*J. Mol. Biol.* (1993) **233**, 231–244

# The Canine Parvovirus Empty Capsid Structure

**Hao Wu† and Michael G. Rossmann‡**

*Department of Biological Sciences*
*Purdue University, West Lafayette*
*IN 47907, U.S.A.*

*(Received 23 February 1993; accepted 6 May 1993)*

The structure of empty canine parvovirus capsids shows that residues 37 to the carboxy-terminal residue 584 (VP2 numbering) are ordered in each of the 60 subunits. The central structural motif of each subunit is the eight-stranded antiparallel $\beta$-barrel that has been found in many other virus structures. Five $\beta$-hairpin turns form a $\beta$-cylindrical structure at each icosahedral 5-fold axis. The N-terminal glycine-rich sequence can be accommodated within this cylinder without excessive steric hindrance, consistent with the electron density distribution. By far the largest conformational differences between the full and empty virus were found in the region where some ordered DNA has been observed to bind in canine parvovirus full particles. Extensive interactions among 3-fold related subunits indicate that a trimeric subunit might be a viral assembly intermediate.

*Keywords:* parvovirus structure; empty capsids; assembly; conformational differences; very large unit cell

## 1. Introduction

Parvoviridae, containing primarily the autonomous and the dependo parvoviruses, are the etiologic agents of many diseases in man and other animals (Berns, 1984). Younger animals are most susceptible to parvovirus infection because parvoviruses replicate in proliferating cells (Tattersall, 1972; Berns, 1984). The first episode of canine parvovirus (CPV§) infection occurred around 1978 in various geographically separated areas (Parrish *et al.*, 1988), causing acute enteritis and panleukopenia in dogs of all ages and myocarditis in puppies, all of which could be highly lethal (Berns, 1984). CPV infection is now endemic in most populations of wild and domestic canines (Parrish *et al.*, 1991).

CPV capsids consist of 60 copies of a combination of viral coat proteins VP1 (82·3 kDa), VP2 (67·3 kDa) and VP3 (63·5 kDa) (Paradiso *et al.*, 1982; Cotmore & Tattersall, 1987). While VP1 and VP2 are translated from spliced mRNAs so that VP1 contains the entire sequence of VP2 with a unique N-terminal extension of about 15 kDa, VP3 is derived by cleaving VP2 at the tryptic sites close to the VP2 N terminus (Tattersall & Cotmore, 1988). Considerable conservation of amino acid sequences and similar post-translational processing of viral coat proteins (Tattersall & Cotmore, 1988; Chapman & Rossmann, 1993) characterize the rat virus-like (RV-like) subgroup of autonomous parvoviruses. This group includes CPV, feline panleukopenia virus (FPV), mink enteritis virus (MEV), rat virus (RV), minute virus of mice (MVM), H-1 virus in rats (H-1) and porcine parvovirus (PPV).

CPV packages single-stranded DNA as its genetic material. Both full (DNA-containing) and empty (non-DNA-containing) particles could be readily formed from *in vitro* and *in vivo* infections. They are indistinguishable using immunological and hemagglutinating assays (Saliki *et al.*, 1992). It has been suggested that during the virus life cycle, DNA is encapsidated into at least partly preformed empty capsids (Tattersall & Cotmore, 1988). The relative independence of viral assembly to the presence of DNA suggests that protein–protein interactions dominate the assembly process. Coat protein analyses have shown that CPV full particles contain VP1, VP2 and VP3, whereas CPV empty particles contain only VP1 and VP2. In addition, if CPV full particles are treated with trypsin, their VP2 proteins are cleaved into a VP3-like peptide, but a similar treatment of empty capsids does not

† Present address: Department of Biochemistry, College of Physicians and Surgeons of Columbia University, 630 West 168th Street, New York, NY 10032, U.S.A.

‡ Author to whom all correspondence should be addressed.

§ Abbreviations used: CPV, canine parovovirus; FPV, feline panleukopenia virus; MEV, mink enteritis virus; RV, rat virus; MVM, minute virus of mice; PPV, porcine parvovirus, r.m.s., root-mean-square; SBMV, southern bean mosaic virus, FMDV, foot and mouth disease virus.



0022–2836/93/180231–14 $08.00/0

produce cleavage of VP2 (Clinton & Hayashi, 1976; Tattersall *et al.*, 1977). The absence and presence of VP3 in CPV empty and full capsids, respectively. indicate that the DNA encapsidation is associated with a conformational change of CPV capsids, exposing the N-terminal tryptic sites of VP2 to proteases. It has been suggested that the cleavage of VP2 to VP3 is essential for the infectivity of the virus because the cleavage exposes the conserved glycine-rich sequence which may be important in interacting with cellular membranes (Tattersall & Cotmore, 1988).

The CPV full particle structure determination (Tsao *et al.*, 1991) showed that the virus possesses $T=1$ (Caspar & Klug, 1962) icosahedral symmetry. The 60 copies of the structural motif common to the three viral coat proteins extended from residue 37 to the carboxy-terminal residue 584 of VP2. The polypeptide fold consisted of an eight-stranded antiparallel $\beta$-barrel common to most (but not all) viral capsid proteins determined to date (Rossmann & Johnson, 1989). The $\beta$-strands within this barrel ($\beta$B...$\beta$I), which form two opposing $\beta$-sheets, contain the antiparallel $\beta$-sheets BIDG and CHEF (that is, $\beta$B is antiparallel and adjunct to $\beta$I, $\beta$I is antiparallel to $\beta$D and between $\beta$B and $\beta$D, etc.). Electron density difference maps of the full minus empty particle structure for a very limited data set in monoclinic crystals suggested that in CPV full virus, the conserved amino-terminal glycine-rich sequence of VP2 protrudes to the viral exterior by way of the 5-fold axial channel in approximately one out of five subunits (Tsao *et al.*, 1991; Wu *et* al., 1993). This observation was consistent with the trypsin treatment described above. The difference map also demonstrated the presence of some ordered DNA structure associated with the inside of the viral capsid (Tsao *et al.*, 1991; Wu *et al.*, 1993). Presumably, the icosahedral protein coat has imposed its symmetry onto the secondary structure of part of the DNA molecule within the virion.

The structure determination and refinement of CPV empty particles in a tetragonal space group will be described elsewhere (Wu *et al.*, 1993). The tetragonal crystal presented a particularly difficult crystallographic data collection problem because of the very long $c$ axis of 795 Å. The CPV particles are situated on a crystallographic 2-fold axis giving a 30-fold non-crystallographic redundancy. Phases were determined by non-crystallographic real space averaging to 3·0 Å resolution. The X-ray data restrained in the Hendrickson–Konnert refinement (Hendrickson & Konnert, 1980; Hendrickson, 1985) included both the structure amplitudes and the molecular replacement determined phases (Arnold & Rossmann, 1988). In addition, the 30 non-crystallographically related asymmetric units of the structure were constrained to be identical. At the completion of refinement between 6·0 and 3·0 Å resolution, the crystallographic $R$-factor was 21·1%, the root-mean-square deviation from idealized bond lengths was 0·02 Å and 84 water molecules had been included per non-crystallographic asymmetric unit (Wu *et al.*, 1993). Here we analyze this refined structure and compare it with that of full particles. Their differences might suggest how the viral capsid protein recognizes and encapsidates its own genomic DNA.

## 2. The Model and the Unassigned Density

The final atomic model of the CPV empty particle structure contains residues 37 to 584 according to the numbering of the VP2 polypeptide. The CPV primary sequence (Rhode, 1985) has four cysteine residues, at positions 270, 273, 490 and 494. The two free cysteine residues, C270 and C273, are the mercury binding sites in the fluorescein mercuric acetate ($C_{24}H_{18}Hg_2O_9$) derivative of the CPV full particle crystals (Tsao *et al.*, 1992). Residues C490 and C494 form a disulfide bond and the loop between these cysteines interacts with the icosahedrally ordered DNA inside the full viral particle (Tsao *et al.*, 1991). Residue P465 has a *cis* conformation based on its better fit to the electron density than a *trans* conformation. This proline residue is conserved among the RV-like viruses (Chapman & Rossmann, 1993) and might, therefore, play some role in the correct folding of CPV.

The main chain $\phi$ and $\psi$ torsion angles (Ramachandran & Sasisekharan, 1968) of the refined CPV empty particle structure are shown in a Ramachandran plot (Fig. 1). For each five-residue segment centered at a given residue in the CPV model, 20 segments from a data base of well refined structures (Jones & Thirup, 1986) were selected based on their best fits to the $C^\alpha$ positions of the CPV fragment. The "pep_flip" score of the given residue is the r.m.s. deviation between the position

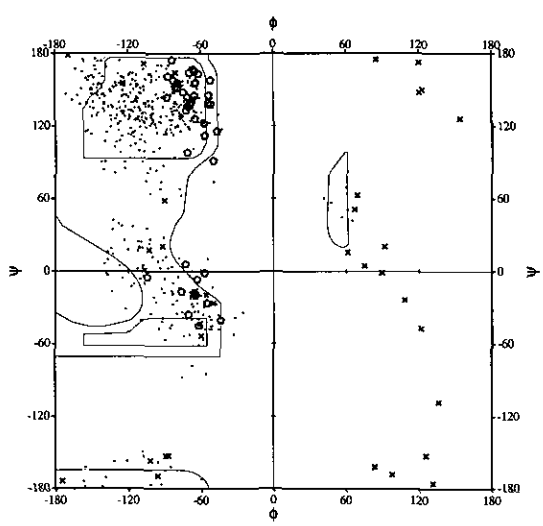



**Figure 1.** Ramachandran diagram of the refined CPV structure. Glycyl residues are marked by crosses, prolines by pentagons and the remaining non-glycyl residues by dots. The boundary of energetically favorable conformations (Ramachandran & Sasisekharan, 1968) is superimposed.

**Table 1**
*Residues with unfavorable main-chain conformations*

| Residue | Segment | Pep_flip (Å) | Remarks[a] |
|---|---|---|---|
| S226 | GTSGT | 2·9 | Involves glycines |
| H234 | IYHGT | 2·8 | Involves glycine |
| V336 | AEVGY | 2·8 | Involves glycine |
| G363 | RGGAQ | 2·8 | Involves glycines |
| D373 | AADGN | 2·5 | Involves glycine |
| T391 | TTTGE | 2·8 | Involves glycine |
| L422 | FNLPV | 2·5 | |
| Q491 | VCQNN | 3·0 | May be a result of the -S-S- bond |

[a] See text for details.

**Table 2**
*The $\chi_1$ side-chain conformations*

| Residue type | $g^-$ | $t$ | $g^+$ | $e$ |
|---|---|---|---|---|
| Arg | 4 | 2 | 14 | 1 |
| Asn | 11 | 10 | 26 | 2 |
| Asp | 2 | 11 | 13 | 0 |
| Cys | 1 | 1 | 2 | 0 |
| Gln | 3 | 8 | 14 | 1 |
| Glu | 3 | 9 | 11 | 0 |
| His | 3 | 1 | 7 | 0 |
| Ile | 2 | 3 | 20 | 1 |
| Leu | 2 | 14 | 18 | 0 |
| Lys | 3 | 11 | 5 | 1 |
| Met | 2 | 4 | 6 | 0 |
| Phe | 5 | 10 | 13 | 0 |
| Ser | 9 | 11 | 6 | 1 |
| Thr | 26 | 3 | 20 | 5 |
| Trp | 5 | 4 | 5 | 0 |
| Tyr | 6 | 7 | 9 | 0 |
| Val | 4 | 2 | 29 | 1 |

The $g^-$, $t$ and $g^+$ conformations are defined by $\chi_1 = 60(\pm 40)°$, $\chi_1 = 180(\pm 40)°$ and $\chi_1 = -60(\pm 40)°$, respectively. The $e$ conformation is for the remaining $\chi_1$ angles.

of the observed central carbonyl oxygen and those of the 20 best fits. Residues with "unusual" main chain conformations, as assessed by having larger than 2·5 Å pep_flip scores, are listed in Table1 and often involved glycines. The high pep_flip score of residue Q491 is possibly a result of the tight loop formed by the disulfide bond between C490 and C494. The side-chain conformations of each type of residue (Table 2) showed that the three staggered conformations are highly populated, which is consistent with studies of well-refined crystal structures (Janin *et al.*, 1978; Benedetti *et al.*, 1983; Ponder & Richards, 1987).

Residues that have poorly defined electron density either for their main-chain atoms or only for their side-chain atoms are shown in Table 3. These residues mostly reside on the surface of the viral protein shell, either externally, exposed to the solvent, or internally, facing the nucleic acid. Most notably, two stretches of residues, from 155 to 163 and from 361 to 372, are associated with the weakest interpretable density in the map. There are

**Table 3**
*Residues associated with weak density and high temperature factors*

| Residue | $B_{max}$ (Å$^2$)[a] | Residue | $B_{max}$ (Å$^2$)[a] | Residue | $B_{max}$ (Å$^2$)[a] |
|---|---|---|---|---|---|
| A. *Main-chain (at least one main-chain atom with* B > *40 Å$^2$)* | | | | | |
| S154 | 41 | G299 | 42 | E368 | 63 |
| E155 | 66 | A300 | 40 | N369 | 61 |
| S156 | 79 | G360 | 41 | Q370 | 56 |
| A157 | 89 | R361 | 49 | A371 | 55 |
| T158 | 91 | G362 | 59 | A372 | 50 |
| Q159 | 94 | G363 | 61 | N492 | 46 |
| P160 | 96 | A364 | 58 | N493 | 43 |
| P161 | 84 | Q365 | 62 | S515 | 44 |
| T162 | 71 | T366 | 63 | A516 | 44 |
| K163 | 51 | D367 | 65 | N517 | 41 |
| **Residue** | **$B_{max}$ (Å$^2$)[c]** | **Residue** | **$B_{max}$ (Å$^2$)[c]** | **Residue** | **$B_{max}$ (Å$^2$)[c]** |
| B. *Side-chain only (at least one side-chain atom with* B > *40 Å$^2$)*[b] | | | | | |
| E50 | 41 | S226 | 46 | E393 | 49 |
| E55 | 65 | K271 | 45 | E411 | 43 |
| N56 | 43 | Q296 | 52 | Q491 | 42 |
| E60 | 40 | Q309 | 51 | Q509 | 44 |

[a] Maximum $B$-value among all main-chain atoms for the residue.
[b] Those residues whose main-chain atom $B$-values are also greater than 40 Å$^2$ (listed in A) are not listed here.
[c] Maximum $B$-value among all side-chain atoms for the residue.

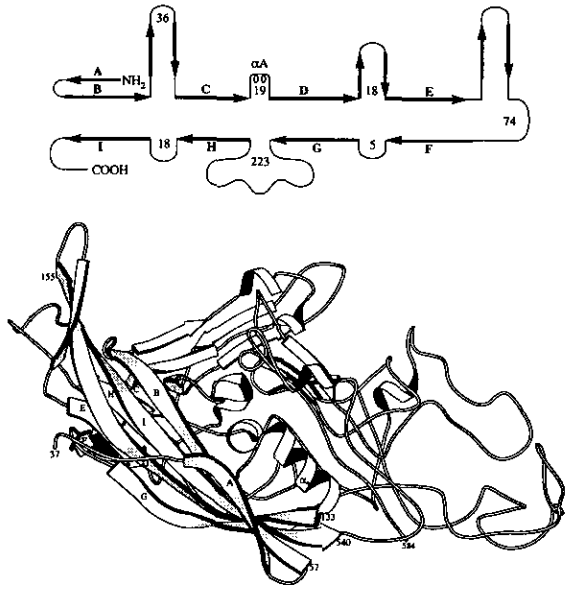



**Figure 2.** The linear representation of the CPV capsid protein structure and the ribbon diagram where the central $\beta$-strands (A to I) and the conserved helix ($\alpha$A) are labeled. The ribbon diagram, as well as Figs 4, 8 and 9, was created using the program MOLSCRIPT (Kraulis, 1991).

two pieces of unmodeled density along each of the icosahedral 3-fold symmetry axes. One (0·75$\sigma$ to 1·0$\sigma$ in height, where $\sigma$ is the r.m.s. deviation of the electron density map) is surrounded by the three symmetry-related carbonyl oxygen atoms (3·8 Å distant) of residue I447 and the other (1·5$\sigma$ to 2·0$\sigma$) by the three main-chain nitrogen atoms (3·5 Å away) of residue N437, indicating that some molecules or ions sit on the 3-fold axes to stabilize the partial charges of oxygen or nitrogen atoms. There is also some diffuse density connected to the most amino-terminal residue 37.

## 3. Secondary and Tertiary Structure

The CPV tertiary fold is shown schematically as a ribbon diagram (Fig. 2) and as a hydrogen bonding diagram (Fig. 3). The central structural motif is the canonical jelly-roll type $\beta$-barrel that has been found in many other virus structures (Rossmann & Johnson, 1989). Extensive insertions or loops between the strands in the CPV $\beta$-barrel form two-thirds of the total structure. In particular, the GH loop alone (the loop between the G and H strands) contains 220 residues.

The secondary structural elements of the CPV VP2 polypeptide (Table 4) were assigned on the

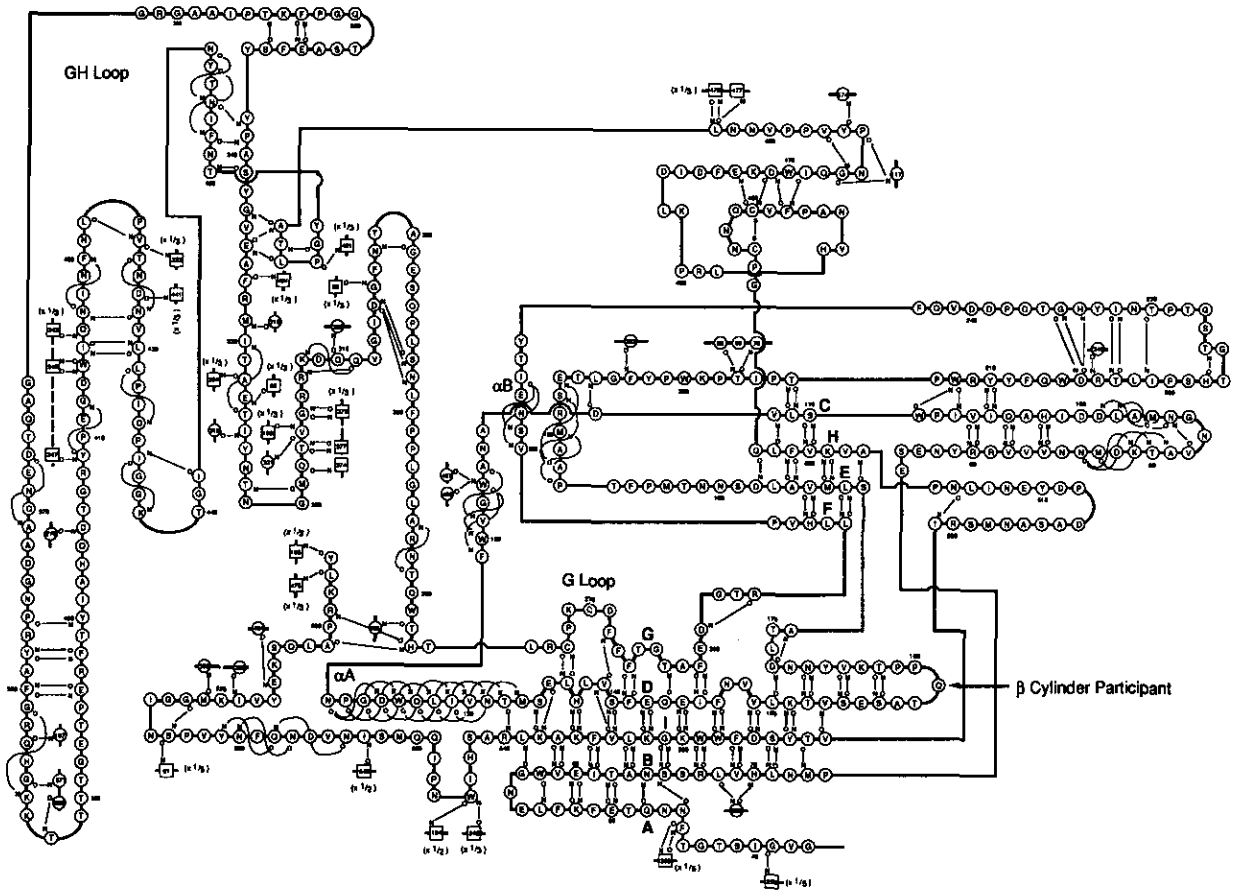



**Figure 3.** Main-chain hydrogen bonding diagram of the CPV capsid protein. A hydrogen bond is defined by the O...N distance less than 3·5 Å and the C=O...N angle greater than 125°. Interactions with neighboring subunits (residues in squares) are indicated by $\times 1/2$, $\times 1/3$ or $\times 1/5$ for 2, 3 or 5-fold rotationally related subunits. Secondary structural elements are also identified. $\beta$-Strands in the $\beta$-barrel are labeled A, B,... I and the two conserved $\alpha$-helices are labeled $\alpha$A and $\alpha$B.

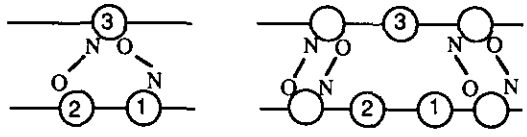
**Table 4**
*Secondary structural elements*

A. *β-Strands and α-helices*[a]

| | Residues | | Residues |
|---|---|---|---|
| $\beta_A$ | 46–54 | Within the EF loop: | |
| $\beta_B$ | 57–73 | $\alpha_{EF}$ | 187–193 |
| Within the BC loop: | | $\beta_{EF}$ | 206–210 |
| $\beta_{BC}$ | 80–84 | $\beta'_{EF}$ | 215–219 |
| $\alpha_{BC}$ | 86–91 | $\beta''_{EF}$ | 230–235 |
| $\alpha'_{BC}$ | 95–99 | $\beta_F$ | 253–255 |
| $\beta''_{BC}$ | 102–109 | $\beta_G$ | 261–273 |
| $\beta_C$ | 110–112 | Within the GH loop: | |
| Within the CD loop: | | $\beta_{GH}$ | 376–380 |
| $\alpha_{CD}$ | 116–120 | $\beta'_{GH}$ | 396–400 |
| $\alpha_A$ | 122–133 | $\beta''_{GH}$ | 469–473 |
| $\beta_D$ | 132–156 | $\beta'''_{GH}$ | 488–490 |
| Within the DE loop: | | $\beta_H$ | 497–503 |
| $\beta_{DE}$ | 162–168 | $\beta_I$ | 522–540 |
| $\beta_E$ | 172–178 | C terminus: | |
| | | $\alpha_{cterm}$ | 557–561 |

B. *β-Bulges*

*Classical bulge* *Wide bulge*




| Position 1 | Position 2 | Position 3 | Classification |
|---|---|---|---|
| S134 | E135 | K538 | Classical |
| V139 | S140 | V534 | Classicial |
| N147 | V148 | F527 | Wide |
| G264 | T265 | E142 | Wide |

[a] Within the virus β-barrel, the strands are named $\beta_A$ to $\beta_I$. Within the loops, the β-strands that contain at least 4 hydrogen bonds and the α-helices that contain at least 2 consecutive turns are listed by the name of loops. The $\alpha_A$ helix is named so because of its homology to other viruses.

basis of main-chain hydrogen bonding patterns. The central antiparallel β-barrel consists of two opposing β-sheets. The strands B, I, D and G form the highly twisted BIDG sheet. Strand A hydrogen bonds with the B strand in the BIDG sheet, which extends the latter to a five-stranded β-sheet in the same manner as the C subunit of $T=3$ plant and insect viruses or VP2 of picornaviruses. The strands C, H, E and F form the opposing CHEF sheet of the β-barrel. There are several β-structures within the loops that connect these central β-strands, including the DE β-hairpins that form the 5-fold β-cylindrical structure.

There are four β-bulges in the VP2 polypeptide structure (Table 4), two classical bulges and two wide β-bulges (Richardson *et al.*, 1978). The β-bulges at residues 527 and 534 (numbering is on position 3) are at similar positions along the length of the BIDG sheet as those in HRV14 and Mengo virus (Arnold & Rossmann, 1990; Krishnaswamy & Rossmann, 1990), although the extra residues are in different strands. β-Bulges accentuate the right-handed twist of β-sheets (Richardson *et al.*, 1978). Since the BIDG sheet is usually highly twisted in all virus capsid proteins, these β-bulges maintain the twist of the sheet.

There are only a few helical structures. Except helix αA, which has eight consecutive turns, other helices have only two or three turns. Helix αA occurs at an equivalent position as a conserved helix in many virus structures including HRV14 (Arnold & Rossmann, 1990), Mengo (Krishnaswamy & Rossmann, 1990), southern bean mosaic virus (SBMV) (Silva & Rossmann, 1987) and tomato bushy stunt virus (Harrison *et al.*, 1978). Helix αB, which occurs in many virus strucures (Rossmann & Johnson, 1989), is represented only by two α-helix-like hydrogen bonds (Fig. 3).

## 4. The 5-fold β-Cylindrical Structure

Five symmetry-related β-hairpins form a β-cylinder around each of the icosahedral 5-fold axes (Fig. 4). In picornaviruses (Arnold & Rossmann, 1990; Krishnaswamy & Rossmann, 1990), the amino termini of VP3 twist together at each 5-fold axis to form a cylindrical parallel β-sheet, which could play a role in pentamer formation. The five β-hairpins in CPV, however, are too distant from one another to make main-chain hydrogen bonding interactions. Residue Q159 sits at the most exterior of the β-cylinder at a viral radius of 132 Å. Proceeding into the interior of the virus, the cylinder is lined by consecutive layers of amino acid side-chains facing the 5-fold axis. These are, starting from the outside, T158, (S156 and T162), (S154 and V164), (T152 and N166), (L150 and D168), and finally L169 closest to the inside. What is remarkable is the number of small polar residues represented by threonine and serine. The base of the cylinder is stabilized by a hydrogen bonding network among residues K151, N167 and D168. The narrowest part of the cylinder is at a viral radius of 101 Å with the $C^{\delta 1}$ atom of L169 only 4 Å away from the 5-fold axis.

Residue 37 is the first amino-terminal residue that shows completely ordered electron density. Although the β-cylindrical structure was seen for both full and empty viruses, the full minus empty difference electron density and biochemical information (Clinton & Hayashi, 1976; Tattersall *et al.*, 1977) suggest that the amino ends of some VP2 polypeptides emerge to the outside through the 5-fold axes only in the case of the full virus (Tsao *et al.*, 1991; Wu *et al.*, 1993). Since the sequence (Rhode, 1985) prior to residue 37 is (22)GSGNGSG-GGGGGGSG(36) and residue 37 is situated at the base of the cylinder, these small residues should be able to thread through the β-cylinder for at least one of the surrounding subunits. A model of this sequence inserted along the 5-fold axis could be built with little steric hindrance (Fig. 4). Furthermore, where the channel becomes wider at the outside of the virion, there is no further density in the full particle structure, presumably because the glycine-rich sequence is unrestrained and, hence, appears to be entirely disordered. This unusual

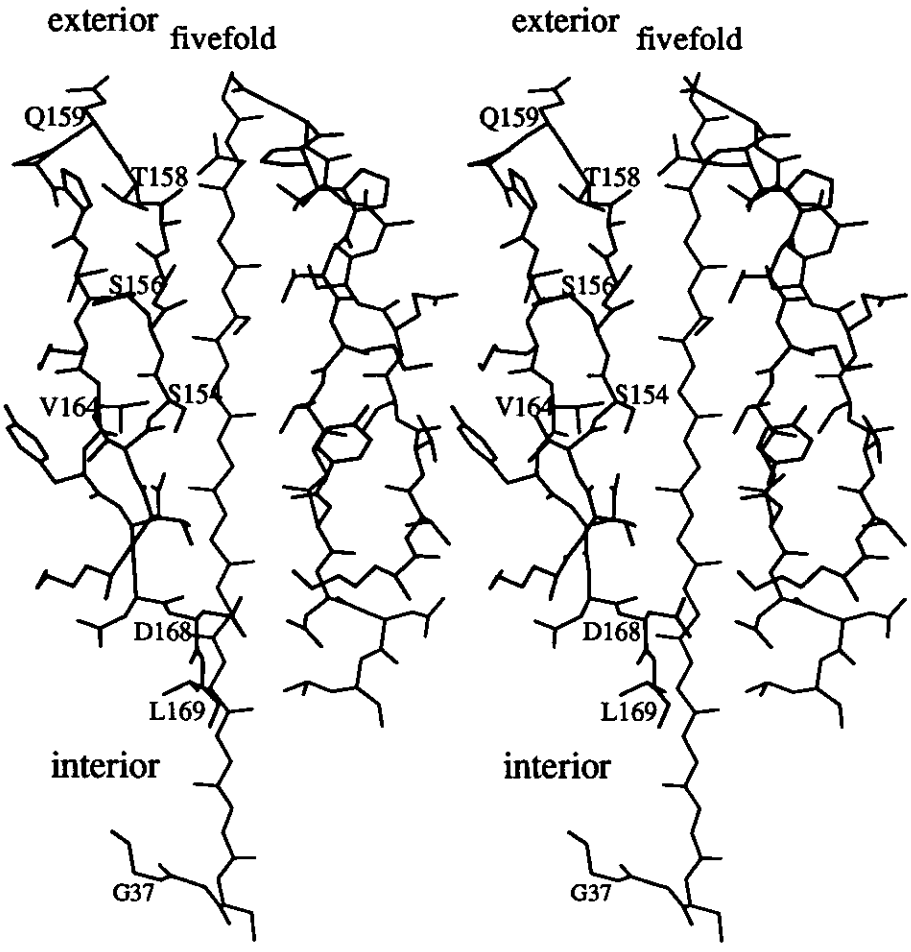



**Figure 4.** Two of the $\beta$-hairpins that make up the 5-fold $\beta$-cylindrical structure are shown with respect to the 5-fold axis. The glycine-rich sequence preceding residue G37 could be built through the inside of the $\beta$-cylinder with little steric hindrance.

situation of a peptide running through a 5-fold channel is possibly important for membrane translocation of the infectious virion (Tattersall & Cotmore, 1988) and is reminiscent of the loss of the internal VP4 protein from picornaviruses, possibly by way of the 5-fold axis, soon after viral attachment to cell receptors (Lonberg-Holm *et al.*, 1976; Giranda *et al.*, 1992). It is noteworthy that the bacteriophage $\phi$X174 has an extensive ion channel running along the 5-fold axis (McKenna *et al.*, 1992) and that the ion transporting properties of the 5-fold channel in SBMV have also been noted (Silva & Rossmann, 1987).

## 5. Subunit Interactions

Subunit interactions related by icosahedral 2-, 3- and 5-fold axes are shown in the $C^{\alpha}$ trace of Figure 5(a), (b) and (c), respectively. The number of close interactions (<3·5 Å) was analyzed and subdivided into non-polar, mixed (non-polar–polar) and polar interactions, the latter including charged attractions (Tables 5, 6, 7 and 8).

The most extensive interactions are between 3-fold-related subunits. These subunits are highly intertwined with one another (Fig. 5(b)), forming the prominent 3-fold spike at the virus surface. The surface loops around the 3-fold axes are loosely packed (Fig. 2) and perhaps are ordered only when the subunits are associated with each other. A similar situation exists for the hexon of adenovirus where the loops forming the 3-fold "tower" would probably only be ordered when interacting with one another (Roberts *et al.*, 1986; Stewart *et al.*, 1991). Most of the 3-fold related interactions are between the GH loops of neighboring subunits, although the BC loop, the EF loop and part of the C-terminal segment also contribute. A large proportion of the interactions between the 3-fold-related subunits is polar (Tables 5 and 7), including a

**Table 5**
*Number of interactions between icosahedrally related subunits*

| | Total | Hydrophobic | Mixed | Polar | Ion pairs |
|---|---|---|---|---|---|
| 2-fold | 53 | 11 | 27 | 15 | 1 |
| 3-fold | 296 | 38 | 156 | 102 | 12 |
| 5-fold | 82 | 14 | 36 | 32 | 1 |

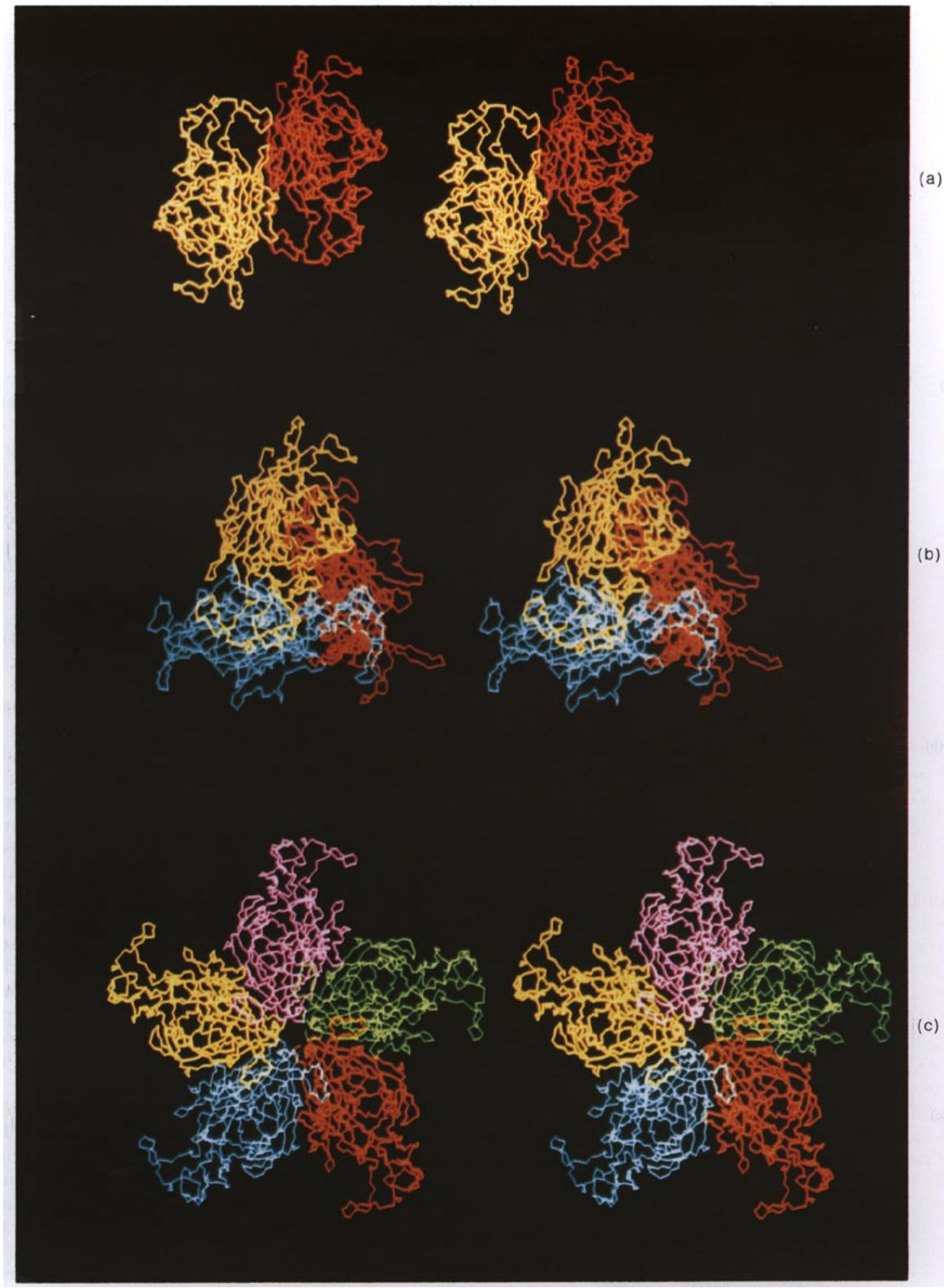



**Figure 5.** (a) A dimer of subunits; (b) a trimer of subunits; (c) a pentamer of subunits.

significant number of main-chain to main-chain hydrogen bonds (Fig. 6). These main-chain interactions are between residues from the 3-fold-related GH loops which are long and yet have only a few defined secondary structural elements. The 3-fold interactions, therefore, not only fill the empty spaces within the loops, but also help to satisfy the buried polar atoms, which may add to the specificity of subunit–subunit recognition.

The viral capsid is thinnest at the 2-fold axes, as

**Table 6**
*Possible hydrogen bonding between 2-fold-related subunits*

| Atom | Atom (× 1/2) | Distance (Å) | Atom | Atom (× 1/2) | Distance (Å) |
|---|---|---|---|---|---|
| G124 N | W545 O | 2·8 | E298 $O^{\varepsilon 2}$ | S564 $O^{\gamma}$ | 2·3 |
| Q127 $O^{\varepsilon 1}$ | I548 N | 2·9 | E298 $O^{\varepsilon 2}$ | T389 $O^{\gamma 1}$ | 2·4 |
| Q127 $N^{\varepsilon 2}$ | I548 O | 2·7 | N302 O | N565 $N^{\delta 2}$ | 3·1 |
| S297 O | N565 $N^{\delta 2}$ | 2·9 | I548 O | I553 N | 3·1 |
| E298 $O^{\varepsilon 1}$ | K387 $N^{\zeta}$ | 3·1 | Q550 N | S552 $O^{\gamma}$ | 3·2 |
| E298 $O^{\varepsilon 1}$ | T389 $O^{\gamma 1}$ | 3·3 | | | |

Interactions are selected for appropriate atom pairs separated by less than 3·5 Å without regard for direction.

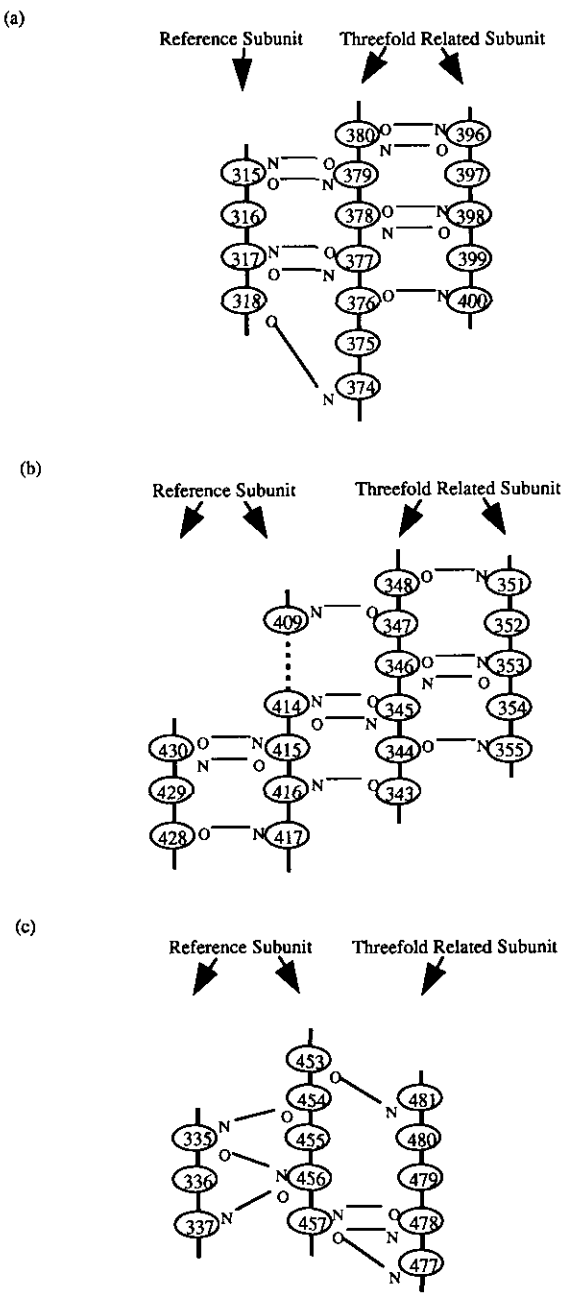



**Figure 6.** Three different sets of main-chain hydrogen bonding interactions between 3-fold-related subunits.

manifested by the 2-fold depression on the viral surface. Least extensive interactions exist between the 2-fold-related subunits (Fig. 5(a)). The dimer interface is rather hydrophobic (Tables 5 and 6), being mostly formed by the αA helix, part of the GH loop around residue 300 and part of the C-terminal region. Although the αA helix contributes to the 2-fold interactions in many other viruses (Rossmann *et al.*, 1983; Fisher & Johnson, 1993), the interaction of the helices in CPV is end-on *instead of antiparallel. The 5-fold-related interactions* (Fig. 5(c)) are intermediate in quantity to those of the 3-fold and 2-fold-related subunit interfaces (Tables 5 and 8).

Buried surface area is an approximate measure of the strength of subunit interactions. The total surface area per subunit in a monomer, dimer, trimer and pentamer (Table 9) was computed using the Kabsch & Sander (1983) algorithm and is broken down into that from hydrophobic, aromatic, non-charged hydrophilic and charged residues. The area that becomes buried on trimer formation is 10,689 $Å^2$ per subunit, which is much larger than that hidden by dimer (1806 $Å^2$) or pentamer (4440 $Å^2$) formation. Although a substantial surface area of charged residues is buried in the trimeric unit, the atoms holding the charge are mostly exposed. The surface area of the GH loop, the principal contributor to trimer interaction, is reduced by 7500 $Å^2$ on trimer formation, representing 70% of the total buried surface.

Every 40 $Å^2$ of buried surface area equals roughly 1 kcal/mol of association energy (Richards, 1977), assuming that the increase in entropy of solvent molecules is the driving force for subunit association. The free energy decrease by forming a trimer (3 × 10,689 $Å^2$ buried surface area) is therefore about 802 kcal/mol. In HRV14 for which the assembly pathway is known, the free energy decrease of the first step of assembly, the association of VP1, VP0 and VP3 into a protomer, was estimated to be 475 kcal/mol (Arnold & Rossmann, 1990). In FMDV, the same protomer formation step has a free energy decrease of 386 kcal/mol (Fry *et al.*, 1990). Thus, if the assembly process of CPV is energetically driven, it is highly probable that the assembly goes through highly stable trimeric inter-

## Table 7

*Possible hydrogen bonding between 3-fold-related subunits*

| Atom | Atom (×1/3) | Distance (Å) | Atom | Atom (×1/3) | Distance (Å) |
|---|---|---|---|---|---|
| R274 $N^{\eta 1}$ | D475 $O^{\delta 2}$ | 2·9 | G320 O | T399 $O^{\gamma 1}$ | 2·8 |
| R274 $N^{\eta 2}$ | D475 $O^{\delta 1}$ | 2·8 | N321 $O^{\delta 1}$ | Y343 $O^{\eta}$ | 3·0 |
| R274 $N^{\eta 2}$ | D475 $O^{\delta 2}$ | 3·2 | T322 O | R377 $N^{\varepsilon}$ | 3·1 |
| H277 $N^{\varepsilon 2}$ | D240 $O^{\delta 1}$ | 2·8 | N323 $O^{\delta 1}$ | R377 $N^{\eta 2}$ | 2·7 |
| H277 $N^{\varepsilon 2}$ | D239 O | 3·0 | N323 $O^{\delta 1}$ | R377 $N^{\varepsilon}$ | 3·0 |
| W279 $N^{\varepsilon 1}$ | Q242 $O^{\varepsilon 1}$ | 3·2 | T326 $O^{\gamma 1}$ | A97 O | 2·9 |
| W279 $N^{\varepsilon 1}$ | D240 O | 3·2 | E327 N | D100 $O^{\delta 1}$ | 2·9 |
| R283 $N^{\eta 1}$ | P352 O | 2·3 | E327 N | D99 O | 3·3 |
| R283 $N^{\eta 1}$ | D99 $O^{\delta 2}$ | 3·1 | E327 $O^{\varepsilon 1}$ | Y211 $O^{\eta}$ | 2·9 |
| R283 $N^{\eta 2}$ | D99 $O^{\delta 2}$ | 2·6 | E327 $O^{\varepsilon 2}$ | R191 $N^{\eta 2}$ | 2·5 |
| R283 $N^{\eta 2}$ | I101 O | 3·2 | A328 O | K354 N | 3·0 |
| L287 O | R191 $N^{\varepsilon}$ | 2·7 | R332 $N^{\eta 2}$ | Y211 $O^{\eta}$ | 3·0 |
| L287 O | R191 $N^{\eta 2}$ | 3·2 | R332 $N^{\eta 2}$ | Y211 $O^{\eta}$ | 3·5 |
| P288 O | R209 $N^{\eta 2}$ | 2·9 | P333 O | V484 N | 3·0 |
| P288 O | R209 $N^{\varepsilon}$ | 3·0 | S339 $O^{\gamma}$ | E346 $O^{\varepsilon 1}$ | 2·9 |
| P288 O | R191 $N^{\varepsilon}$ | 3·4 | Q365 $N^{\varepsilon 2}$ | T217 $O^{\gamma 1}$ | 3·2 |
| P295 O | R80 $N^{\eta 1}$ | 3·1 | G407 O | S348 $O^{\gamma}$ | 2·7 |
| P295 O | R80 $N^{\eta 2}$ | 3·3 | R408 $N^{\varepsilon}$ | E346 $O^{\varepsilon 1}$ | 2·8 |
| G304 N | V83 O | 2·8 | R408 $N^{\eta 2}$ | E346 $O^{\varepsilon 1}$ | 3·3 |
| Q310 $N^{\varepsilon 2}$ | L98 O | 2·7 | Y409 N | A347 O | 2·8 |
| Q310 $N^{\varepsilon 2}$ | D100 $O^{\delta 2}$ | 3·0 | D413 $O^{\delta 1}$ | A347 N | 2·8 |
| K312 $N^{\zeta}$ | T394 $O^{\gamma 1}$ | 3·5 | W414 N | F345 O | 3·1 |
| R313 $N^{\eta 1}$ | D100 $O^{\delta 2}$ | 2·8 | W414 O | F345 N | 3·0 |
| R313 $N^{\eta 2}$ | D100 $O^{\delta 2}$ | 2·9 | Q416 N | Y343 O | 2·7 |
| R314 $N^{\eta 1}$ | S192 O | 3·0 | N417 $N^{\delta 2}$ | G441 O | 2·6 |
| G315 N | A379 O | 2·9 | P423 O | T223 $O^{\gamma 1}$ | 3·3 |
| G315 O | A379 N | 3·1 | V424 N | G94 O | 2·9 |
| V316 N | M190 O | 3·3 | V424 O | T223 N | 2·6 |
| T317 N | R377 O | 3·0 | N426 $O^{\delta 1}$ | H222 $N^{\varepsilon 2}$ | 2·9 |
| T317 O | R377 N | 2·7 | N426 $O^{\delta 1}$ | K439 $N^{\zeta}$ | 3·0 |
| Q318 $O^{\varepsilon 1}$ | I357 N | 3·0 | N426 O | K439 $N^{\zeta}$ | 2·6 |
| Q318 $N^{\varepsilon 2}$ | I357 O | 2·8 | D427 O | G441 N | 2·7 |
| V429 O | K439 $N^{\zeta}$ | 2·6 | N459 N | D477 $O^{\delta 1}$ | 3·0 |
| D434 $O^{\delta 2}$ | K439 $N^{\zeta}$ | 2·9 | R481 $N^{\eta 2}$ | K479 O | 3·1 |
| N449 O | N449 $N^{\delta 2}$ | 2·7 | W545 N | F243 O | 3·0 |
| T450 $O^{\gamma 1}$ | E346 $O^{\varepsilon 1}$ | 2·7 | N546 $O^{\delta 1}$ | F243 N | 2·7 |
| T450 $O^{\gamma 1}$ | E346 $O^{\varepsilon 2}$ | 3·2 | N546 $N^{\delta 2}$ | V241 O | 3·0 |
| Y451 O | N449 $N^{\delta 2}$ | 3·5 | Q549 $O^{\varepsilon 1}$ | D239 $O^{\delta 1}$ | 3·0 |
| P453 O | R481 N | 2·8 | R581 $N^{\eta 1}$ | D475 O | 2·2 |
| L457 N | L478 O | 3·3 | E583 O | D475 N | 2·9 |
| L457 O | L478 N | 2·7 | Y584 $O^{\eta}$ | D471 $O^{\delta 2}$ | 2·7 |
| L457 O | D477 N | 3·3 | Y584 $O^{\eta}$ | D471 $O^{\delta 1}$ | 3·5 |
| N458 $O^{\delta 1}$ | D477 N | 2·7 | Y584 O | K472 $N^{\zeta}$ | 3·0 |
| N458 $O^{\delta 1}$ | T476 $O^{\gamma 1}$ | 3·3 | Y584 O | F185 N | 3·2 |

Interactions are selected for appropriate atom pairs separated by less than 3·5 Å without regard for direction.

## Table 8

*Possible hydrogen bonding between 5-fold-related subunits*

| Atom | Atom (×1/5) | Distance (Å) | Atom | Atom (×1/5) | Distance (Å) |
|---|---|---|---|---|---|
| R80 $N^{\eta 1}$ | N565 O | 2·6 | R256 $N^{\eta 1}$ | V38 O | 3·4 |
| R81 N | S564 O | 3·2 | R256 $N^{\eta 2}$ | V148 O | 3·1 |
| R81 $N^{\eta 1}$ | P563 O | 2·9 | G258 N | G39 O | 2·7 |
| R81 O | N565 N | 3·5 | D259 $O^{\delta 2}$ | S41 N | 2·8 |
| D168 $O^{\delta 1}$ | N167 $N^{\delta 2}$ | 2·8 | D259 $O^{\delta 2}$ | S41 $O^{\gamma}$ | 3·3 |
| D168 $O^{\delta 2}$ | K151 $N^{\zeta}$ | 3·1 | L506 O | Y524 $O^{\eta}$ | 3·1 |
| T170 $O^{\gamma 1}$ | N167 $O^{\delta 1}$ | 3·1 | T507 O | H70 $N^{\varepsilon 2}$ | 2·8 |
| S172 $O^{\gamma}$ | N147 $N^{\delta 2}$ | 3·2 | N508 $O^{\delta 1}$ | Y165 $O^{\eta}$ | 3·0 |
| T236 O | Q558 $N^{\varepsilon 2}$ | 3·2 | N508 $O^{\delta 1}$ | K163 $N^{\eta}$ | 3·3 |
| N248 $N^{\delta 2}$ | N47 O | 3·5 | N508 O | H70 $N^{\varepsilon 2}$ | 3·2 |
| S249 O | Q48 $N^{\varepsilon 2}$ | 2·8 | N508 O | N72 $N^{\delta 2}$ | 3·4 |
| V252 N | F45 O | 3·2 | Y510 $O^{\eta}$ | P202 O | 2·6 |
| V252 O | F45 N | 2·8 | D513 $O^{\delta 2}$ | Q383 N | 3·3 |
| L254 O | S41 $O^{\gamma}$ | 2·7 | D513 $O^{\delta 2}$ | R382 N | 3·4 |
| R256 $N^{\eta 1}$ | I40 O | 3·2 | D513 O | R382 $N^{\eta 1}$ | 3·2 |

Interactions are selected for appropriate atom pairs separated by less than 3·5 Å without regard for direction.

**Table 9**

*The surface area (Å$^2$) per non-hydrated subunit in a monomer, dimer, trimer or pentamer, broken down into several classes*

| | Monomer | Dimer | Trimer | Pentamer |
|---|---|---|---|---|
| Total capsid protein | 31,536 | 29,674 | 20,509 | 26,758 |
| | | (1862) | (11,027) | (4778) |
| Hydrophobic residues[a] | 10,018 | 9405 | 5923 | 8459 |
| | | (613) | (4095) | (1559) |
| Aromatic residues[b] | 3939 | 3558 | 2140 | 3205 |
| | | (381) | (1799) | (734) |
| Non-charged hydrophilic residues[c] | 9198 | 8623 | 7149 | 7864 |
| | | (575) | (2049) | (1334) |
| Charged residues[d] | 8381 | 8088 | 5297 | 7230 |
| | | (293) | (3084) | (1151) |
| BC loop residues[e] | 1893 | 1893 | 1107 | 1560 |
| | | (0) | (786) | (333) |
| DE loop residues[f] | 1453 | 1453 | 1438 | 974 |
| | | (0) | (15) | (479) |
| GH loop residues[g] | 13,824 | 13,597 | 6004 | 13,402 |
| | | (227) | (7820) | (422) |

Numbers in parentheses are the surface area buried per subunit on oligomer formation.
[a] Residues A, I, L, P, V, G, M, C.
[b] Residues F, Y, W.
[c] Residues N, Q, S, T.
[d] Residues D, E, H, K, R.
[e] Residues 74 to 109.
[f] Residues 154 to 171.
[g] Residues 274 to 496.

mediates. Unlike the assembly of picornaviruses, there is no apparent contribution to the assembly of the capsid by post-translational polypeptide processing.

## 6. Temperature Factor Variation

The mean temperature factor is 13·9 Å$^2$ for all protein atoms, 13·7 Å$^2$ for main-chain atoms alone and 14·0 Å$^2$ for side-chain atoms. This is remarkable both because of the unusually low average temperature factor and because of the rough equivalence of side-chain and main-chain temperature factors. The absolute value of the temperature factors is, however, somewhat arbitrary depending on the basis used for scaling the data films to each other. Comparison with other refined virus structures shows similar results, with the mean temperature factor of main-chain and side-chain atoms being 15·2 Å$^2$ and 15·5 Å$^2$, respectively, for HRV14; 15·4 Å$^2$ and 16·1 Å$^2$, respectively, for Mengo virus; and 20·6 Å$^2$ and 21·7 Å$^2$, respectively, for poliovirus.

The average temperature factor of each residue roughly correlates with the mean accessibility and the secondary structure of each residue (Fig. 7). The accessibility of each residue was defined by the ratio of exposed surface area (Kabsch & Sander, 1983), either exposed to solvent or nucleic acids, to its total surface area (Shrake & Rupley, 1973); 85% of the residues whose mean temperature factors are higher than twice the average temperature factor of the whole protein are at least 20% accessible. Only 11% of the residues whose temperature factors are lower than the mean have larger than 20% accessibility. However, not all exposed regions have high temperature factors. Protein flexibility has been shown to be an important functional determinant (Huber & Bennett, 1983) in antigenicity (Tainer *et*

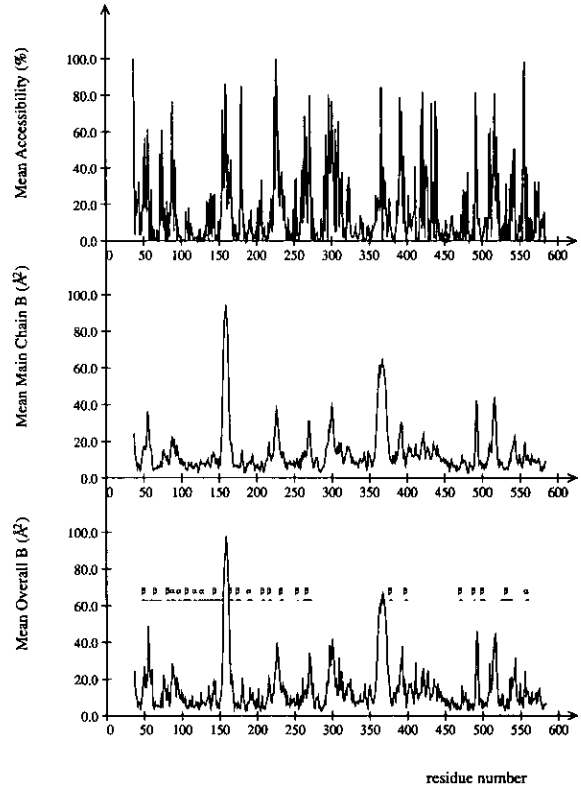



**Figure 7.** The average temperature factor of each residue is compared with its solvent accessibility. The secondary structural elements are also shown.

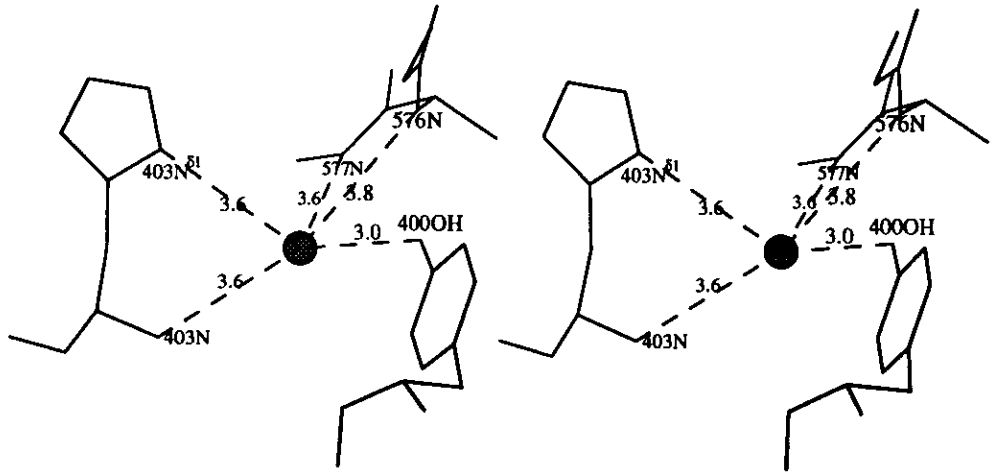



**Figure 8.** The environment of a putative negatively charged group (dark shaded circle).

*al.*, 1984; Sheriff *et al.*, 1985), receptor binding (Ruoslahti, 1988; Kim *et al.*, 1990) and substrate binding (Huber & Bode, 1978). Residues 156 to 161, which have the highest temperature factors in the whole structure, lie at the top of the 5-fold $\beta$-cylindrical structure. The flexibility of these residues may be required for the participation of the $\beta$-cylinder in the early stages of viral infection (Tattersall & Cotmore, 1988; Tsao *et al.*, 1991). Residues 360 to 370, which is the second most flexible region in the structure, are adjacent to residue 377, shown to be important for erythrocyte binding in hemagglutination (Barbis *et al.*, 1992). These residues are more ordered in FPV (Agbandje *et al.*, 1993) relative to CPV, a conformational change associated with a lack of ability for FPV to hemagglutinate at neutral pH (Barbis *et al.*, 1992).

Within $\beta$-strands or $\alpha$-helices, the temperature factors are low (11·9 Å$^2$) as compared to the regions connecting these secondary structures (15·7 Å$^2$). The only piece of secondary structure within which there is a stretch of residues with higher temperature factor is $\beta$G. This high temperature factor region corresponds to the large bulge (or loop) within the strand, present in most viral canonical $\beta$-barrels.

## 7. Solvent Structure

A total of 84 water molecules per VP2 subunit (2520 per crystallographic asymmetric unit) were included in the atomic model with an average temperature factor of 32·7 Å$^2$ (Wu *et al.*, 1993). One putative water molecule has an excessively low temperature factor of 2·0 Å$^2$. The electron density of this putative water molecule is at least as high as the protein region. It interacts with five neighboring protein atoms less than 3·8 Å away, all of which are hydrogen bond donors (Fig. 8). This suggests that the putative water molecule might be a negatively charged group, although neither phosphate nor sulfate was used in the crystallization of CPV.

## 8. DNA Binding Surface

Difference maps between data for full and empty particles have shown some moderately ordered DNA structure (11 nucleotides) associated with the internal surface of the CPV full particle protein coat (Tsao *et al.*, 1991; Wu *et al.*, 1993). The disulfide loop, from residues C490 to C494, forms part of the DNA interacting surface, where there are the largest conformational differences between CPV full and empty particles (Fig. 9). Side-chains of residues N492 and N493 in the CPV full particle structure would clash with the putative pyrimidine base of NA4 and the putative purine base of NA3, respectively, had they remained in the conformation seen in the CPV empty particle structure. The greatest movement of $C^{\alpha}$ positions is 2·1 Å and of side-chain atoms is 7·9 Å, both in residue N493. However, residues Q491 and N492 also showed dramatically different conformations. In the CPV empty particle structure, the side-chain of residue D269 is positioned towards the disulfide loop, while in full particles it was "pushed" away by the conformational change in the loop. Conformational changes were also seen in residues P495 and F266. The phenyl ring of F266 possibly becomes stacked with the putative purine base of NA3 in CPV full particles.

Residues in the loop between C490 and C494 are not particularly conserved among parvoviruses in general (Chapman & Rossmann, 1993), although in RV-like parvoviruses, residues 492 and 493 are invariant and residue 491 is either glutamine or lysine (Chapman & Rossmann, 1993). While on the one hand the region interacting with DNA might be particularly conserved to maintain the DNA binding ability, on the other hand it may vary to recognize different DNA sequences.

## 9. Evolutionary Comparisons

Most viral coat protein structures solved to date have the similar canonical $\beta$-barrel fold (Rossmann

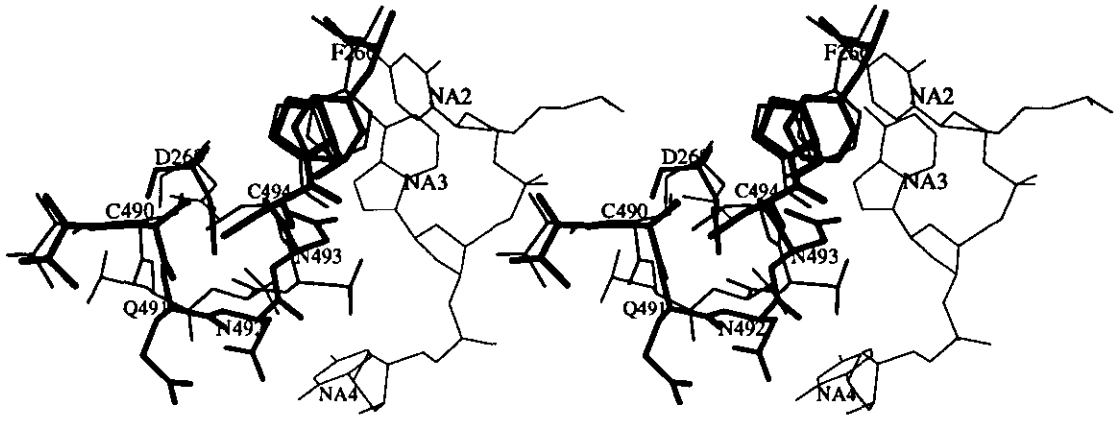



**Figure 9.** The conformational difference between CPV empty and full capsids in the region of icosahedrally ordered DNA. Thick lines represent the CPV empty capsid structure found in the tetragonal crystal form and reported in this paper. The thinner lines represent the CPV full particle structure found in the monoclinic crystal form and reported by Tsao *et al.* (1991). Nucleotides, as they are in full particles, are indicated by the thinnest lines and labeled NA2, NA3 and NA4.


**Figure 10.** Detailed alignment of CPV coat protein with the 3 major viral proteins of HRV14. Residues shown for HRV14 are those for which at least 2 consecutive residues superimpose structurally onto CPV. The conservation of the CPV and the HRV14 amino acid sequence within the parvoviruses and the picornaviruses, respectively, was based on the sequence alignments of parvoviruses (Chapman & Rossmann, 1993) and picornaviruses (Palmenberg, 1989). The parvovirus sequence alignment included 10 parvoviruses, of which 9 were autonomous parvoviruses and 1 was a dependo parvovirus; 6 of the autonomous parvoviruses belong to the RV-like subgroup. The picornavirus alignment included sequences of 6 polioviruses, 3 rhinoviruses, 2 coxsackieviruses (PRC set), 3 hepatitis A viruses, 3 cardioviruses and 3 aphthoviruses. Sequence similarity was defined by the following groupings: V, I, L and M; N and D; Q and E; T, S and V; A and G; Y, F and W; R and K; and L, F and M. For CPV, regular capital letters represent similarity within the 6 RV-like parvoviruses and bold capital letters represent similarity within the 10 aligned parvoviral sequences. One exception was allowed in each case. For HRV14, regular capital letters represent conservation within the PRC set, where there are at least 3 observations of the HRV14 amino acid with others identical or similar, and bold capital letters represent residues conserved among all picornaviruses, where the PRC conservation had to extend to at least 1 other entire family. No more than 3 amino acid types and no more than 2 outliers to the above rule were allowed. The definition of conservation for HRV14 followed that of Arnold & Rossmann (1990).

& Johnson, 1989). The three-dimensional structures of the coat proteins of many viruses can be superimposed remarkably well, suggesting that many viruses may have evolved from a common ancestor (Rossmann & Johnson, 1989). The $C^{\alpha}$ backbone of VP1, VP2 and VP3 of HRV14 (Rossmann *et al.*, 1985; Arnold & Rossmann, 1990) was superimposed (Rao & Rossmann, 1973) onto the CPV structure (Fig. 10). The degree of sequence conservation within parvoviruses (Chapman & Rossmann, 1993) and picornaviruses (Arnold & Rossmann, 1990) is marked on the sequence of CPV and HRV14, respectively, by different letter types. Most of the superimposed residues are within the virus $\beta$-barrel and other secondary structural elements, e.g. $\alpha$A and $\beta$A, that are also present in many other virus

structures (Rossmann & Johnson, 1989). In CPV, the structurally superimposable residues make up a major proportion of residues whose sequences are conserved among parvoviruses.

We thank Drs John Johnson, Janet Smith and Liang Tong for useful discussions and suggestions. We are grateful to Helene Prongay and Sharon Wilder for assistance in the preparation of the manuscript. The work was supported by a Howard Hughes Medical Institute predoctoral fellowship to H.W. and grants from the National Science Foundation and National Institutes of Health to M.G.R. The work also benefited from a Lucille P. Markey Foundation Award for the expansion of structural studies at Purdue University.

## References

Agbandje, M., McKenna, R., Rossmann, M. G., Strassheim, M. L. & Parrish, C. R. (1993). Structure determination of feline panleukopenia virus empty particles. *Proteins*, **16**, 155–171.

Arnold, E. & Rossmann, M. G. (1988). The use of molecular-replacement phases for the refinement of the human rhinovirus 14 structure. *Acta Crystallogr. sect. A*, **44**, 270–282.

Arnold, E. & Rossmann, M. G. (1990). Analysis of the structure of a common cold virus, human rhinovirus 14, refined at a resolution of 3·0 Å. *J. Mol. Biol.* **211**, 763–801.

Barbis, D. P., Chang, S.-F. & Parrish, C. R. (1992). Mutations adjacent to the dimple of the canine parvovirus capsid structure affect sialic acid binding. *Virology*, **191**, 301–308.

Benedetti, E., Morelli, G., Némethy, G. & Scheraga, H. A. (1983). Statistical and energetic analysis of side-chain conformations in oligopeptides. *J. Peptide Protein Res.* **22**, 1–15.

Berns, K. I. (1984). Editor of *The Parvoviruses*. Plenum Press, New York.

Caspar, D. L. D. & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.

Chapman, M. S. & Rossmann, M. G. (1993). Comparison of surface properties of picornaviruses: strategies for hiding the receptor site from immune surveillance. *Virology*, in the press.

Clinton, G. M. & Hayashi, M. (1976). The parvovirus MVM: a comparison of heavy and light particle infectivity and their density conversion *in vitro*. *Virology*, **74**, 57–63.

Cotmore, S. F. & Tattersall, P. (1987). The autonomously replicating parvoviruses of vertebrates. *Advan. Virus Res.* **33**, 91–174.

Fisher, A. & Johnson, J. E. (1993). Ordered duplex RNA controls capsid architecture in an icosahedral animal virus. *Nature (London)*, **361**, 176–179.

Fry, E., Logan, D., Acharya, R., Fox, G., Rowlands, D., Brown, F. & Stuart, D. (1990). Architecture and topography of an aphthovirus. *Sem. Virol.* **1**, 439–451.

Giranda, V. L., Heinz, B. A., Oliveira, M. A., Minor, I., Kim, K. H., Kolatkar, P. R., Rossmann, M. G. & Rueckert, R. R. (1992). Acid-induced structural changes in human rhinovirus 14: possible role in uncoating. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 10213–10217.

Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. (1978). Tomato bushy stunt virus at 2·9 Å resolution. *Nature (London)*, **276**, 368–373.

Hendrickson, W. A. (1985). Stereochemically restrained refinement of macromolecular structures. *Methods Enzymol.* **115**, 252–270.

Hendrickson, W. A. & Konnert, J. H. (1980). Incorporation of stereochemical information into crystallographic refinement. In *Computing in Crystallography* (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds), pp. 13.01–13.23, Indian Academy of Sciences, Bangalore.

Huber, R. & Bennett, W. S., Jr (1983). Functional significance of flexibility in proteins. *Biopolymers*, **22**, 261–279.

Huber, R. & Bode, W. (1978). Structure basis of the activation and action of trypsin. *Acct. Chem. Res.* **11**, 114–122.

Janin, J., Wodak, S., Levitt, M. & Maigret, B. (1978). Conformation of amino acid side-chains in proteins. *J. Mol. Biol.* **125**, 357–386.

Jones, T. A. & Thirup, S. (1986). Using known substructures in protein model building and crystallography. *EMBO J.* **5**, 819–822.

Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.

Kim, S., Boege, U., Krishnaswamy, S., Minor, I., Smith, T. J., Luo, M., Scraba, D. G. & Rossmann, M. G. (1990). Conformational variability of a picornavirus capsid: pH-dependent structural changes of Mengo virus related to its host receptor attachment site and disassembly. *Virology*, **175**, 176–190.

Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.

Krishnaswamy, S. & Rossmann, M. G. (1990). Structural refinement and analysis of Mengo virus. *J. Mol. Biol.* **211**, 803–844.

Lonberg-Holm, K., Gosser, L. B. & Shimshick, E. J. (1976). Interaction of liposomes with subviral particles of poliovirus type 2 and rhinovirus type 2'. *J. Virol.* **19**, 746–749.

McKenna, R., Xia, D., Willingmann, P., Ilag, L. L., Krishnaswamy, S., Rossmann, M. G., Olson, N. H., Baker, T. S. & Incardona, N. L. (1992). Atomic structure of single-stranded DNA bacteriophage $\phi$X174 and its functional implications. *Nature (London)*, **355**, 137–143.

Palmenberg, A. C. (1989). Sequence alignments of picornaviral capsid proteins. In *Molecular Aspects of Picornavirus Infection and Detection* (Semler, B. L. & Ehrenfeld, E., eds), pp. 211–241, American Society for Microbiology, Washington, DC.

Paradiso, P. R., Rhode, S. L., III & Singer, I. I. (1982). Canine parvovirus: a biochemical and ultrastructural characterization. *J. Gen. Virol.* **62**, 113–125.

Parrish, C. R., Have, P., Foreyt, W. J., Evermann, J. F., Senda, M. & Carmichael, L. E. (1988). The global spread and replacement of canine parvovirus strains. *J. Gen. Virol.* **69**, 1111–1116.

Parrish, C. R., Aquadro, C. F., Strassheim, M. L., Evermann, J. F., Sgro, J.-Y. & Mohammed, H. O. (1991). Rapid antigenic-type replacement and DNA sequence evolution of canine parvovirus. *J. Virol.* **65**, 6544–6552.

Ponder, J. W. & Richards, F. M. (1987). Tertiary templates for proteins. Use of packing criteria in the

enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* **193**, 775–791.

Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **23**, 283–438.

Rao, S. T. & Rossmann, M. G. (1973). Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **76**, 241–256.

Rhode, S. L., III (1985). Nucleotide sequence of the coat protein gene of canine parvovirus. *J. Virol.* **54**, 630–633.

Richards, F. M. (1977). Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176.

Richardson, J. S., Getzoff, E. D. & Richardson, D. C. (1978). The $\beta$ bulge: a common small unit of non-repetitive protein structure. *Proc. Nat. Acad. Sci., U.S.A.* **75**, 2574–2578.

Roberts, M. M., White, J. L., Grütter, M. G. & Burnett, R. M. (1986). Three-dimensional structure of the adenovirus major coat protein hexon. *Science*, **232**, 1148–1151.

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

Rossmann, M. G., Abad-Zapatero, C., Hermodson, M. A. & Erickson, J. W. (1983). Subunit interactions in southern bean mosaic virus. *J. Mol. Biol.* **166**, 37–83.

Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B. & Vriend, G. (1985). Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature (London)*, **317**, 145–153.

Ruoslahti, E. A. (1988). Fibronectin and its receptors. *Annu. Rev. Biochem.* **57**, 375–413.

Saliki, J. T., Mizak, B., Flore, H. P., Gettig, R. R., Burand, J. P., Carmichael, L. E., Wood, H. A. & Parrish, C. R. (1992). Canine parvovirus empty capsids produced by expression in a baculovirus vector: use in analysis of viral properties and immunization of dogs. *J. Gen. Virol.* **73**, 369–374.

Sheriff, S., Hendrickson, W. A., Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1985). Influence of solvent accessibility and intermolecular contacts on atomic mobilities in hemerythrins. *Proc. Nat. Acad. Sci., U.S.A.* **82**, 1104–1107.

Shrake, A. & Rupley, J. A. (1973). Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* **79**, 351–371.

Silva, A. M. & Rossmann, M. G. (1987). Refined structure of southern bean mosaic virus at 2·9 Å resolution. *J. Mol. Biol.* **197**, 69–87.

Stewart, P. L., Burnett, R. M., Cyrklaff, M. & Fuller, S. D. (1991). Image reconstruction reveals the complex molecular organization of adenovirus. *Cell*, **67**, 145–154.

Tainer, J. A., Getzoff, E. D., Alexander, H., Houghten, R. A., Olson, A. J., Lerner, R. A. & Hendrickson, W. A. (1984). The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. *Nature (London)*, **312**, 127–134.

Tattersall, P. (1972). Replication of the parvovirus MVM. I. Dependence of virus multiplication and plaque formation on cell growth. *J. Virol.* **10**, 586–590.

Tattersall, P. &. Cotmore, S. F. (1988). The nature of parvoviruses. In *Parvoviruses and Human Disease* (Pattison, J. R., ed.), pp. 5–41, CRC Press, Boca Raton, FL.

Tattersall, P., Shatkin, A. J. & Ward, D. C. (1977). Sequence homology between the structural polypeptides of minute virus of mice. *J. Mol. Biol.* **111**, 375–394.

Tsao, J., Chapman, M. S., Agbandje, M., Keller, W., Smith, K., Wu, H., Luo, M., Smith, T. J., Rossmann, M. G., Compans, R. W. & Parrish, C. R. (1991). The three-dimensional structure of canine parvovirus and its functional implications. *Science*, **251**, 1456–1464.

Tsao, J., Chapman, M. S., Wu, H., Agbandje, M., Keller, W. & Rossmann, M. G. (1992). Structure determination of monoclinic canine parvovirus. *Acta Crystallogr. sect. B*, **48**, 75–88.

Wu, H., Keller, W. & Rossmann, M. G. (1993). Determination and refinement of the canine parvovirus empty capsid structure. *Acta Crystallogr. sect. D*, in the press.

*Edited by R. Huber*