reprocessed at home, data were seen to extend from 87 Å to beyond 1.6 Å. As the data are incomplete and weak at the highest resolution, a limit of 1.65 Å was imposed. The P3 structure reported here was refined with *CNS* (Brunger *et al.*, 1998) using the maximum-likelihood target (Pannu & Read, 1996), which reduced model bias and produced electron density that could be interpreted more accurately. Improvement in some of the previously trouble-some areas was particularly noticeable. The resultant $2F_o - F_c$ electron-density map at 1.65 Å resolution displays more detail and so more precisely defines some features of the crystal structure. These include alternative conformations for some residues and an improved description of ordered solvent, including precipitant molecules. The final native P3 model has R_{work} = 19.0% and R_{free} = 20.8% (Table 3).

3.2. P3 subunit structure

The initial methionine residue of the 395 amino-acid P3 protein is cleaved after translation (Bamford & Bamford, 1990), but its remaining 394 residues are numbered 2–395 for consistency with published biochemical and genetic data. As the complete P3 trimer lies in the crystallographic asymmetric unit, its subunits (I, II and III) are independent. The electron density for subunit I is contiguous over residues 15–384, for II over residues 11–384 and for III over residues 14–384. As both the N- and C-terminal residues are disordered, they were not modeled. In addition, 25 residues were modeled as alanine as

their side-chain density was not well defined or missing. All these residues fall in areas with higher thermal mobility than the rest of the molecule, such as the N-terminal α-helix, the turns between β-strands and the last modeled residue at the C-terminus. The minor differences between the residues in these sections seem solely to arise from their high thermal mobility and their correspondingly weak density.

The main feature of the P3 subunit is the 'double-barrel' of two eight-stranded viral β-barrels or jelly rolls (V1 and V2; Fig. 2). The jelly-roll fold is a common motif in capsid proteins from a wide array of viruses infecting bacteria, plants, and animals (Chelvanayagam *et al.*, 1992; McKenna *et al.*, 1992; Rossmann & Johnson, 1989). Although the two jelly rolls of P3 have the same topology, their sequences exhibit no apparent conservation. Their antiparallel β-strands are labeled B through I (Table 4; Fig. 2), following the first example discovered in tomato bushy stunt virus (TBSV; Harrison *et al.*, 1978). The loops between the strands are named after the strands they connect and any secondary structure in the loop is named correspondingly (Table 4). For example, the DE1 loop occurs between D1 and E1 and its β-strand is DE1-β.

The two jelly rolls have the same width (15 Å), but V1 is taller than V2 (47 Å *versus* 31 Å). Each contains a 12-residue α-helix that follows its F-strand, FG1-α (114–125) in V1 and FG2-α (321–332) in V2, and is perpendicular to the β-strands of its barrel. FG1-α is positioned between V1 and V2, where it locks the two jelly rolls together. FG2-α is in an analogous position in the trimer, where it lies between V2 in one subunit and V1 in the neighboring subunit (V1′). The portion of FG1 which immediately follows FG1-α forms a long 'trimerization loop' (126–158) that makes an important contribution to the stability of the trimer. This loop extends out to overlap one neighboring subunit and to interact with its counterpart (FG1″) in the third subunit (see Fig. 4b). The analogous portion of the FG2 loop in V2 (333–339) is far shorter and does not interact with other subunits. The viral jelly roll can also be viewed as a β-sandwich, where B, I, D and G form one sheet and C, H, E and F form the other sheet. The two sheets are separated as the hydrogen bonding is not continuous: B does not hydrogen bond with C, or F with G. Additionally, although β-strands within sheets are roughly parallel, the sheets themselves have a slight offset and are not exactly parallel. As the fold of V1 and V2 is the same, the jelly rolls can easily be superimposed with a root-mean-square deviation of 1.30 Å for their 84 common C$^\alpha$ positions (Fig. 3a). The mutual orientation of the sheets and their separation is roughly the same in both barrels.

Three surface protrusions around the edge of the P3 trimer are visible in the EM reconstruction of the PRD1 virion (Butcher *et al.*, 1995). These correspond to the 'towers' formed at the top of the molecule by the extended loops above V1 (Fig. 2). The tower is composed of the last few residues of D1 (74–78), the DE1 loop (79–95), the end of the FG1 loop (149–158), the HI1 loop (197–222) and the first few residues of G1 (159–162) and I1 (223–227). The tower rises ~22 Å above the V1 jelly roll and is analogous to the tower in adenovirus hexon, although its position relative to the base is somewhat

different. The hexon tower is more complicated as both jelly rolls contribute loops and an individual tower is formed by three interacting loops from different subunits.

The N-terminal α-helix and the I1B2 loop extend below the jelly rolls and are positioned where they can interact with the membrane in the intact virion (Fig. 3b). From results on the related bacteriophage PR4, the viral membrane is composed mainly of two lipids with charged head groups: phosphatidylethanolamine (56% of the membrane) and phosphatidylglycerol (37%) (Davis *et al.*, 1982). Cross-linking experiments on PR4 that targeted these lipids show that P3 is in direct contact with the membrane and has a preference for phosphatidylglycerol (Davis & Cronan, 1985). Although P3 could also contact one or more of the ~15 membrane-bound proteins, none were identified in protein–protein cross-linking experiments (Luo *et al.*, 1993). The I1B2 loop (239–255) extends 20 Å beneath the jelly rolls and has a sharp β-turn (Gly243, Gln244, Asn245 and Gly246) at its lowest point (Fig. 3b). The glutamine and asparagine could contact the membrane and form hydrogen bonds with charged lipid head groups. The N-terminal α-helix (NT-α) extends 22 residues (13–34) from the first residue visible in the electron density and drops ~18 Å below V1. It contains a strip of glutamine residues (Gln19, Gln26 and Gln33) along one of its sides where they also could contact the membrane. Thus, six patches of polar residues lie in the same plane beneath the P3 trimer, the NT-α helix and the I1B2 loop from each subunit, where they can interact with the viral membrane (Fig. 3b). Recently, the EM reconstruction of the PRD1 virion has been extended to higher resolution and a fitting of the P3 trimer into its EM density supports the idea that NT-α and I1B2 contact the membrane (San Martín *et al.*, 2001).

Laser Raman spectroscopy experiments on P3 trimers in solution showed a secondary-structure content of $18 \pm 6\%$ (~71 residues) α-helix and $44 \pm 5\%$ (~173 residues) β-strand (Tuma *et al.*, 1996). The secondary structure in the X-ray model, defined with the program *DSSP* (Kabsch & Sander, 1983), is 14.5% (57 residues) α-helix and 37.3% (147 residues) β-strand, based on the 394 residues of the subunit. The 20 disordered residues at the N- and C-termini were obviously not assigned to secondary structure in the X-ray model. However, weak electron density suggests that the N-terminus could contain additional α-helical content. The good agreement between the two sets of experimental values indicates that P3 trimers have comparable conformations in solution and in the crystal.

3.3. Correlation of the P3 trimer with EM reconstructions of PRD1 virions

The morphology of the P3 trimer was first indicated by cryo-EM image reconstructions of PRD1 virions. Density presumed to correspond to the P3 trimer was isolated from the 28 Å resolution EM image reconstructions of the entire virion and the 20 Å resolution P3 shell (Butcher *et al.*, 1995) to reveal a molecule with three towers and a large central cavity. The EM density agrees very well with the X-ray model. This corre-

spondence and evidence from antibody-binding experiments (Bamford *et al.*, 1993) that shows the N-terminus is internal to the virion positions the X-ray structure relative to the viral capsid. The trimer sits with the N-terminal α-helix (NT-α) and the I1B2 loop pointed toward the viral membrane, with the towers exposed on the surface of the capsid (Fig. 4a). The
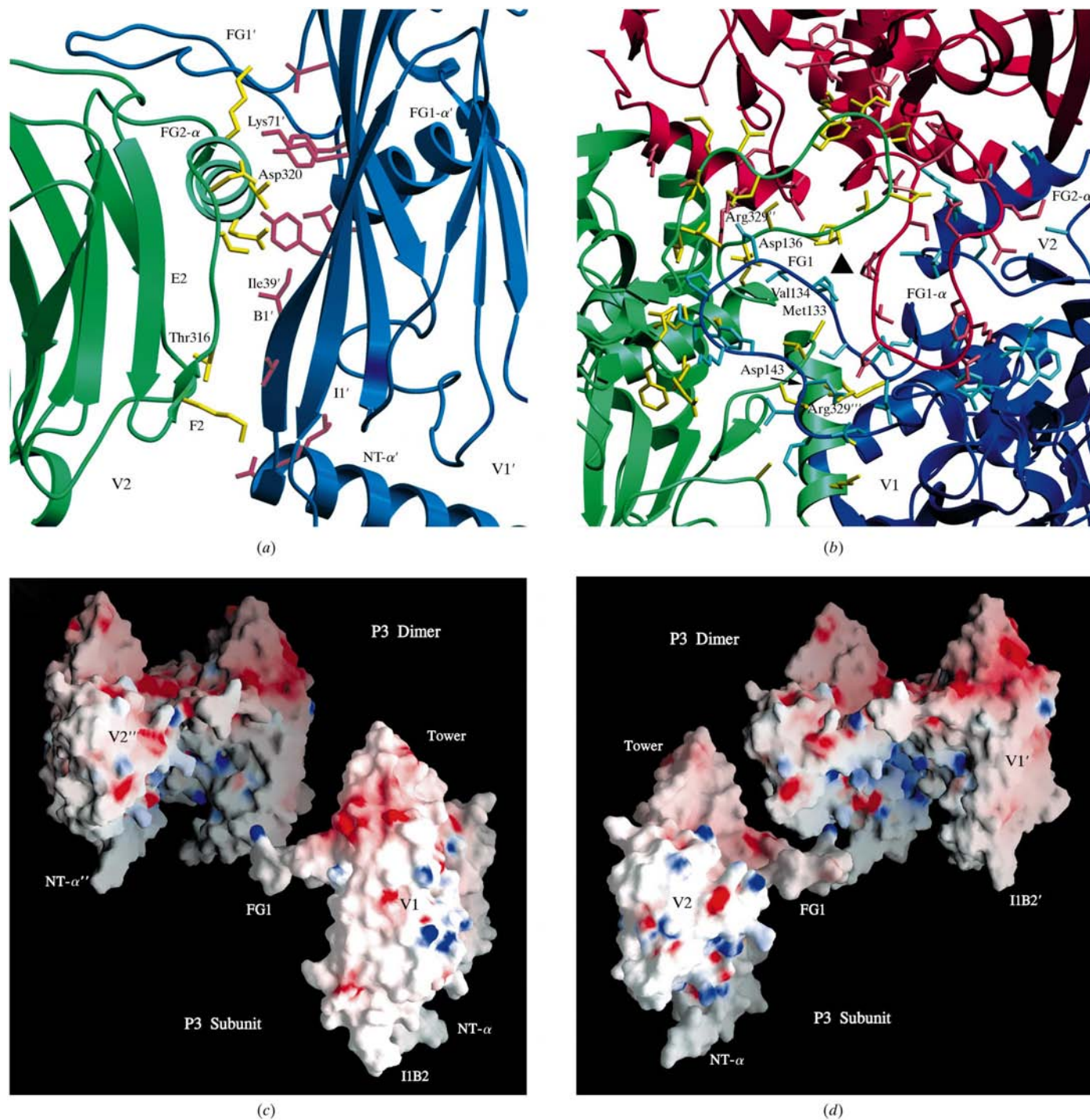

Figure 5
Interactions within the P3 trimer. (*a*) Residues in the interface between two subunits that are within 3.5 Å of each other (Table 5). The atoms of Thr316 and Ile39′ come into closest main-chain contact and the side chains of Asp320 and Lys71′ form a salt link. (*b*) The P3 trimer viewed as in Fig. 4(*b*). The residues within 3.5 Å of each other that are involved in interactions along the FG1 loop (Table 6). Asp136 and Arg329″ form salt links, as do Asp143 and Arg329‴. The hydrophobic residues Met133 and Val134 seal the pore at the molecular threefold (triangle). (*c*) A surface representation of P3 with the electrostatic potential displayed. The blue color is more positive and red more negative. A single subunit has been extracted and rotated so the interacting face of the V1 jelly roll can be viewed. It would fit back in with this label against the V2″ label. (*d*) Similar to (*c*) but viewing the subunit from the V2 side, which will fit back against V1′. No large patches of complementary charges are observed. (*a*) and (*b*) were produced with *SETOR* (Evans, 1993) and (*c*) and (*d*) with *GRASP* (Nicholls *et al.*, 1991).

helices FG1-α and FG2-α radiate out from the molecular threefold like spokes of a wheel to lie between the jelly rolls and form a jelly roll–helix–jelly roll–helix alternating pattern (Fig. 4b). Thus, in the trimer, when viewed from the top, the molecule appears as a daisy with a jelly roll forming each of the six petals. This pseudo-hexagonal shape allows P3 molecules to fill the hexavalent positions on the viral capsid and interdigitate to form a closely packed surface. While all the trimers in each facet form a p3 array, the arrays in adjacent facets are rotated by 60° at their mutual edge so that the P3 molecular contacts within a facet are different to those between facets. This arrangement is equivalent to that for hexon in the adenovirus virion (Burnett, 1985; Oostrum et al., 1987).

Negatively stained P3 trimers appear as hollow rings in the electron microscope (Bamford & Mindich, 1982). The EM image reconstruction showed that this feature arose from the large central cavity of the P3 trimer, but also revealed that a thin layer of density caps the protein cylinder (Butcher et al., 1995). The P3 X-ray structure shows the origins of these morphological features. The six jelly rolls in the trimer form a ring with a large cavity in its center, but the FG1 loops of each subunit approach each other to create a cap over the upper portion of this tube (Fig. 4b). These loops closely approach to form a small closed pore at the central threefold molecular axis.

3.4. P3 is a stable trimer

P3 forms a very stable trimer and must be boiled in sodium dodecyl sulfate (SDS) to dissociate (Mindich, Bamford, McGraw et al., 1982), but it cannot be refolded into a trimer (Bamford & Bamford, 1990), nor are there reports of the soluble monomeric form. In fact, the bacterial GroEL/ES chaperonins are required to assist in proper trimer formation after P3 translation (Hänninen et al., 1997). The stability of P3 can be understood from its subunit interactions, although these are restricted to two areas. The first is between V2 and helix FG2-α in one subunit and V1′ in the next subunit. This contact area is not that extensive (Table 5; Fig. 5a). Only one salt link is observed, between Asp320 of FG2-α and Lys71′ of D1′. The β-strands F2 and B1′ come into closest contact, although only the main-chain atoms of Thr316 approach within hydrogen-bonding distance of the main-chain atoms of Val38′ and Ile39′ (Fig. 5a). F2 and B1′ are not aligned for extensive hydrogen bonding to occur, with only one being observed between the main-chain N atom of Thr316 and the main-chain O atom of Ile39′, and so they do not form a continuous β-sheet.

As trimerization still seems to occur when most of V2, including the F2 strand and FG2-α, is eliminated by engineered C-terminal truncations (Bamford et al., 1993), this part is not critical. That V2 does not mediate trimerization is consistent with the fact that most major stabilizing interactions occur at the second area of contact through the trimerization loop FG1 (Table 6; Fig. 5b). The FG1 loop from one subunit

Table 5
Intersubunit contacts between V2 and V1′.

Contacts are defined as residue pairs with main-chain or side-chain atoms within 3.5 Å.

Residues in V2 and FG2-α	Residues in V1′
Thr316	Val38, Ile39
Leu319	Tyr233
Asp320	Lys71, Tyr233
Lys322	Met145, Asn146, Tyr165
Thr323	Tyr165
Gln327	Leu69
Lys348	Leu32, Gln33, Ser35

overlaps a neighboring subunit and extends sufficiently far that it comes into contact with the tip of the FG1 loop of the third subunit (Fig. 5b). Most of the residues of this loop form hydrogen bonds, including water-mediated ones, or engage in van der Waals contacts. However, there are no apparent unusual interactions. Of the three charged residues of FG1, only two make salt links: Asp136[I] with Arg329[II] of subunit II and Asp143[I] with Arg329[III] of subunit III. Lys140, the third charged residue, has its N$^\zeta$ directed into solvent. A surface representation of the electrostatic potential of P3 shows very little charge complementarity for either contact region (Figs. 5c and 5d).

The total surface area of one P3 subunit is ~17 195 Å^2, but that of the trimer is 41 520 Å^2. Therefore, 10 065 Å^2, or
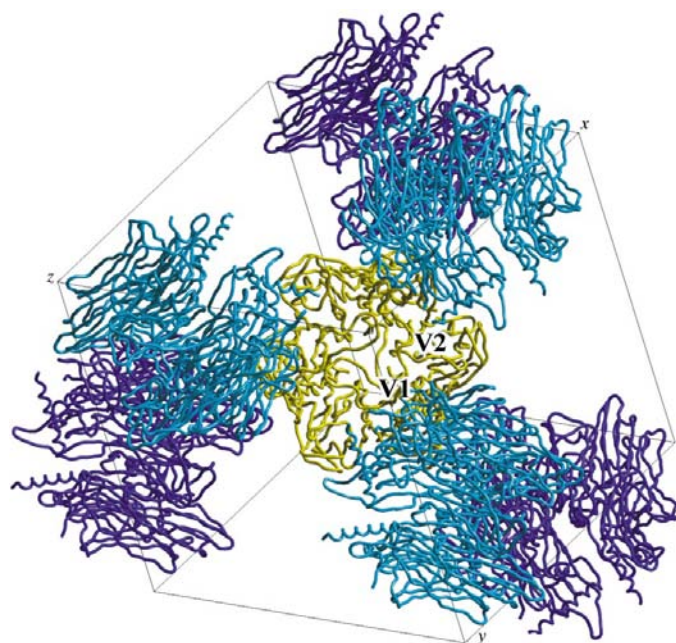

Figure 6
A view of the crystal packing of P3 in the $P2_12_12_1$ unit cell. Looking down the threefold axis of the central molecule (yellow), it can be seen that this axis is aligned almost along the body diagonal of the unit cell. The symmetry-related molecules interact in a tetrahedral arrangement that is close to that in space group $P2_13$. The interactions with the front three molecules (blue) and the back three (magenta) are with V1 of the central molecule (Table 7). V2 and the top and bottom of the molecule are exposed to solvent. The figure was produced with SETOR (Evans, 1993).

Table 6
Contacts between FG1 loop and the other subunits.

Contacts are defined as residue pairs with main-chain or side-chain atoms within 3.5 Å. The residue's superscript identifies its subunit.

Residues in FG1	Residues in the other subunits
Gln125[I]	Arg330[III]
Leu130[I]	Ala326[III], Arg329[III]
Ser131[I]	Thr135[II], Asp336[II], Ala333[III], Thr334[III]
Met133[I]	Met133[II], Met133[III], Ala333[III]
Val134[I]	Val134[II], Val134[III]
Thr135[I]	Ser132[II]
Asp136[I]	Arg329[II], Ser132[II], Asp143[III]
Ile139[I]	Phe90[III]
Lys140[I]	Leu87[III], Asp89[III]
Tyr141[I]	His118[III], Phe119[III], Thr122[III], Asp335[III]
Gly142[I]	Thr334[III], Asp335[III]
Asp143[I]	Asp136[II], Arg329[III]
Val144[I]	Ala299[III], Asp335[III]
Met145[I]	Ala299[III], Lys322[III], Asp335[III]
Asn146[I]	Lys322[III]
Val147[I]	Arg329[III]
Ile148[I]	Pro138[II]
Thr163[I]	Lys322[III]
Tyr165[I]	Lys322[III], Thr323[III]

3355 Å^2 per subunit, is buried upon trimer formation. The interaction between V2 and FG2-α with V1′ accounts for only 569 Å^2, so the two edges of the subunit would account for 1138 Å^2 (34%) of the total buried area, meaning that most of the buried surface is formed by FG1. The total (3355 Å^2) is more than the average of 1600 ($\pm$400) Å^2 found for protein–protein recognition sites (Lo Conte *et al.*, 1999), but only represents ~20% of the surface area of each P3 subunit, which is abnormally large due to the internal cavity. In the trimer, the trimerization loops overlap like the flaps on a cardboard box to lock the subunits together (Fig. 5*b*). This entanglement of the subunits makes it difficult to disrupt the molecule and is probably the most important factor in stabilization. However, all points of interaction contribute to make P3 a protein of unusually high stability.

3.5. Crystal contacts

The crystal contacts reflect some flexibility in the P3 trimer. The P3 molecule crystallizes in the orthorhombic space group $P2_12_12_1$ (*a* = 118.0, *b* = 121.3, *c* = 126.4 Å) with a trimer in the asymmetric unit. The solvent content of ~65% is on the high side for macromolecular crystals, which usually range from 27 to 65% (Matthews, 1968). P3 packs in the crystal with a tetrahedral pattern (Fig. 6) reminiscent of adenovirus hexon (Athappilly *et al.*, 1994; Rux & Burnett, 2000). The interactions between the P3 molecules, as for hexons, are not the same in the crystal as they are in the viral capsid (Butcher *et al.*, 1995; San Martín *et al.*, 2001). Hexon packs in the cubic space group $P2_13$, with a trimer axis aligned with a crystallographic threefold along the body diagonal of the unit cell, thus having a single subunit in the asymmetric unit. As the hexon subunits obey the crystal symmetry, they appear identical in the elec-

tron density. P3 almost forms a cubic unit cell, as its molecular threefold is close to a body diagonal, but the lower symmetry of its packing may reflect a molecule with more flexibility than hexon.

In P3, the molecular threefold provides only local and not crystallographic symmetry and so the intermolecular contacts made by each subunit are slightly different (Table 7). This difference is evident in the distribution of major heavy-atom binding sites in the uranyl acetate and uranyl nitrate derivatives. The uranium ion associates with Glu195 of the H1 strand in two of the three subunits. These glutamate residues lie on the exterior of V1, but they are not completely solvent accessible owing to the close proximity of symmetry-related P3 molecules in the crystal. The Glu195 residue from subunit II (Glu195[II]) binds a uranium ion with the aid of Glu224[III] in a symmetry-related molecule, as do Glu195[III] and Glu224[I]. However, the third pair, Glu195[I] and Glu224[II], are not close enough to jointly bind a uranium ion and there is also an MPD molecule between them to deny access. When crystals are soaked in uranyl acetate or uranyl nitrate for long times or at high concentrations, they crack. This probably reflects competition of uranium ions with the MPD to disrupt the crystal packing and so crack the crystal.

In the crystal, the outer edge of each V1 interacts with two different symmetry mates. One symmetry mate contacts the lower portion of V1 and part of the molecular base, while the second contacts the upper part of V1 and the tower region (Table 7). These two symmetry mates are antiparallel to one another and nearly perpendicular to the central molecule (Fig. 6). The tetrahedral arrangement leaves V2 completely exposed to solvent, along with the top and bottom of the molecule.

3.6. Flexibility in the P3 trimer and its role in membrane interactions

The P3 molecule shows a degree of flexibility that may be important in its interactions with the membrane. Overall, the subunits of the P3 trimer are very similar. The average root-mean-square deviation when they are superimposed is 0.37 Å over all C$^{\alpha}$ positions, with subunits II and III agreeing better (0.30 Å) than I and II (0.38 Å) or I and III (0.43 Å). The largest variations in the subunit structures occur at NT-α, the C-terminus and in some of the loop areas (GH1, I1B2, EF2, and GH2) that are solvent exposed (Fig. 7*a*). During the early stages of refinement when strict NCS was imposed, the threefold-averaged density in these areas was not well defined and difficult to interpret. Later, the removal of the NCS restraints from the model in these difficult areas did not affect the residual *R* values, but the corresponding density became better resolved. Although residues could then be modeled, their thermal mobility constants remained higher than those in the rest of the molecule. Another sign that these regions have unusual flexibility is that they contain all the residues that required truncation to alanine in the final model because of poor side-chain density. The GH2 loop posed a particular challenge in refinement. Simulated-annealing (SA) omit maps

Table 7
Native and Se-Met P3 crystal contacts.

Contacts are defined as residue pairs with main-chain or side-chain atoms within 3.5 Å. The residue's superscript identifies its subunit. Residues displayed in square brackets are found only for native P3 and in parentheses only for Se-Met P3; the rest are common to both.

Subunit I.

Tower and upper part of V1		Molecular base and lower part of V1	
Symmetry†: $\frac{1}{2} - x, -y, \frac{1}{2} + z$		Symmetry: $-\frac{1}{2} + x, \frac{1}{2} - y, -z$	
Residue	Symmetric residues contacted	Residue	Symmetric residues contacted
[Arg52^{I}]	[Asp48^{II}], [Asn51^{II}]	Asp48^{I}	Asp56^{III}, Lys193^{III}
Arg98^{I}	Asp222^{II}	Asn51^{I}	Arg52^{III}, Val54^{III}
Tyr102^{I}	Asn79^{II}	Thr77^{I}	Tyr102^{III}, (Arg108^{III})
Arg108^{I}	Asn79^{II}, His80^{II}, Ala81^{II}, Gly157^{II}	Asn79^{I}	(Arg98^{III}), [Arg108^{III}], Glu111^{III}, (Glu261^{III})
Glu111^{I}	Asn79^{II}, (Ala81^{II})	(Gly157^{I})	(Arg108^{III})
Lys193^{I}	[Asn79^{II}], Glu224^{II}, [Glu226^{II}]	Glu161^{I}	Asn106^{III}
		Glu224^{I}	Arg98^{III}, Glu111^{III}
		Phe225^{I}	Lys193^{III}
		Glu226^{I}	Tyr102^{III}, Asn106^{III}, Arg191^{III}, Lys193^{III}
		[Glu227^{I}]	[Arg191^{III}]

Subunit II.

Tower and upper part of V1		Molecular base and lower part of V1	
Symmetry: $-x, \frac{1}{2} + y, \frac{1}{2} - z$		Symmetry: $\frac{1}{2} - x, -y, -\frac{1}{2} + z$	
Residue	Symmetric residues contacted	Residue	Symmetric residues contacted
Asn51^{II}	Asn51^{III}	[Asp48^{II}]	[Arg52^{I}]
Arg52^{II}	Asn51^{III}	[Asn51^{II}]	[Arg52^{I}]
Ser53^{II}	Asn51^{III}	Asn79^{II}	Tyr102^{I}, Arg108^{I}, Glu111^{I}, [Lys193^{I}]
Val54^{II}	Asp48^{III}, Asn51^{III}	His80^{II}	Arg108^{I}
Arg98^{II}	Glu224^{III}	Ala81^{II}	Arg108^{I}, (Glu111^{I})
Arg108^{II}	Asn79^{III}	Gly157^{II}	Arg108^{I}
Glu111^{II}	Asn79^{III}	Asp222^{II}	Arg98^{I}
Arg191^{II}	Glu226^{III}	Glu224^{II}	Lys193^{I}
Lys193^{II}	Glu226^{III}	[Glu226^{II}]	[Lys193^{I}]
Glu195^{II}	Glu224^{III}		

Subunit III.

Tower and upper part of V1		Molecular base and lower part of V1	
Symmetry: $\frac{1}{2} + x, \frac{1}{2} - y, -z$		Symmetry: $-x, -\frac{1}{2} + y, \frac{1}{2} - z$	
Residue	Symmetric residues contacted	Residue	Symmetric residues contacted
Arg52^{III}	Asn51^{I}	Asp48^{III}	Val54^{II}
Val54^{III}	Asn51^{I}	Asn51^{III}	Asn51^{II}, Arg52^{II}, Ser53^{II}, Val54^{II}
Asp56^{III}	Asp48^{I}	Asn79^{III}	Arg108^{II}, Glu111^{II}
Arg98^{III}	(Asn79^{I}), Glu224^{I}	Glu224^{III}	Arg98^{II}, Glu195^{II}
Tyr102^{III}	Thr77^{I}, Glu226^{I}	Glu226^{III}	Arg191^{II}, Lys193^{II}
Asn106^{III}	[Glu161^{I}], Glu226^{I}		
Arg108^{III}	(Thr77^{I}), [Asn79^{I}], (Gly157^{I})		
Glu111^{III}	Asn79^{I}, Glu224^{I}		
Arg191^{III}	Glu226^{I}, [Glu227^{I}]		
Lys193^{III}	Asp48^{I}, Phe225^{I}, Glu226^{I}		
(Glu261^{III})	(Asn79^{I})		

† Symmetry operator used to create symmetry-related molecule.

for part of this loop (352–357) in all three subunits showed areas of positive density. In subunit I, the electron density was sufficiently clear that two alternate paths of the loop could be traced (Figs. 7a and 7b). The main chain of the GH2 loop can be built equally well in either tracing, although individual side-chain atoms can be better defined in one or the other.

Much of the flexibility occurs in regions that are important for P3 interactions in the capsid. P3 has been fitted into an EM reconstruction of the PRD1 virion, showing that GH1, EF2, and GH2 interact with other P3 trimers when packed in the capsid (San Martín *et al.*, 2001). The contacts between the P3 trimer and others within a facet, or with P31 at the capsid vertices, would occur along the outer portions of the jelly rolls (Fig. 4b). Loops EF2 and GH2, which lie at the base of the V2 jelly roll (Fig. 2), are positioned where they would be involved in capsid contacts. In the capsid, these loops would be more constrained by interactions with other capsid proteins and the membrane and would then be likely to exhibit a more ordered conformation. A similar loop, GH1, runs along the lower edge of V1 and contains a sharp kink at Pro138. It is also solvent exposed and conformationally variable in the crystal structure, but is likely to engage in intermolecular contacts in the capsid.

The putative membrane-interacting regions of P3, NT-α and the I1B2 loop also exhibit structural variability in the three subunits. There is weak density with a helical shape preceding Gln11 at the N-terminus, suggesting that NT-α may contain one or more additional turns. As this density is disjointed and so it is difficult to identify C$^{\alpha}$ positions, it has not been modeled. The latest EM reconstructions of the virion and the *sus1* particle also suggest that the α-helix could continue for several of the remaining N-terminal residues (San Martín *et al.*, 2001). The path of the main chain of NT-α in the model is fairly distinct in the electron density, although its atoms have high thermal mobility constants (46.8 Å^2 on average compared with 25.8 Å^2 for the whole molecule), but its side-chain atoms are not well defined. NT-α extends ~18 Å below V1, but its direction relative to the rest of the molecule is slightly different in each of the three subunits. This small flexibility is provided by Ser35, which acts as a hinge. However, the side chain of Arg28 tethers NT-α to the bottom of V1 by making salt links with Asp176 and Asp253 (Fig. 3b). In the virion, the flexibility of NT-α and I1B2 could be utilized
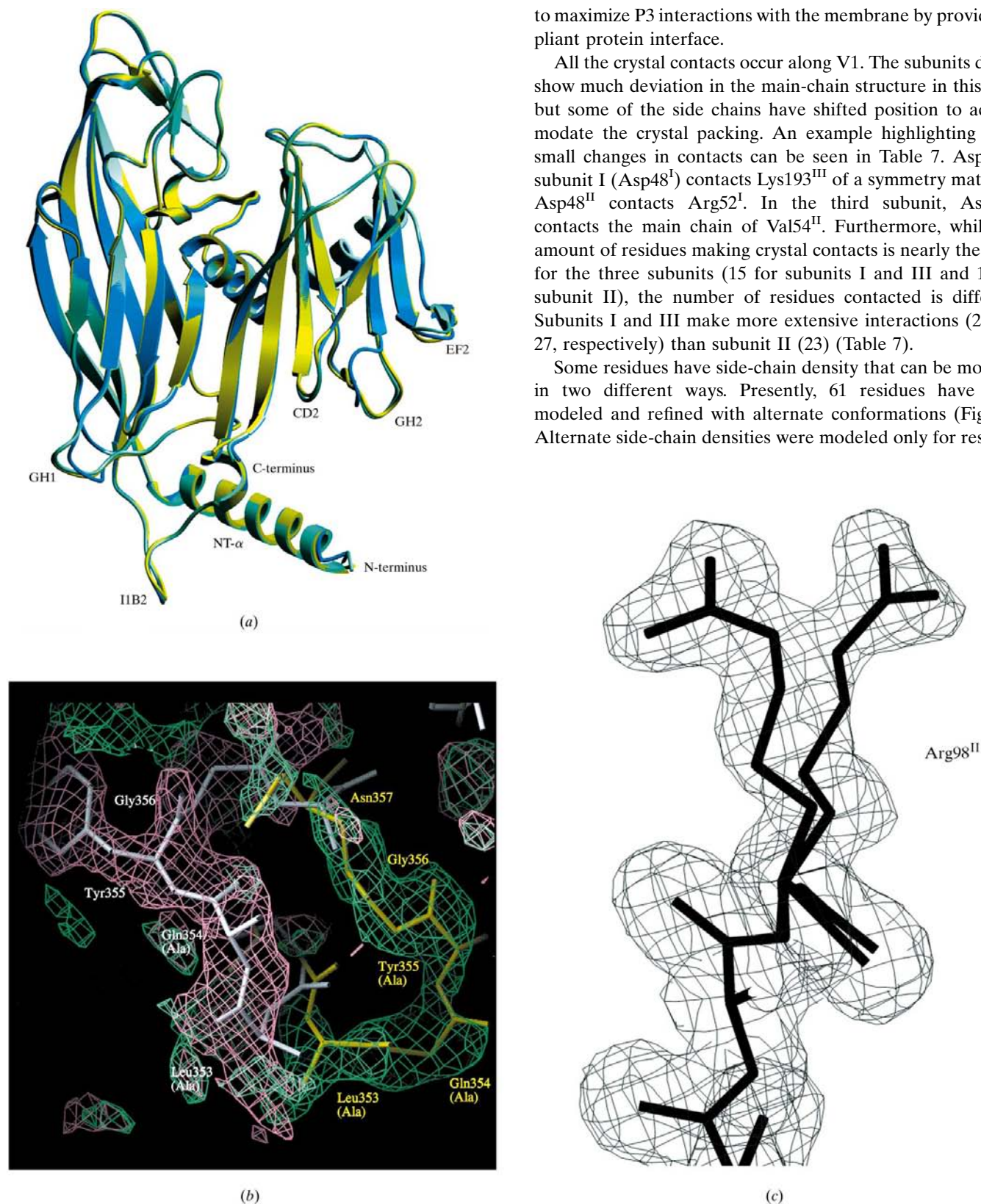
to maximize P3 interactions with the membrane by providing a pliant protein interface.

All the crystal contacts occur along V1. The subunits do not show much deviation in the main-chain structure in this area, but some of the side chains have shifted position to accommodate the crystal packing. An example highlighting these small changes in contacts can be seen in Table 7. Asp48 in subunit I (Asp48[I]) contacts Lys193[III] of a symmetry mate, but Asp48[II] contacts Arg52[I]. In the third subunit, Asp48[III] contacts the main chain of Val54[II]. Furthermore, while the amount of residues making crystal contacts is nearly the same for the three subunits (15 for subunits I and III and 18 for subunit II), the number of residues contacted is different. Subunits I and III make more extensive interactions (28 and 27, respectively) than subunit II (23) (Table 7).

Some residues have side-chain density that can be modeled in two different ways. Presently, 61 residues have been modeled and refined with alternate conformations (Fig. 7c). Alternate side-chain densities were modeled only for residues

Figure 7
Comparison of the three subunits of the P3 trimer. (a) Superposition of the three subunits (blue, green and yellow) showing their overall similarity with relatively small differences in some areas (labeled). A portion of the GH2 loop (352–358) in subunit I (blue) has two conformations. (b) $F_o - F_c$ electron-density maps showing 3σ positive density for the GH2 loop in subunit I, which has two possible conformations. The green density is generated after the model is refined when built into the pink density and the pink density when it is built in the green. The figure was produced with *O* (Jones *et al.*, 1991). (c) The alternative conformations of Arg98[II]. (a) and (c) were produced with *SETOR* (Evans, 1993).

that displayed supporting electron density above 1σ in $2F_o - F_c$ maps and 3σ in $F_o - F_c$ maps. These residues are not the same for each subunit and no general pattern was observed in the variability.

3.7. The pore region at the molecular threefold

The three FG1 loops not only overlap their neighboring subunits and contribute to molecular stability, but also create a cap over the internal cavity between the jelly rolls (Fig. 4b). At the molecular threefold, the three loops form a pore that is plugged by a pair of hydrophobic residues from each of the subunits, three valine residues (Val134) at the top and three methionine residues (Met133) at the bottom (Fig. 5b). Notably, a water molecule lies at the center of this pore, where it forms hydrogen bonds, as it is 2.8 Å from the main-chain O atoms of all three Val134 residues. As the pore is small, this water molecule completely fills its central portion. It is possible that the residues at the pore could shift sufficiently to permit the pore to open, thus allowing exchange of the central water or the transport of ions or small molecules across the protein capsid.

In the virion, an ion entering the pore would still have to contend with the membrane or a membrane-associated protein before it could enter the interior. The P3 trimer has a slight negative charge on the top of the molecule from the residues in the FG1 loop (Fig. 8). This would tend to repulse negatively charged ions and molecules, but to attract those positively charged. If a molecule did cross the hydrophobic plug, it would then encounter a cavity with a positive charge (Figs. 5c and 5d). This charge distribution would tend to prevent charged ions and molecules from crossing the pore, but could allow free transport of small neutral molecules such as water. In human rhinoviruses, cations (possibly a calcium or zinc) at the fivefold axes (Zhao et al., 1997) are thought to aid in stability and the water in P3's pore may also play this role. As the pore is only large enough to accommodate a water molecule, it is too narrow for the transport of nucleic acid, which appears to be inserted and removed through the icosahedral fivefold (Bamford et al., 1995; Grahn et al., 1999; Rydman et al., 1999). Notably, the potassium tetra-chloroplatinate (II) derivative shows a platinum ion bound to the sulfurs of the three Met133 residues forming the lower part of the hydrophobic plug in the pore. This demonstrates that ions are attracted to this area, but access in the crystal is through the basal cavity rather than by passage through the pore.

3.8. Contribution of solvent atoms, ions and small molecules to molecular stability

The importance of bound water in protein structure is becoming more widely appreciated. Water molecules are essential for the folding of proteins and allowing them to enter solution (Fye et al., 1998) and water-mediated hydrogen bonds add stability to protein–protein interactions (Janin, 1999). A large number of water molecules (782) are interspersed in and around the P3 molecule (Figs. 9a and 9b). Water molecules were only modeled in density where they formed hydrogen bonds with protein atoms, other waters, or both. A high percentage (35%) is at the top of the molecule between the three towers, where a puddle of water forms extensive hydrogen-bonding networks that extend three to four solvent layers from the protein. The puddle is a feature of the trimer and would not be present in the monomer. The water network covers ~ 9900 Å^2 (or $\sim 24\%$ of the total accessible surface area) and is broken only at the hydrophobic plug formed by the three Val134 residues in the pore (Fig. 8). It seems highly likely that this network plays an important role in stabilizing the trimer by enhancing hydrogen bonding for both the intra- and intersubunit contacts. The water interactions also reinforce the thin FG1 loop by building on the protein framework to produce a thicker cap over the P3 cylinder.

Although the first solvent layer around proteins is the best ordered and is necessary for protein solubility, ordered bulk solvent can extend to as many as three layers (Badger & Caspar, 1991). However, computer simulations and NMR experiments suggest that only a few waters near proteins will have a low exchange rate, while most will be nearly as mobile as those in bulk solvent, so that there is a rapid turnover of the individual waters in the layers (Karplus & Faerman, 1994; Levitt & Park, 1993). In X-ray structures, waters in the outer layers are usually not visualized because of mobility, but some protein-surface features, such as crystal contact areas and small crevices, will lead a few waters to have the preferred positional minima that are revealed as electron-density peaks
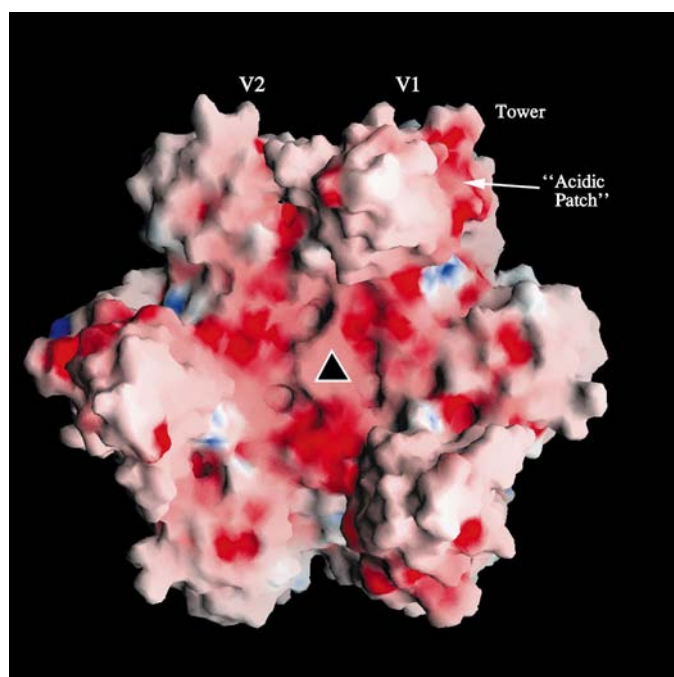

Figure 8
Electric potential for the upper surface of P3, with blue positive charge and red negative charge. The acidic patch along the outer edge of the tower is labeled (see also Fig. 10a). The hydrophobic plug formed by a Val134 triplet in close contact seals the pore at the molecular threefold (black triangle). The figure was produced with GRASP (Nicholls et al., 1991).

(Burling *et al.*, 1996). The clearest example is seen in crambin (Teeter, 1991), where distinct pentagonal ordering of the water molecules is maintained by certain anchoring positions on the protein. In P3, the most extensive ordering of solvent is not at the crystal contacts or in small crevices, but occurs at the top of the molecule between the towers. The cup shape and polar acidic nature of this region (Figs. 5*c*, 5*d* and 8) may help maintain the hydrogen-bonding network by reducing water fluctuation and turnover. It would be interesting to examine P3 with NMR to investigate if the water-exchange rate of the puddle is lower than normal and to perform molecular-dynamic simulations to see if waters would congregate to this area.

The extensive water network at the top of P3 is a striking feature that suggests a functional role in trimer stabilization. The importance of water-mediated hydrogen bonding for protein–protein stability is well established (Janin, 1999) and extending this concept to water networks seems logical, but the idea is controversial (Levitt & Park, 1993). These authors and others claim that as the ordered waters are still in constant exchange with bulk solvent or have increased vibrational energy when extensively hydrogen bonded, the free energy of the system is not significantly affected (Creighton, 1993; Levitt & Park, 1993). Resolving this question for P3 would be difficult, but one possible approach would be to determine its thermal melting curve in the presence of different detergents or other solute particles to disrupt the water network (Parsegian *et al.*, 1995). The difficulty in such experiments is to distinguish the effect on water from that on protein.

A further feature of some 'waters' is that they display an unusual octahedral coordination (Fig. 9*c*). These are usually associated with glutamate or aspartate residues and are all in the puddle of the tower region. Six such molecules were modeled as sodium ions (Na^+; Nayal & Di Cera, 1996). Two abnormally large volumes of density that are associated by crystallographic symmetry with Arg28 of all three subunits, originally modeled as water molecules with very low thermal mobility constants, were later assigned as chloride ions (Cl^-). These ions may further stabilize the ordered solvent. Sodium chloride is a component of the protein and crystallization buffers and the latter also contains sodium acetate.

The high-resolution structure also reveals the role of MPD in crystal formation. Some areas of density have shapes in good agreement with MPD molecules (Fig. 9*c*), which was the precipitant used in crystallization, and 18 molecules have been modeled. Most are in the crystal contact areas and are oriented to mediate contacts between hydrophobic residues on one trimer and hydrophilic residues on its neighbor. Their positions show how MPD aids in crystal formation and stability.

3.9. Comparison of native P3 with Se-Met P3

Careful analysis of the diffraction data has cast some light on the effects of radiation damage. After MAD phasing, a molecular model of the Se-Met P3 structure was built and refined to 2.2 Å resolution using the programs *X-PLOR*

(Brunger, 1992) and *CNS* (Brunger *et al.*, 1998) and the MAD data at the remote wavelength L3. The three subunits of Se-Met P3 (I, II and III) have contiguous density over residues 16–384 for I, 15–384 for II and 14–385 for III. The residues whose side-chain densities are not well defined have been modeled as alanine. These 34 residues fall in the same regions of high mobility observed in the native structure. The 437 water molecules included in the Se-Met P3 structure correspond to a subset of the 782 waters that are observed in the higher resolution 1.65 Å native P3 model.

Although an Se-Met P3 model produced midway through its refinement was used to phase the native data, the subsequent refinement of the two models was performed independently. The aim was to investigate whether any structural alterations occur in the P3 molecule upon substituting selenium for sulfur in the five methionine residues of the subunit. The refinements showed that the Se-Met molecule has a 1° rotation about its molecular threefold axis relative to the native structure. This leads to a difference of ~0.7 Å in atom positions at the edge of the ~80 Å diameter P3 trimer. Despite this rotation, the native P3 and the Se-Met P3 models are very similar, with a root-mean-square deviation of 0.36 Å on all atoms of the trimer. These results were surprising, as very little difference in structure had been expected and Se-Met and native crystals had similar diffraction quality. Sulfur and selenium have similar chemical characteristics and their radii are not that different: 1.05 Å for sulfur and 1.20 Å for selenium. The detailed alterations in the structure leading to the overall movement were less easy to tease out. In the immediate vicinity of the methionine residues, the only apparent difference is slightly higher electron density corresponding to the additional 18 electrons that selenium contains compared to sulfur. Additionally, the Met133 residues surrounding the pore region in Se-Met P3 only appear to have a single preferred orientation, presumably owing to their slightly larger size.

The Se-Met P3 crystal contacts are slightly different from those in the native P3 structure (Table 7). Almost the same number of residues produced contacts in each subunit as in the native structure, but the burden shifts slightly to subunit III in Se-Met P3. Subunit I has 13 interacting residues (compared with 15 in native P3), subunit II has 16 (18) and subunit III has 16 (15). The total number of residues involved drops from 78 (or 48 unique residues) in the native structure to 74 (45) for Se-Met P3. The particular interactions that the residues make also change.

It is possible that the differences in the Se-Met and native structures arose from the freezing step. However, identical techniques and solvents were used and the crystals appeared to behave similarly in all ways. As a test of variability between crystals, the refined native and Se-Met P3 models were subjected to conjugate-gradient minimization against data collected on the laboratory source from various native and Se-Met P3 crystals. None of the resultant models showed a rotation. A more likely possibility is damage to the crystal from synchrotron radiation. When various protein crystals at cryotemperatures (hen egg-white lysozyme, winged-bean

chymotrypsin inhibitor and *Torpedo californica* acetyl-cholinesterase) were examined, all showed molecular rotations upon long exposure (Ravelli & McSweeney, 2000; Weik *et al.*, 2000). It is notable that the Se-Met P3 crystal used in the MAD experiment was mounted and exposed to synchrotron X-radiation but, owing to technical problems at the beamline, the experiment was halted and then restarted more than 10 h later. Any rotation arising from radiation damage should be enhanced by this delay and models minimized against all three Se-Met MAD data sets showed the same 1° rotation from native, as if this might be the maximum rotation allowed. To investigate the radiation-damage hypothesis, the refined native and Se-Met P3 models were also minimized against three incomplete data sets for another Se-Met P3 crystal from the same beamline (data not shown). The P3 molecule steps from the 'native' position found for the first data set towards the 'Se-Met' position found for the last. This is compelling evidence that the observed rotation was a consequence of radiation damage, but the structural mechanism for the change remains a mystery. Such effects may be enhanced by the absorption of energy by the Se atoms at their resonance frequency during a MAD experiment.

3.10. Major coat proteins in other PRD1-related viruses

PRD1-like viruses have been independently isolated from different parts of the world. At present, six are known: the prototypic PRD1 and the highly related PR3, PR4, PR5, L17 and PR772 (Bamford *et al.*, 1981). Among these, there are only very minor differences in the gene sequences for the major coat proteins (A.-M. Saren, JKHB, DHB and L. Paulin, work yet to be published). To highlight this similarity, there are only 12 differences in the nucleotide sequences encoding the major capsid proteins for PRD1 and PR4 and only one alters the amino-acid sequence: Thr77 in PRD1 becomes Lys77 in PR4 (T77K) (Myung *et al.*, 1994). The other sequence alterations are two for PR3 (T77K and V362L), two for PR5 (T77K and E224K), one for L17 (T77K) and one for PR772 (T77K) (Fig. 10).

PRD1-P3 is the only one of the coat proteins to have a threonine instead of a lysine at position 77. This residue lies in the tower region at the end of D1. The electron density for P3 confirms the assignment of a threonine; a lysine could not easily be accommodated in the crystal because of steric hindrance with a symmetry-related molecule. The tower has a patch of negative charge along its outer edge at the base of the tower, just above the BC1 loop (Figs. 2, 8 and 10). Thr77 occurs just above this negative patch. If this was a lysine, its long side chain could interact with one of the neighboring glutamate or aspartate residues and neutralize some of the charge in the negative patch. A similar alteration is found in bacteriophage PR5 (E224K) involving Glu224, which is located at the start of I1 and is part of the acidic patch. With positive lysines at residues 77 and 224, the PR5 tower would be much less acidic than that of PRD1. These changes might have arisen to allow bacteriophages to better cope with individual environments with different pH values.

The second change in PR3 (V362L) is more complex. Val362 is on the H2 strand and lies at the bottom of a pocket surrounded by hydrophilic residues (Asp273, Tyr305, Arg309, Asp315 and Lys318). It also pairs with Val275 in strand C2. If Val362 were altered to the more bulky leucine, the residues surrounding the pocket would be pushed out to accommodate the change. This modification most likely would not have a noticeable effect on the structure and would only cause the side chains aligning the pocket to adjust, but it could also cause some minor changes to strands H2 and C2.

The few differences between the major coat proteins of the PRD1-like viruses seem to be easily accommodated by the structure. It is therefore surprising that these viruses, which have been isolated from different locations all over the world, should maintain such highly conserved major coat proteins. Myung *et al.* (1994) examined amber mutants in the major coat protein from a nitrosoguanidine-mutagenized (*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine) stock of bacteriophage PR4. They noted that the low number of mutations isolated and their pattern suggested that the major coat protein P2 (P3 in PRD1) might have a low tolerance for amino-acid substitutions. Alternatively, this intolerance might arise from the requirements of the capsid-assembly pathway. The bacterial GroEL/ES chaperonins are involved in P3 trimerization (Hänninen *et al.*, 1997). The trimers then must interact with two assembly-factor proteins expressed by the bacteriophage, the membrane-bound P10 and the soluble P17 (as named in PRD1), for proper capsid formation (Caldentey *et al.*, 1999). Altogether, P3 must have fairly stringent structural requirements to satisfy its various roles in protein synthesis, chaperonin-mediated trimerization, membrane association, viral assembly and function as a protective coat.

3.11. Relationship of P3 to adenovirus hexon

The structural and evolutionary relationship of P3 (394 residues) to adenovirus type 2 hexon (967 residues, PDB code 1dhx; Athappilly *et al.*, 1994) has been reported (Benson *et al.*, 1999). P3 and hexon both contain two jelly rolls with their β-strands parallel to the molecular threefold, which in turn is normal to the viral surface. In both proteins, an α-helix follows the F-strand in each jelly roll to produce a helix–jelly roll–helix–jelly roll motif. The disposition of the tower loops with respect to the β-strands in P3 and hexon are also very similar, but those in hexon are much longer. Although its jelly rolls have the same width as those in P3, hexon forms a molecule with a broader base. This is because of the presence of an additional three-stranded β-sheet, called the pedestal connector (PC) domain, and the L_3 loop between its two jelly rolls (Benson *et al.*, 1999). This expansion of the hexon base allows adenovirus to form a capsid with a vertex–vertex diameter of 920 Å (Stewart *et al.*, 1991), compared with the much smaller value of 698 Å in bacteriophage PRD1 (San Martín *et al.*, 2001).

Recently, Rux & Burnett (2000) solved the structure of the hexon for type 5 adenovirus (951 residues; PDB code 1rux). As type 5 hexon has 86% amino-acid identity to type 2 hexon

and they crystallize isomorphously, the type 2 hexon structure was used as the starting type 5 model. Surprisingly, the refinement was not straightforward and the final model for type 5 hexon contained several significant revisions to the starting model. These included a reverse tracing of a 65 amino-acid loop, a 41 amino-acid shift in the sequence registry from the N-terminus to residue 130 and small registry shifts in other portions of the model. These changes produced a type 5 structure with much lower residuals than for the type 2 hexon structure, suggesting that its model agrees much better with the experimental data. In recent work, the type 2 hexon structure has been revisited to produce a revised model exhibiting the same alterations (J. J. Rux and RMB, personal communication). Although the changes in the sequence registry of hexon did not change its overall architecture, they revealed that the flexible N-terminus is not as long as thought previously (Athappilly *et al.*, 1994). The revised hexon model has two new features relevant to the P3–hexon comparison. It contains a more pronounced α-helix at the N-terminus, which matches the position of NT-α in P3. In addition, an acidic stretch that was partially buried in the original type 2 hexon model has been relocated to a loop in the tower. Therefore, the towers in hexon would be negatively charged as in P3. The correspondence between P3 and hexon is thus even greater than was first apparent.

4. Conclusions

The X-ray crystal structure of PRD1 coat protein P3 reveals a second and more primitive version of the 'double-barrel' viral motif first observed in adenovirus type 2 hexon. A recent
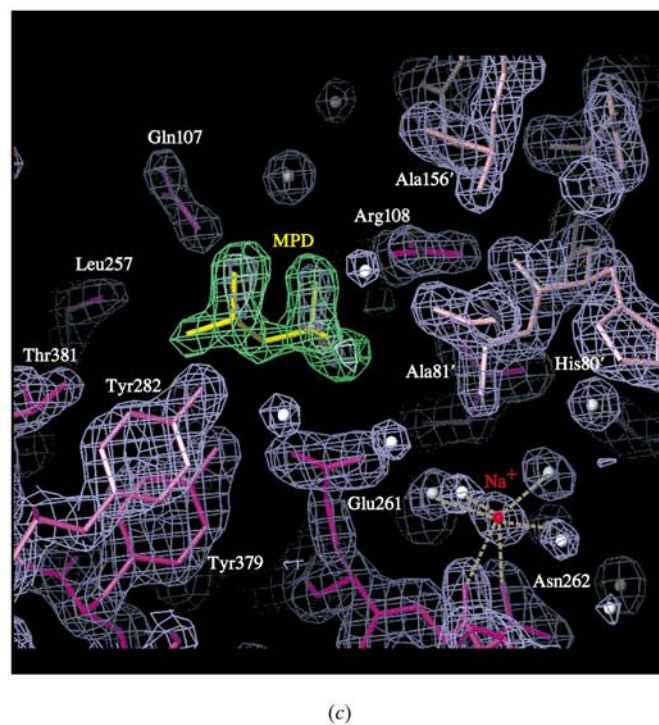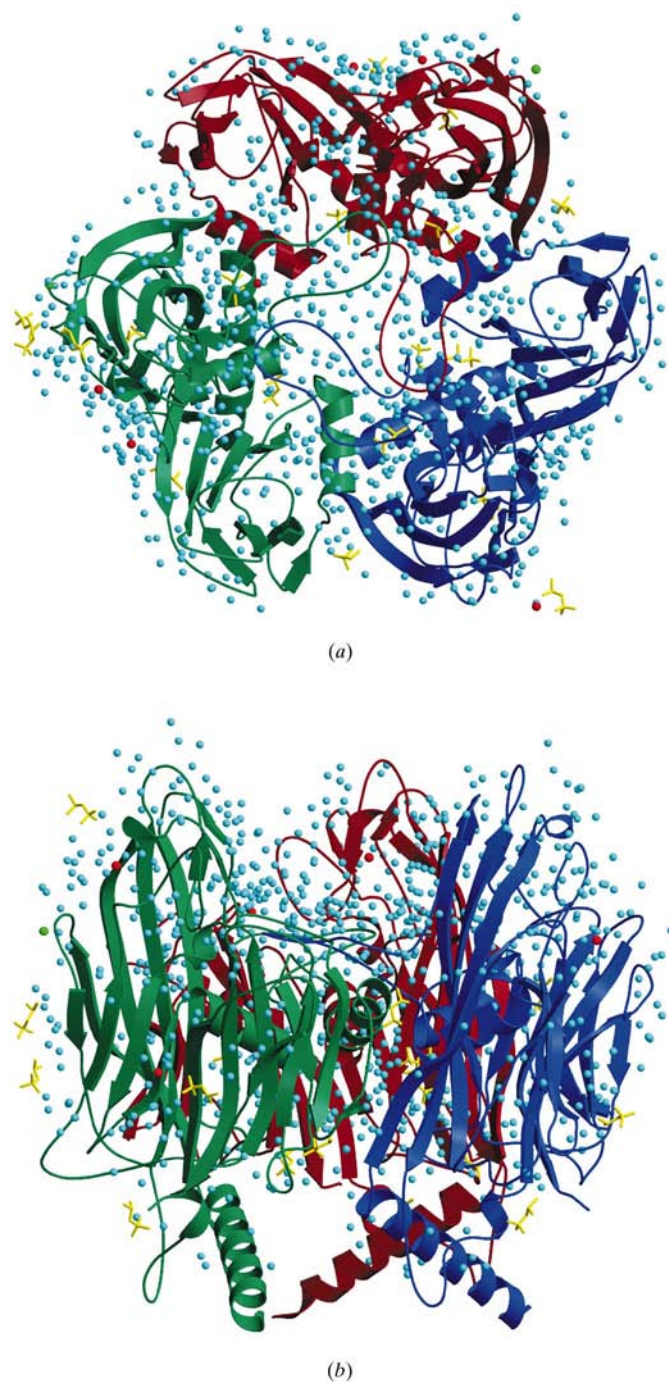

Figure 9
Solvent atoms. (*a*) A view of the P3 trimer, as in Fig. 4(*b*), that shows water molecules (light blue), MPD molecules (yellow), sodium ions (red) and chloride ions (green). Water molecules surround the tower region and flood across the top of the molecule, aiding in hydrogen bonding and adding to the stability of the molecule. (*b*) Colored as in (*a*) but with P3 oriented as in Fig. 4(*a*). (*c*) An MPD molecule (yellow) placed within positive 2σ density from an $F_o - F_c$ map (green). This MPD is in the interface between symmetry-related P3 trimers (magenta and pink). A sodium ion (red) is also visible with protein and water molecules surrounding it with octahedral coordination (marked with dashed lines for clarity). The P3 molecule is shown within a 1.7 Å resolution $2F_o - F_c$ electron-density map contoured at 2σ. (*a*) and (*b*) were produced with *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Bacon, 1997) and (*c*) was produced with *O* (Jones *et al.*, 1991).

structure determination for type 5 hexon (Rux & Burnett, 2000) shows that this similarity is now even closer than was first noted (Benson *et al.*, 1999). With two examples of the motif, it is possible to make some observations about how and why the single viral jelly roll found in structurally simpler virions such as plant viruses has evolved into a very stable trimeric coat protein with six barrels. It seems clear that the P3 double barrel arose by gene duplication. This event must have been relatively common, as examples are found in the human picornoviruses as well as the plant comoviruses (Rossmann & Johnson, 1989) and so it is easy to see how two independent jelly rolls can become integrated into a single subunit (see Lynch & Conery, 2000). Such duplication would considerably reduce the number of separate 'building blocks' and so significantly enhance the probability for their accurate assembly into a viral capsid (Burnett, 1985). Trimerization would not only reduce the number of building blocks still further, but would provide a pseudo-hexagonal shape conducive to forming closely packed coat-protein arrays and interdigitating facets. These features, which characterize adenovirus-like architecture, permit the assembly of large capsids with only a very small number of well defined interactions.

An essential characteristic of such a building block is that it be very stable. In P3, the stability arises from the overlap of the long trimerization loop with neighboring subunits. This arrangement is simpler than in hexon, where three loops interact to form each tower, but illustrates the underlying principle of stability gained by tethering subunits to each other. The complexity obviously comes at a metabolic price, as host-cell chaperonins are required to fold and trimerize P3 and adenovirus has evolved its specialized 100 K protein for this task. The interaction between loops seems to be the primary force driving trimerization in P3, as subunit–subunit interactions themselves are not critical. C-terminal truncation experiments show that trimers still form when most of V2 is removed (Bamford *et al.*, 1993; JKHB, unpublished results). An interesting feature in P3, which may also contribute to stabilization, is the pool of ordered water between the towers. It is possible that this is also present in hexon, but that it was not observed owing to the lower resolution of its structure determination.

The shape of a double-barrel trimeric protein is ideal for forming a closely packed array. This would tend to be planar or modestly curved and so would suit the construction of virions with relatively large facets. As evolution led to greater complexity, growing genomes would require correspondingly larger virions. These could be constructed by increasing the size of the coat proteins, while maintaining the pseudo-triangulation number at 25 as seen in PRD1 and adenovirus, where 240 trimers of P3 or hexon form the 20 facets. In PRD1, the daisy shape of P3 allows its petal-like jelly rolls to interdigitate. In the larger hexon, insertions between the barrels lead to a more hexagonal shape and less interdigitation. The result is an expansion of the facet and a corresponding increase in the virion's internal volume. Even larger volumes must require an increase in the triangulation number, as

polypeptides are limited in size. It is notable that the 1850 Å diameter Chilo iridescent virus ($T = 147$) and the 1900 Å *Paramecium bursaria chlorella* virus ($T = 169$) (Yan *et al.*, 2000) also have trimeric pseudo-hexagonal coat proteins and it seems likely that these may also contain a double-barrel motif.

We envisage that there is a PRD1–adenovirus lineage originating from an early ancestor. The PRD1 coat protein appears to have evolved to fulfill its functions extremely effectively, but to have reached a point where its sequence and structure are optimal. It is possible that P3 reveals a minimal double-barrel structure, as the molecule is tightly constrained by its various roles as trimerization substrate, coat protein and membrane anchor. Even mutations in its tower may change its charge distribution or shape and so interfere with its interactions with GroEL/ES upon trimerization or with the assembly factors P10 and P17 in virion formation. This invariability in the coat does not seem to preclude adaptation and lead to extinction, as PRD1 is quite promiscuous and infects a wide variety of Gram-negative bacteria (Olsen *et al.*, 1974).

The structural role of P3 in the PRD1 capsid is being investigated with complementary EM studies. The P3 X-ray structure has been positioned by rigid-body refinement methods within new high-resolution cryo-EM image reconstructions of the native PRD1 virion, a particle lacking DNA and the P3 shell without a membrane (San Martín *et al.*, 2001).
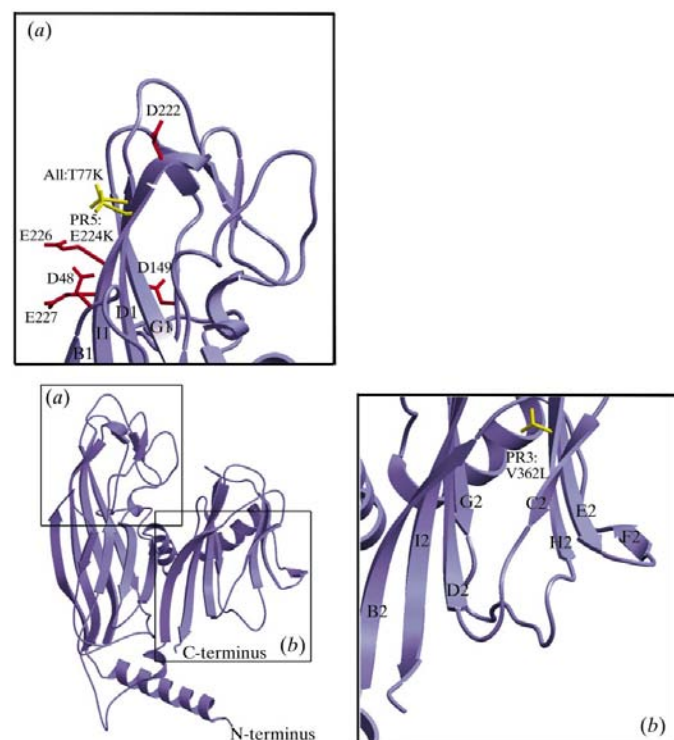

Figure 10
The acidic patch of P3 and the differences between the PRD1 family members. (*a*) The location of Thr77 and Glu224 (yellow) and the residues that make up the acidic patch (red). (*b*) Val362, the other altered residue in the PRD1 family. The figure was produced with *SETOR* (Evans, 1993) and *MOLSCRIPT* (Kraulis, 1991).

research papers

This work, which will explore the organization and interactions of P3, is analogous to the earlier fitting of hexon into adenovirus EM density (Stewart, Fuller *et al.*, 1993). It defines the regions of P3 that interact in the PRD1 capsid and will allow difference imaging, as was performed for adenovirus, to visualize these interactions. An example is the ten C-terminal residue stretch that is disordered in the crystal structure, where the corresponding EM density should reveal its role in capsid formation or stability. The EM work is also probing P3–membrane interactions and will reveal any minor capsid proteins that may be present. The EM fitting confirms the suggestion from the P3 structure that NT-α and I1B2 contact the membrane and reveals that this interaction is present in the wild-type virion where the membrane closely follows the inside of the capsid. However, this interaction must be flexible as the membrane is not flush against the facet in *sus1*, where DNA is absent, and the only P3 contacts are then with the part of the N-terminus that is disordered in the crystal structure. The EM work is being extended by refining the P3 positions by molecular-dynamics methods to determine the exact chemistry of the P3 capsid interactions (C. San Martín, RMB & C. L. Brooks III, personal communication).

As PRD1 virions can be obtained in large quantities (Bamford & Bamford, 1991) with high purity (Walin *et al.*, 1994), PRD1 may be useful as a foreign epitope display system. The obvious candidate for alteration is P3, but the very low sequence variability in the major coat proteins of the PRD1-family, together with data from recombinant P3 studies, suggest that it is intolerant to change. Although foreign epitopes have been successfully inserted into the P3 tower region, most alterations cause a severe loss in the ability of the recombinant protein to complement the mutation in the virus, so attention has turned to the vertex (J. T. Huiskonen, L. Laakkonen, DHB and JKHB, personal communication). These findings contrast with hexon's wide sequence variation in the different adenovirus serotypes, particularly in the tower region, and its tolerance to epitope insertion (Crompton *et al.*, 1994).

The stage is also set to explore the various roles of P3 by mutating residues in potential interaction regions. We can probe the residues involved in trimer formation and define those interacting with chaperonins and assembly factors. Of particular interest are residues that may be involved in membrane interactions and translocation, such as the glutamine and asparagine residues in NT-α and I1B2. We will characterize the resultant virions using methods similar to our recent EM studies on PRD1 (San Martín *et al.*, 2001). The known P3 structure and the PRD1 system make a powerful tool for the exploration of viral assembly processes and protein–membrane interactions.

References

Aleström, P., Stenlund, A., Li, P. & Pettersson, U. (1982). *Gene*, **18**, 193–197.

Athappilly, F. K., Murali, R., Rux, J. J., Cai, Z. & Burnett, R. M. (1994). *J. Mol. Biol.* **242**, 430–455.

Badger, J. & Caspar, D. L. D. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 622–626.

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

Bamford, D. H. & Mindich, L. (1982). *J. Virol.* **44**, 1031–1038.

Bamford, D. H. & Mindich, L. (1984). *J. Virol.* **50**, 309–315.

Bamford, D. H., Rouhiainen, L., Takkinen, K. & Söderlund, H. (1981). *J. Gen. Virol.* **57**, 365–373.

Bamford, J. K. H. & Bamford, D. H. (1990). *Virology*, **177**, 445–451.

Bamford, J. K. H. & Bamford, D. H. (1991). *Virology*, **181**, 348–352.

Bamford, J. K. H., Luo, C., Juuti, J. T., Olkkonen, V. M. & Bamford, D. H. (1993). *Virology*, **197**, 652–658.

Barrow, P. A. & Soothill, J. S. (1997). *Trends Microbiol.* **5**, 268–271.

Belnap, D. M. & Steven, A. C. (2000). *Trends Microbiol.* **8**, 91–93.

Benson, S. D., Bamford, J. K. H., Bamford, D. H. & Burnett, R. M. (1999). *Cell*, **98**, 825–833.

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

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Burling, T. F., Weis, W. I., Flaherty, K. M. & Brünger, A. T. (1996). *Science*, **271**, 72–77.

Burnett, R. M. (1985). *J. Mol. Biol.* **185**, 125–143.

Burnett, R. M., Grütter, M. G. & White, J. L. (1985). *J. Mol. Biol.* **185**, 105–123.

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

Caldentey, J., Blanco, L., Savilahti, H., Bamford, D. H. & Salas, M. (1992). *Nucleic Acids Res.* **20**, 3971–3976.

Caldentey, J., Hänninen, A.-L., Holopainen, J. M., Bamford, J. K. H., Kinnunen, P. K. J. & Bamford, D. H. (1999). *Eur. J. Biochem.* **260**, 549–558.

Caldentey, J., Luo, C. & Bamford, D. H. (1993). *Virology*, **194**, 557–563.

Chelvanayagam, G., Heringa, J. & Argos, P. (1992). *J. Mol. Biol.* **228**, 220–242.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowtan, K. D. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

Creighton, T. E. (1993). *Proteins: Structures and Molecular Properties.* New York: W. H. Freeman & Co.

Crompton, J., Toogood, C. I. A., Wallis, N. & Hay, R. T. (1994). *J. Gen. Virol.* **75**, 133–139.

Cudney, B., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.

Davis, T. N. & Cronan, J. E. Jr (1985). *J. Biol. Chem.* **260**, 663–671.

Davis, T. N., Muller, E. D. & Cronan, J. E. Jr (1982). *Virology*, **120**, 287–306.

Desiderio, S. V. & Kelly, T. J. Jr (1981). *J. Mol. Biol.* **145**, 319–337.

Evans, S. V. (1993). *J. Mol. Graph.* **11**, 134–138.

Furey, W. & Swaminathan, S. (1997). *Methods Enzymol.* **277**, 590–620.

Fye, J. L., Woenckhaus, J. & Jarrold, M. F. (1998). *J. Am. Chem. Soc.* **120**, 1327–1328.

Grahn, A. M., Caldentey, J., Bamford, J. K. H. & Bamford, D. H. (1999). *J. Bacteriol.* **181**, 6689–6696.

Hänninen, A.-L., Bamford, D. H. & Bamford, J. K. H. (1997). *Virology*, **227**, 207–210.

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

Hendrix, R. W. (1999). *Curr. Biol.* **9**, R914–R917.

Janin, J. (1999). *Structure*, **7**, R277–R279.

Jiang, J.-S. & Brünger, A. T. (1994). *J. Mol. Biol.* **243**, 100–115.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.

Karplus, P. A. & Faerman, C. (1994). *Curr. Opin. Struct. Biol.* **4**, 770–776.

Kelly, T. J. Jr (1984). *The Adenoviruses*, edited by H. S. Ginsberg, pp. 271–308. New York: Plenum Press.

King, J. & Chiu, W. (1997). *Structural Biology of Viruses*, edited by W. Chiu, R. M. Burnett & R. L. Garcea, pp. 288–311. New York: Oxford University Press.

Kotilainen, M. M., Grahn, A. M., Bamford, J. K. H. & Bamford, D. H. (1993). *J. Bacteriol.* **175**, 3089–3095.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Levitt, M. & Park, B. H. (1993). *Structure*, **1**, 223–226.

Lo Conte, L., Chothia, C. & Janin, J. (1999). *J. Mol. Biol.* **285**, 2177–2198.

Luo, C., Butcher, S. & Bamford, D. H. (1993). *Virology*, **194**, 564–569.

Lynch, M. & Conery, J. S. (2000). *Science*, **290**, 1151–1155.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

McKenna, R., Xia, D., Willingmann, P., Ilag, L. L., Krishnaswamy, S., Rossmann, M. G., Olson, N. H., Baker, T. S. & Incardona, N. L. (1992). *Nature (London)*, **355**, 137–143.

Merril, C. R., Biswas, B., Carlton, R., Jensen, N. C., Creed, G. J., Zullo, S. & Adhya, S. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 3188–3192.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Mindich, L., Bamford, D. H., Goldthwaite, C., Laverty, M. & Mackenzie, G. (1982). *J. Virol.* **44**, 1013–1020.

Mindich, L., Bamford, D. H., McGraw, T. & Mackenzie, G. (1982). *J. Virol.* **44**, 1021–1030.

Myung, H., Vanden Boom, T. & Cronan, J. E. Jr (1994). *Virology*, **198**, 17–24.

Nayal, M. & Di Cera, E. (1996). *J. Mol. Biol.* **256**, 228–234.

Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.

Olsen, R. H., Siak, J.-S. & Gray, R. H. (1974). *J. Virol.* **14**, 689–699.

Oostrum, J. van, Smith, P. R., Mohraz, M. & Burnett, R. M. (1987). *J. Mol. Biol.* **198**, 73–89.

Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, pp. 80–86. Warrington: Daresbury Laboratory.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**52**, 659–668.

Parsegian, V. A., Rand, R. P. & Rau, D. C. (1995). *Methods Enzymol.* **259**, 43–94.

Ramakrishnan, V. & Biou, V. (1997). *Methods Enzymol.* **276**, 538–557.

Ravelli, R. B. G. & McSweeney, S. M. (2000). *Structure*, **8**, 315–328.

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

Rux, J. J. & Burnett, R. M. (2000). *Mol. Ther.* **1**, 18–30.

Rydman, P. S., Caldentey, J., Butcher, S. J., Fuller, S. D., Rutten, T. & Bamford, D. H. (1999). *J. Mol. Biol.* **291**, 575–587.

San Martín, 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.

Savilahti, H. & Bamford, D. H. (1986). *Gene*, **49**, 199–205.

Stewart, P. L., Burnett, R. M., Cyrklaff, M. & Fuller, S. D. (1991). *Cell*, **67**, 145–154.

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

Stewart, P. L., Ghosh, S., Bamford, D. H. & Burnett, R. M. (1993). *J. Mol. Biol.* **230**, 349–352.

Teeter, M. M. (1991). *Annu. Rev. Biophys. Biophys. Chem.* **20**, 577–600.

Tuma, R., Bamford, J. K. H., Bamford, D. H., Russell, M. P. & Thomas, G. J. Jr (1996). *J. Mol. Biol.* **257**, 87–101.

Walin, L., Tuma, R., Thomas, G. J. Jr & Bamford, D. H. (1994). *Virology*, **201**, 1–7.

Weik, M., Ravelli, R. B. G., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J. & Sussman, J. L. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 623–628.

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

Zhao, R., Hadfield, A. T., Kremer, M. J. & Rossmann, M. G. (1997). *Virology*, **227**, 13–23.