

24. A. L. Schechter *et al.*, *Science* **229**, 976 (1985).
25. S. I. Fukushige *et al.*, *Mol. Cell. Biol.* **6**, 955 (1985).
26. C. R. King, M. H. Kraus, S. A. Aaronson, *Science* **229**, 974 (1985).
27. D. Meerakkan and M. Smith, *Cytogenet. Cell. Genet.* **37**, 71 (1984).
28. A. Ullrich *et al.*, *Nature (London)* **309**, 418 (1984).
29. G. M. Clark *et al.*, *N. Engl. J. Med.* **309**, 1343 (1983).
30. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual*, (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982), pp. 282–285.
31. G. J. Dezikes, W. W. Grody, R. M. Kern, S. D. Cedarbaum, *Biochem. Biophys. Res. Commun.*, in press.
32. R. C. Seeger *et al.*, *N. Engl. J. Med.* **313**, 1111 (1985).
33. J. Yokota *et al.*, *Science* **231**, 261 (1986).
34. C. Escot *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4834 (1986).
35. M. Schwab *et al.*, *Nature (London)* **305**, 245 (1983).
36. N. E. Kohl *et al.*, *Cell* **35**, 359 (1984).
37. R. C. Seeger and D. Slamon, unpublished data.
38. E. Silverberg and J. Lubera, *CA Cancer J. Clin.* **36**, 9 (1986).
39. K. B. Horwitz, W. L. McGuire, O. H. Pearson, A. Segaloff, *Science* **189**, 726 (1975).
40. W. L. McGuire, *Semin. Oncol.* **5**, 428 (1978).
41. C. I. Bargmann, M. C. Hung, R. A. Weinberg, *Cell* **45**, 649 (1986).
42. D. R. Cox, *J. R. Stat. Soc.* **34**, 187 (1972).
43. N. E. Breslow, *Int. Stat. Rev.* **43**, 45 (1975).
44. E. L. Kaplan and P. Meier, *J. Am. Stat. Assoc.* **53**, 457 (1958).
45. Supported by U.S. Public Health Service grants CA 36827 and CA 30195 and a grant from Triton Biosciences, Inc. We thank D. Keith, L. Gordon, and W. Aft for their technical assistance and Dr. S. Cedarbaum for providing the human arginase probe.

## Research Articles

# The Atomic Structure of Mengo Virus at 3.0 Å Resolution

MING LUO, GERRIT VRIEND\*, GREG KAMER, IWONA MINOR, EDWARD ARNOLD,  
MICHAEL G. ROSSMANN, ULRIKE BOEGE, DOUGLAS G. SCRABA, GREG M. DUKE,  
ANN C. PALMENBERG

The structure of Mengo virus, a representative member of the cardio picornaviruses, is substantially different from the structures of rhino- and polioviruses. The structure of Mengo virus was solved with the use of human rhinovirus 14 as an 8 Å resolution structural approximation. Phase information was then extended to 3 Å resolution by use of the icosahedral symmetry. This procedure gives promise that many other virus structures also can be determined without the use of the isomorphous replacement technique. Although the organization of the major capsid proteins VP1, VP2, and VP3 of Mengo virus is essentially the same as in rhino- and polioviruses, large insertions

and deletions, mostly in VP1, radically alter the surface features. In particular, the putative receptor binding "canyon" of human rhinovirus 14 becomes a deep "pit" in Mengo virus because of polypeptide insertions in VP1 that fill part of the canyon. The minor capsid peptide, VP4, is completely internal in Mengo virus, but its association with the other capsid proteins is substantially different from that in rhino- or poliovirus. However, its carboxyl terminus is located at a position similar to that in human rhinovirus 14 and poliovirus, suggesting the same autocatalytic cleavage of VP0 to VP4 and VP2 takes place during assembly in all these picornaviruses.

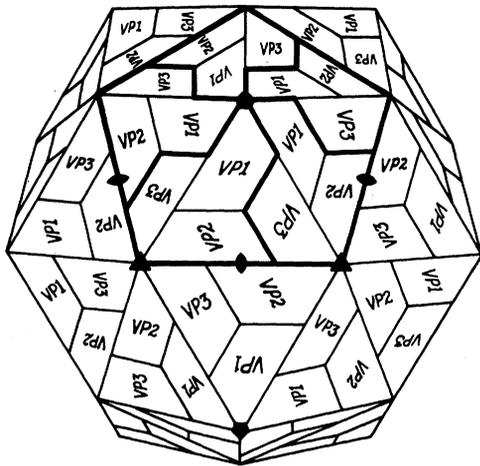
**P**ICORNAVIRUSES (1) COMPRISE A LARGE FAMILY OF SINGLE-STRANDED RNA-CONTAINING animal viruses. They have been classified into four genera on the basis of physical properties of the virions: (i) cardioviruses, such as encephalomyocarditis virus (EMCV), Theiler's murine encephalitis virus, and Mengo virus; (ii) enteroviruses, such as poliovirus, hepatitis A virus, and Coxsackie virus; (iii) aphthoviruses, such as foot-and-mouth disease viruses (FMDV); and (iv) rhinoviruses, of which there are about 100 serotypes. The physiological consequences of picornavirus infection are highly variable because of differing tissue tropism, differing viral stability in acid environment, and differing responses by infected host cells. Among the murine cardioviruses (2), for example, Mengo virus can produce fatal encephalitis, a strain of EMCV can cause diabetes, and a strain of Theiler's virus can cause chronic demyelinating disease.

Picornavirions have relative molecular mass of about  $8.5 \times 10^6$  daltons, of which about 30 percent is RNA. The virus particles are spherical with an external diameter of about 300 Å. The coat proteins form a shell with icosahedral symmetry with 60 protomers,

each composed of a single copy of VP1, VP2, VP3, and VP4 (Fig. 1) with approximate relative molecular masses of 34,000, 30,000, 25,000, and 7,000 daltons, respectively. For Mengo virus, the proteins consist of 277, 256, 231, and 70 amino acids, respectively. On degrading Mengo virus, the virions separate into 13.4S pentamers and then into 6S protomers (3, 4) consistent with structural observations (5). Traditionally, the cardiovirus capsid proteins have been designated  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$  in order of decreasing size; the VP1 ( $\alpha$ ), VP2 ( $\beta$ ), VP3 ( $\gamma$ ), VP4 ( $\delta$ ) nomenclature is used in this article in order to facilitate comparison with other picornaviruses. The viral

M. Luo, G. Vriend, G. Kamer, I. Minor, E. Arnold, and M. G. Rossmann, who did the crystallographic analysis, are in the Department of Biological Sciences, Purdue University, West Lafayette, IN 47907. U. Boege and D. G. Scraba, who produced and purified the 60 milligrams of virus used, are in the Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada. G. M. Duke and A. C. Palmenberg, who determined the sequence of the viral RNA, are in the Biophysical Laboratory of the Graduate School, University of Wisconsin, Madison, WI 53706.

\*Present address: Department of Structural Chemistry, University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands.



**Fig. 1.** Diagrammatic representation of an icosahedral capsid in picornaviruses. The thickly outlined VP1, VP3, and VP2 units correspond to the 6S (VP1, VP3, VP0) protomer and the pentadecamer cap to the 14S pentamer observed in assembly experiments (4).

genome is a single strand of messenger sense RNA (7200 to 8400 bases), which is translated into a single polyprotein by the host protein synthesis machinery and then processed by stepwise proteolysis into its component proteins with the use of virally coded enzymes. During the last step of the processing cascade, most of the VP0 precursors are cleaved into VP4 and VP2 by a mechanism that is probably autocatalytic during completion of virion assembly (5–7). The VP0 cleavage may be necessary for yielding infectious particles. Examination of nucleotide sequences indicates that gene order and peptide arrangements within picornaviral polyproteins are very similar among these viruses. Comparison and alignments confirm a high degree of amino acid homology, especially from viruses of the same genus. For example, the proteins of EMCV (8) and Mengo virus (9) are more than 95 percent identical, and EMCV and Theiler's virus (10) share 80 to 85 percent identity (11).

The atomic structures of several icosahedral insect (12), plant (13, 14), and animal RNA viruses, including human rhinovirus 14 (HRV14) (5) and polio (type 1, Mahoney) virus (15), have now been determined. These share common structural motifs. The individual capsid proteins form eight-stranded antiparallel  $\beta$ -barrels of similar topology. The quaternary structural organization is similar among all these viruses (Fig. 1) with  $T = 3$  (16) or pseudo  $T = 3$  (5) surface lattices. The RNA structure has not been visualized in the crystallographic studies.

The most external and antigenic capsid protein of picornaviruses is VP1 [for Mengo virus, see (17)], while VP4 is entirely internal (18–20). Chemical labeling and cross-linking studies (particularly on Mengo virus) (21, 22), as well as investigation of the dissociation of Mengo virus mediated by chloride ions at pH 6.2 (3, 23), showed that the protein subunits in picornaviruses are arranged essentially as indicated in Fig. 1. These observations were subsequently confirmed by three-dimensional crystallographic investigations on HRV14 (5) and poliovirus type 1 (Mahoney strain) (15).

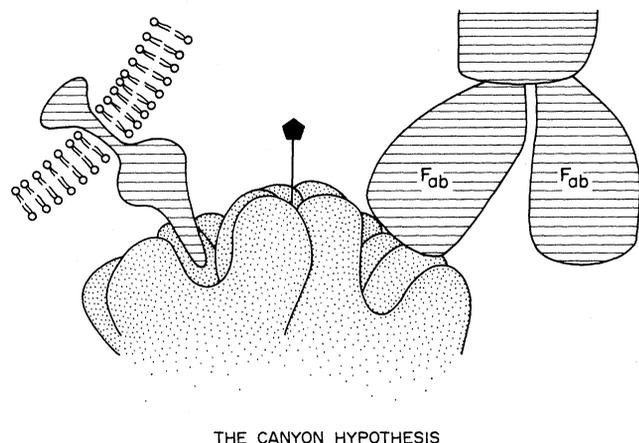
Picornaviruses initiate entry into host cells by attaching to receptors on the host cell membrane. The attachment site on FMDV has been shown to be part of VP1 (24, 25). Receptors for poliovirus (26) and human rhinoviruses (27) have been detected by monoclonal antibodies that inhibit the initial virus-cell interaction, and receptors have been isolated for Coxsackie B3 (28, 29) and the major human rhinovirus group (30). These receptors are glycoproteins with molecular masses ranging from 54,000 to 90,000 daltons. Glycophorin A on erythrocyte membranes is the site of EMCV attachment during hemagglutination and, therefore, provides a

useful model system for virus-receptor interaction (31, 32). A deep "canyon" circulating about each fivefold axis on HRV14, composed primarily of residues from VP1, has been recognized as a putative receptor binding site. It has been hypothesized (5) that receptors, but not antibodies, would be able to bind within the canyon. Hypervariable immunogenic residues on the canyon rim permit the virion to escape the host cell immune system while a constant receptor site remains inaccessible to antibodies (Fig. 2).

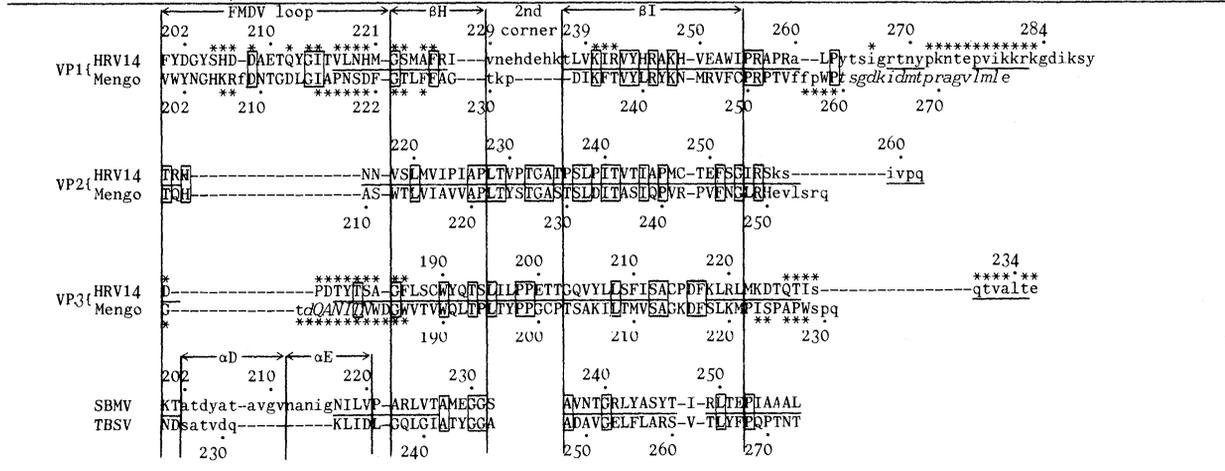
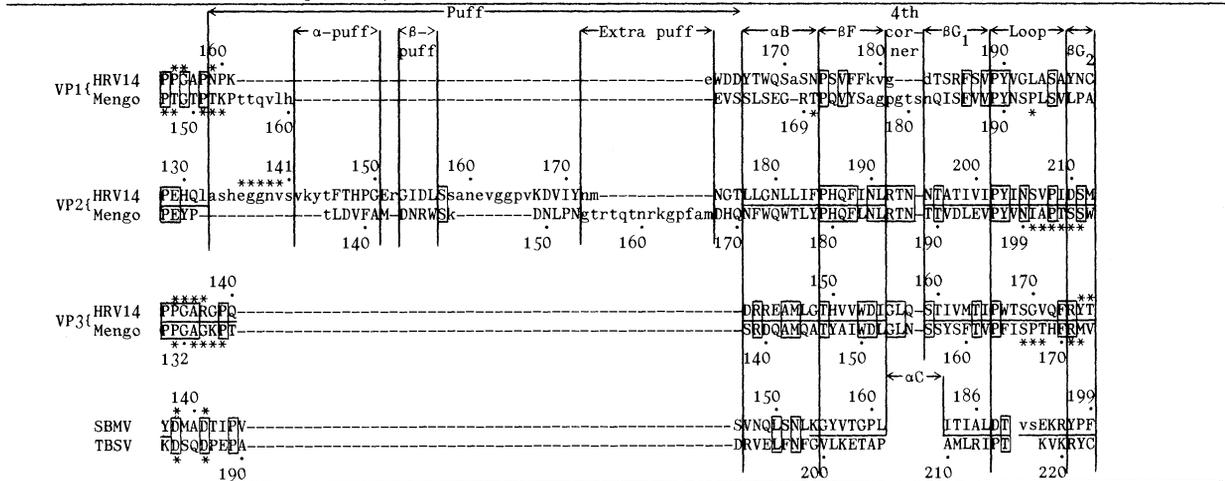
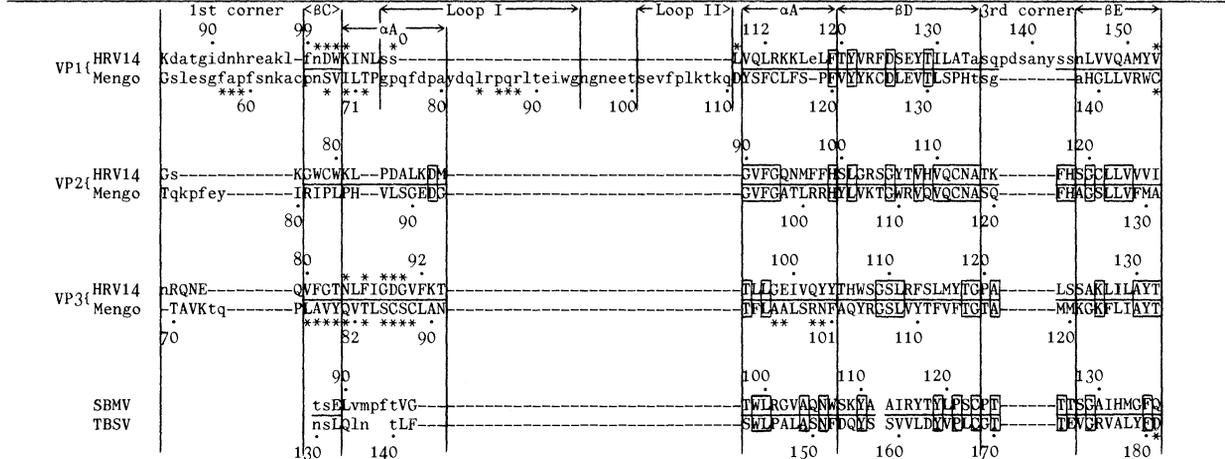
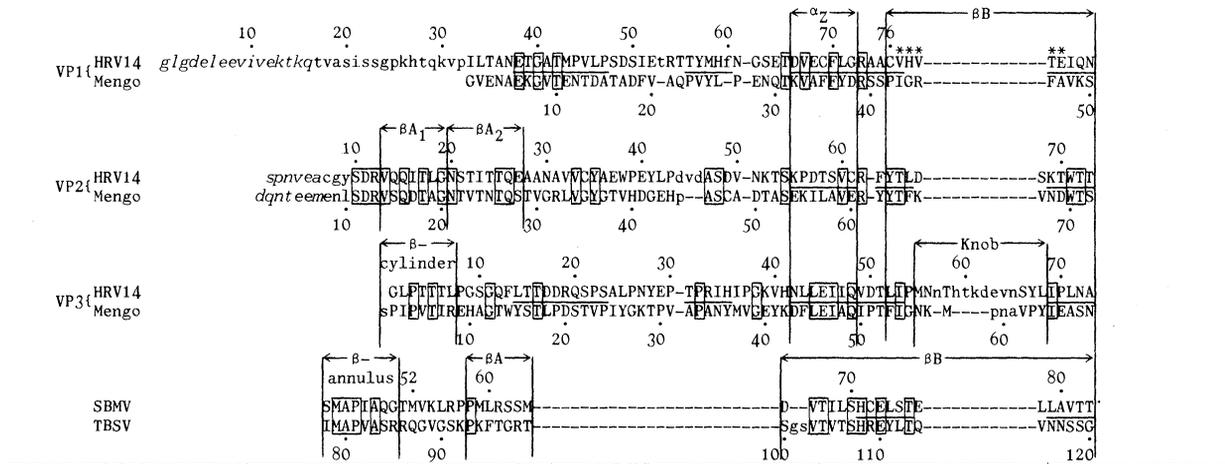
The first Mengo virus crystals obtained had cubic symmetry with  $a = 422 \text{ \AA}$  (33). It was possible to collect x-ray diffraction data to 7  $\text{\AA}$  resolution for these crystals. A comparison of the three-dimensional diffraction data from HRV14 and Mengo virus crystals indicated that their structures might be similar (34). Subsequent sequence comparisons of the two viruses supported these results (9).

An orthorhombic crystal form of Mengo virus was found later which diffracted x-rays to 2.5  $\text{\AA}$  resolution. The three-dimensional structure of Mengo virus, reported here, was determined with the use of these crystals, knowledge of the HRV14 structure (5), the molecular replacement technique (35, 36), and the Mengo virus capsid protein amino acid sequence (Fig. 3). The latter (37) was derived from the RNA sequence. The cleavage sites were determined by direct sequencing of the capsid proteins and alignment with EMCV and other picornavirus sequences (9).

**Structure determination.** The M variant of Mengo virus (38, 39) was propagated and purified as described (33). The orthorhombic crystals were prepared by the hanging drop vapor diffusion method with 2.8 percent polyethylene glycol 8000 in 0.1M phosphate buffer at pH 7.4 in the reservoir with an initial virus concentration of 5 mg/ml and 1.4 percent polyethylene glycol 8000 in the same buffer in the hanging drop. The crystals grew in 1 to 2 days at room temperature to a maximum dimension of 0.8 mm. They were extremely sensitive to x-rays and lasted only about 30 minutes in x-rays from a rotating anode source. However, high-quality oscillation photographs were readily obtained at the Cornell High Energy Synchrotron Source (CHESS). A fresh crystal, maintained at 4°C, was required for each 1-minute exposure of a 0.3° oscillation photograph. A total of 600 film packs was collected with crystals oriented only by eye and an optical microscope. The oscillation films collected in this way had large crystal setting errors (up to 45°), and required the development of new techniques for the determination



**Fig. 2.** The presence of depressions on the picornavirus surface suggests a strategy for the evasion of immune surveillance. The dimensions of the putative receptor binding sites (the canyon in HRV14, the pit in Mengo) sterically hinder an antibody's (top right) recognition of residues at the base of the site, while still allowing recognition and binding by a smaller cellular receptor (top left). This would allow for receptor specificity while at the same time permitting evolution of new serotypes by mutating residues about the rim of the canyon or pit.



of the crystal setting matrix (40). About 220 films were processed, of which 143 films were included in the final data set for the structure determination (Table 1).

“Still” photographs showed that the three principal axes were orthogonal, that the lattice was primitive, and that the cell dimensions were  $a \approx b \approx c \approx 450$  Å. Intensity measurements and “post-refinement” of the oscillation photographs showed that the Laue group was *mmm*, and the cell dimensions were (Å):  $a = 441.4$ ,  $b = 427.3$ , and  $c = 421.9$ . A more complete data set showed that there were systematic absences along each of the axial directions. Thus, the space group was established as  $P2_12_12_1$  with four particles per cell or one particle per asymmetric unit. There is a 60-fold noncrystallographic redundancy because of the icosahedral symmetry of each particle. The orientation of the virus particles in the unit cell was determined with a self-rotation function (41) and refined by fitting a standard icosahedron to the rotation function peaks. One of the icosahedral twofold axes was found to be approximately parallel to the  $b$  axis of the unit cell, generating a large Harker peak at  $(1/2, 1/2, 1/2)$  in a low resolution (13 Å) Patterson. The center of the particle was, therefore, close to  $(1/4, 1/4, 1/4)$ , using the unit cell origin as defined in the *International Tables for Crystallography* (42). A more precise particle position was determined in a translational search around  $(1/4, 1/4, 1/4)$  using the known HRV14 atomic structure as an approximate model. Calculation of the  $R$  factor as a function of resolution showed that the resultant phases based on HRV14 could be accepted to only 8 Å resolution. The  $R$  factor had a minimum (46 percent) at 12 Å resolution. At very low resolution the  $R$  factor was large (69 percent) due to poor solvent modeling, and beyond 8 Å resolution the  $R$  factor assumed a random value ( $R > 54$  percent).

Ten cycles of molecular replacement real-space averaging (35, 43, 44) at 8 Å resolution, with the use of the icosahedral symmetry, increased the correlation coefficient from 0.54 to 0.87 and decreased the  $R$  factor from 54 to 23 percent. Features of the resultant electron density map were consistent with expectations based on sequence alignments (for example, deletion of part of the “puff” in VP2). Phases from 8 to 3 Å were then determined by gradually increasing the resolution in steps of one or two reciprocal lattice points at a time. One or two cycles of phase improvement were performed at each resolution step, although 12 cycles were performed at 3.5 Å resolution prior to the detailed interpretation of the electron density map (Fig. 4). Missing structure amplitude observations were substituted with calculated structure factors from each cycle of Fourier back-transformation of the averaged density (45). (These calculated amplitudes cannot be included in the evaluation of correlation coefficients and  $R$  factors shown in Fig. 4.)

The electron density map of Mengo virus has a quality comparable to that of HRV14 (Fig. 5). The gross polypeptide chain tracing was established at 6 Å resolution, based on knowledge of the HRV14 structure. The detailed chain tracing and amino acid identification were rapidly obtained with a mini-map at 3.5 Å

**Fig. 3.** Alignment of Mengo virus and HRV14 sequences for VP1, VP2, and VP3. Shown also is the alignment of major structural segments in the shell domains of SBMV and TBSV with Mengo virus and HRV14. Horizontal bars indicate regions of structural alignments among VP1, VP2, VP3, and the two plant virus coat proteins. Sequences given in lower case letters have no structural equivalences. Deletions are shown with – relative to other sequences. A blank in the plant viruses indicates that a portion of the sequence has been omitted from the alignments. An asterisk (\*) indicates residues in the pit (Mengo) or canyon (HRV14) lining or ligands for  $\text{Ca}^{2+}$  binding (SBMV and TBSV). Sequences given in italic letters are disordered in the x-ray structures. Sequences that are identical between HRV14 and Mengo virus and between SBMV and TBSV have been boxed. Alignment of sequences that have no structural equivalence is somewhat arbitrary.

**Table 1.** Data used in structure determination. Number of unique observations ( $F^2 > 1\sigma(F^2)$ ) and percentage of possible total.

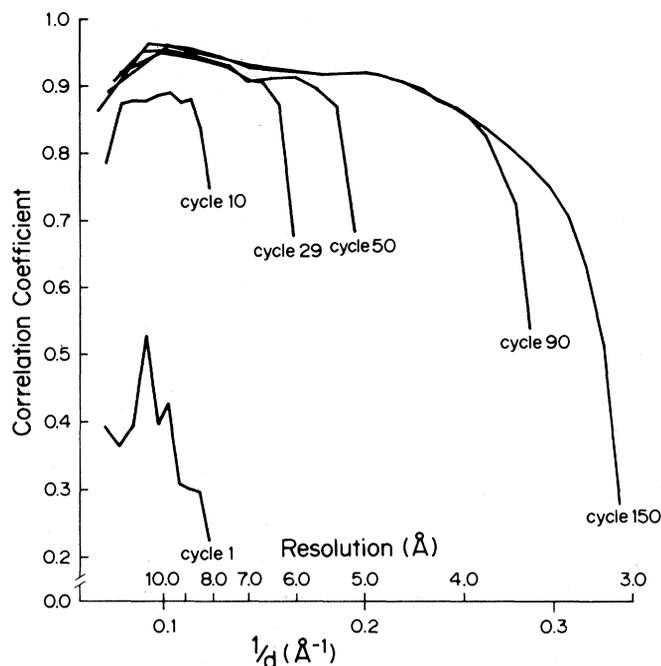
Resolution range (Å)	Number observed	Percentage of possible
∞ to 30	837	54
30 to 15	7,684	71
15 to 10	21,239	72
10 to 7.5	40,196	70
7.5 to 5.0	157,716	67
5.0 to 3.5	396,425	62
3.5 to 3.0	263,549	46
3.0 to 2.75	87,801	19
2.75 to 2.6	19,634	3
Total	995,081	42
Number of film packs	143	
$R$ factor (%)*	11.21	

$$*R = \frac{\sum_b \sum_i |I_h - I_{hi}|}{\sum_b \sum_i I_{hi}} \times 100$$

where  $I_h$  is the mean of the  $I_{hi}$  observations of reflection  $h$ . The missing data are mostly unobserved, since a full data set would require about 300 films with a  $0.3^\circ$  oscillation angle. The films used in the data were chosen randomly for their quality.

resolution. An atomic model was built into the 3.0 Å resolution map with the use of an Evans and Sutherland PS300 computer graphics system and the FRODO program (46). The consistency of the electron density map with the known amino acid sequence showed that there was no bias from the initial 8 Å resolution phasing model based on the HRV14 structure.

Extension of the phases from 8 to 3.0 Å represents the most comprehensive use of the molecular replacement method to date. Previous applications of phase extension at higher resolutions include the results on HRV14 and poliovirus (5 to 3.0 Å) and on

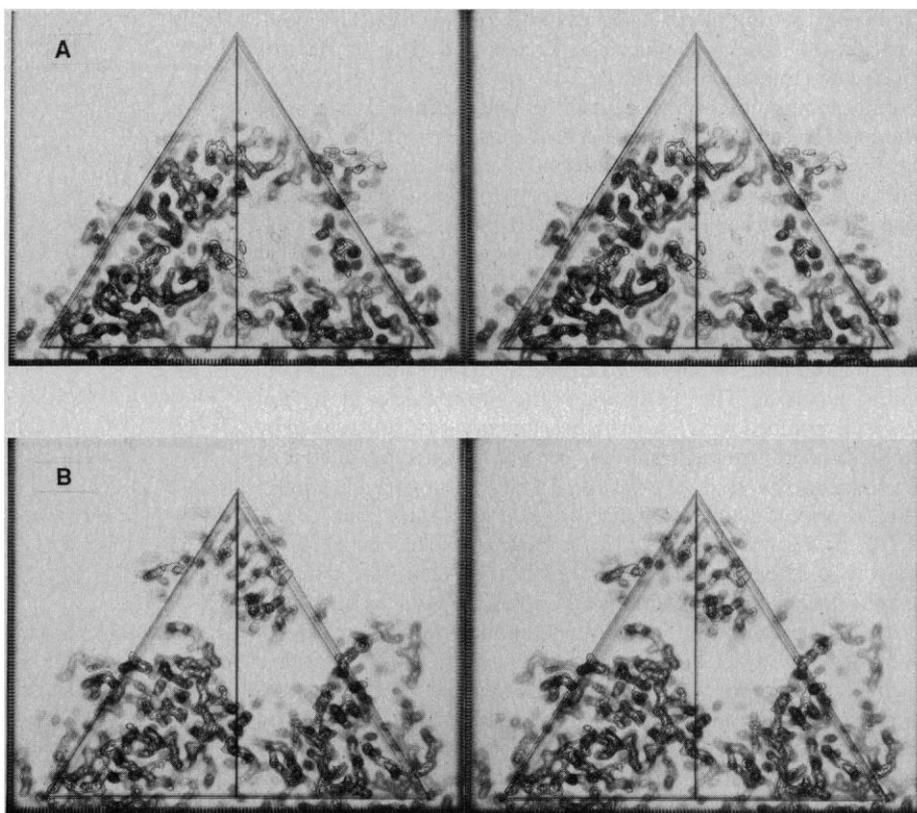


**Fig. 4.** Correlation coefficients,  $C$ , for the phase extension from 8 to 3 Å resolution. The plot shows the last refinement cycle at a given resolution. The definition for  $C$  is:

$$C = \frac{\sum_b (F_{h,\text{calc}} - \langle F_{\text{calc}} \rangle)(F_{h,\text{obs}} - \langle F_{\text{obs}} \rangle)}{[\sum_b (F_{h,\text{calc}} - \langle F_{\text{calc}} \rangle)^2 \sum_b (F_{h,\text{obs}} - \langle F_{\text{obs}} \rangle)^2]^{1/2}}$$

Here  $F_{h,\text{calc}}$  and  $F_{h,\text{obs}}$  are the calculated and observed structure factors of reflection  $h$ , and  $\langle \rangle$  indicates their mean in local resolution shells.

**Fig. 5.** Electron density sections, for (A) Mengo virus and (B) HRV14, cut perpendicular to an icosahedral twofold axis between 144 Å and 135 Å from the particle center. The triangle has an icosahedral fivefold at the top and icosahedral threefolds at the sides, thus covering one icosahedral asymmetric unit. Note (A) the deep external pit in Mengo virus and (B) the canyon in HRV14.



hemocyanin (4.0 to 3.2 Å) (47). The redundancy available for the Mengo virus structure determination was 60-fold, whereas it was only 20-fold for HRV14, 30-fold for poliovirus, and 6-fold for hemocyanin. The power of the method can be shown to be proportional to the square root of the redundancy (43, 48).

The structure of Mengo virus was solved without the use of conventional isomorphous replacement techniques, with only the atomic structure of a related virus being used as a starting model. The success with Mengo virus suggests that *ab initio* determinations for virus structures are likely to be successful even when there is no available information on homologous structure. Indeed, low-resolution *ab initio* determinations have been performed for southern bean mosaic virus (SBMV) (49) and polyoma virus (50).

**Structure.** The overall structure of Mengo virus resembles HRV14 and poliovirus, although the latter are far more alike to each other than they are to Mengo virus, as is consistent with sequence comparisons (9). The largest differences occur in the surface features of VP1 and the relocation of VP4. Each of the three major capsid proteins VP1, VP2, and VP3 consist of an eight-stranded antiparallel  $\beta$ -barrel (Fig. 6) as has been found in all other known spherical RNA viruses (5, 12, 13, 51). These  $\beta$ -barrels are wedge-shaped with the thin ends pointing to the fivefold or pseudo sixfold axis in each case. The thin ends consist of four turns in the polypeptide chain with the outside top corner being the first along the chain between  $\beta$ B and  $\beta$ C, the second corner down from the outside is the last along the polypeptide chain between  $\beta$ H and  $\beta$ I, etc. (Fig. 6). The  $\beta$ -barrels have essentially the same orientation and radius as those in HRV14 (Table 2) and are related by a pseudo threefold axis. The 40 to 70 residues forming the amino terminal arms of VP1 and VP3 as well as all of VP4 intertwine themselves on the interior surface of the  $\beta$ -barrels, indicating the nature of the 6S protomeric assembly unit (3) (Figs. 1 and 7). The amino ends of VP3 form a  $\beta$  cylinder (Fig. 8) around the fivefold axis, associating five protomers corresponding to the 14S pentameric assembly units

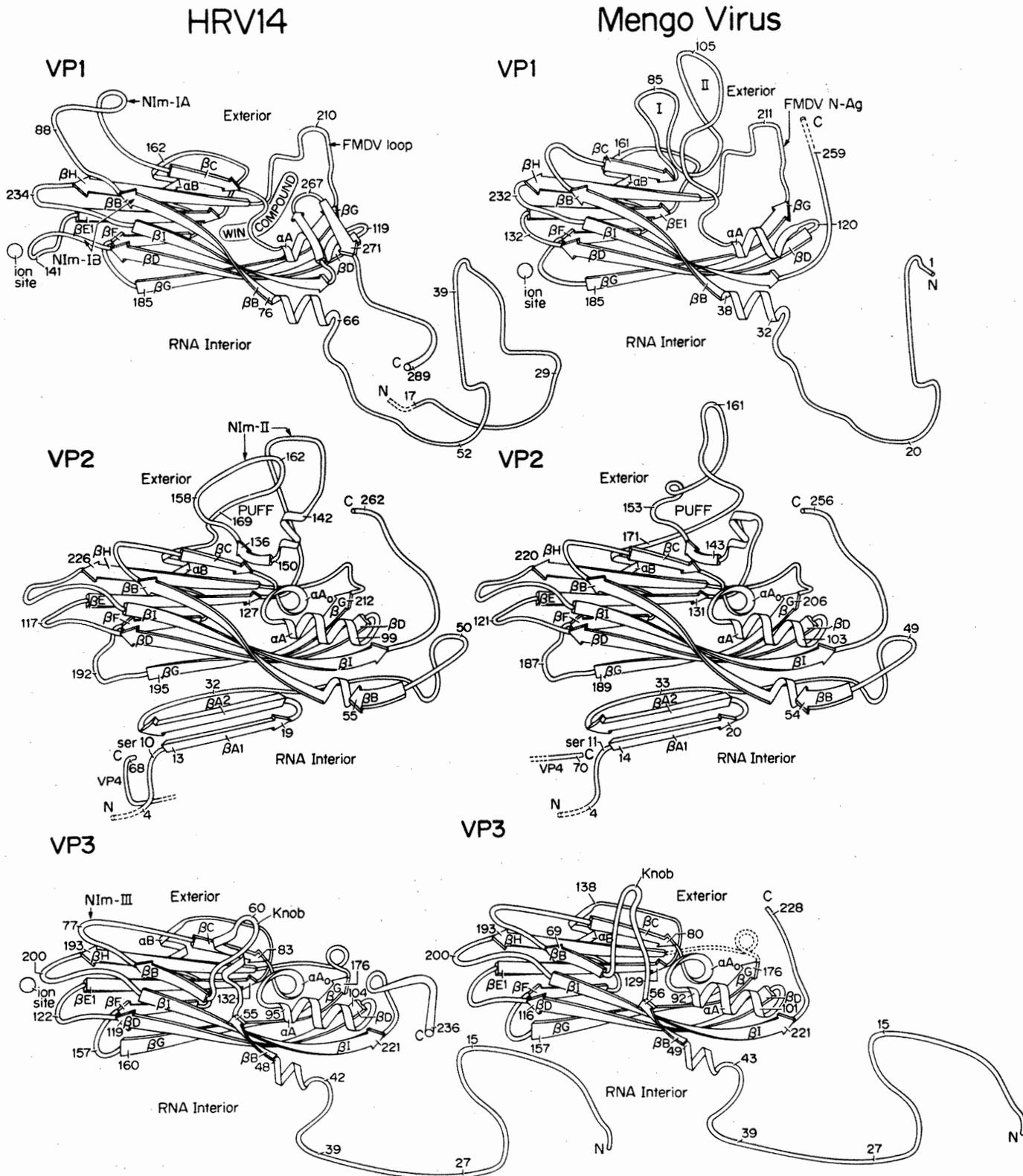
(Figs. 1 and 7) (3, 4). The five parallel  $\beta$  strands in the cylinder have a right-handed twist.

The pentameric clusters of VP1 comprise an even larger proportion of the surface area in Mengo virus than in HRV14 (Fig. 9) because of a series of insertions in VP1 and deletions in VP2 and VP3 (Fig. 6). The amino terminal arm of VP1 is ordered, but is shorter by 32 residues than VP1 of HRV14. The top corner connecting  $\beta$ B and  $\beta$ C, equivalent to a neutralizing immunogen site (NIm-IA) in HRV14, is foreshortened in Mengo virus. Two protrusions, "loop I" and "loop II," are inserted in Mengo virus at the carboxyl end of  $\beta$ C in VP1, and fill the narrow part of the canyon close to the VP2 puff present in HRV14. Loop II is the most external peptide in Mengo virus, and five of the most external residues (96 to 100) show signs of disorder. Mengo and HRV14 show greater similarity between  $\alpha$ A and the carboxyl terminus of VP1. Like the top corner, the third corner down (between  $\beta$ D and  $\beta$ E) and the second corner down (between  $\beta$ H and  $\beta$ I) are much shorter in Mengo virus than in HRV14 VP1. As a result, there is a shallow depression (as opposed to the HRV14 protrusion) at the fivefold vertices. There is an electron density peak on the fivefold axis. The stereochemistry is consistent with this being a cationic ligand chelated by the carbonyl oxygens of the fivefold-related Pro<sup>178</sup> residues on the fourth corner down. A similar electron density peak is observed in HRV14 associated with Asn<sup>141</sup> on the third corner down. The last carboxyl terminal residue of VP1 which can be seen in the electron density map is Pro<sup>259</sup>. The remaining 15 to 18 residues (see below) are disordered and could form a mobile region near the rim of the putative receptor binding "pit" or are missing. Ziola and Scraba (52) determined that the carboxyl terminal sequence of VP1 in purified virions is -Ala-Gly-Val-Leu<sup>274</sup>. However, the carboxyl terminus of VP1 of Mengo virus and EMCV produced and isolated from rabbit reticulocyte lysates was determined as -Val-Leu-Met-Leu-Glu<sup>277</sup> (9), which would shift the cleavage by three amino acids further along the polypeptide. This

might indicate trimming of three residues *in vivo* after translation or after assembly.

The structures of Mengo virus VP2 and VP3 have a far greater similarity to HRV14 VP2 and VP3 than do the structures of VP1

in Mengo virus and HRV14. The first ten residues of Mengo VP2 (Fig. 6) are disordered. The Ser<sup>10</sup> residue in VP2 of HRV14 was identified as a possible nucleophile required for the autocatalytic cleavage of VP0 into VP4 and VP2 during the encapsidation of



**Fig. 6.** Ribbon drawings of the three larger viral proteins VP1, VP2, and VP3 for Mengo virus and HRV14. Each protein has the same wedge-shaped, eight-stranded, antiparallel  $\beta$ -barrel but differs because of insertions

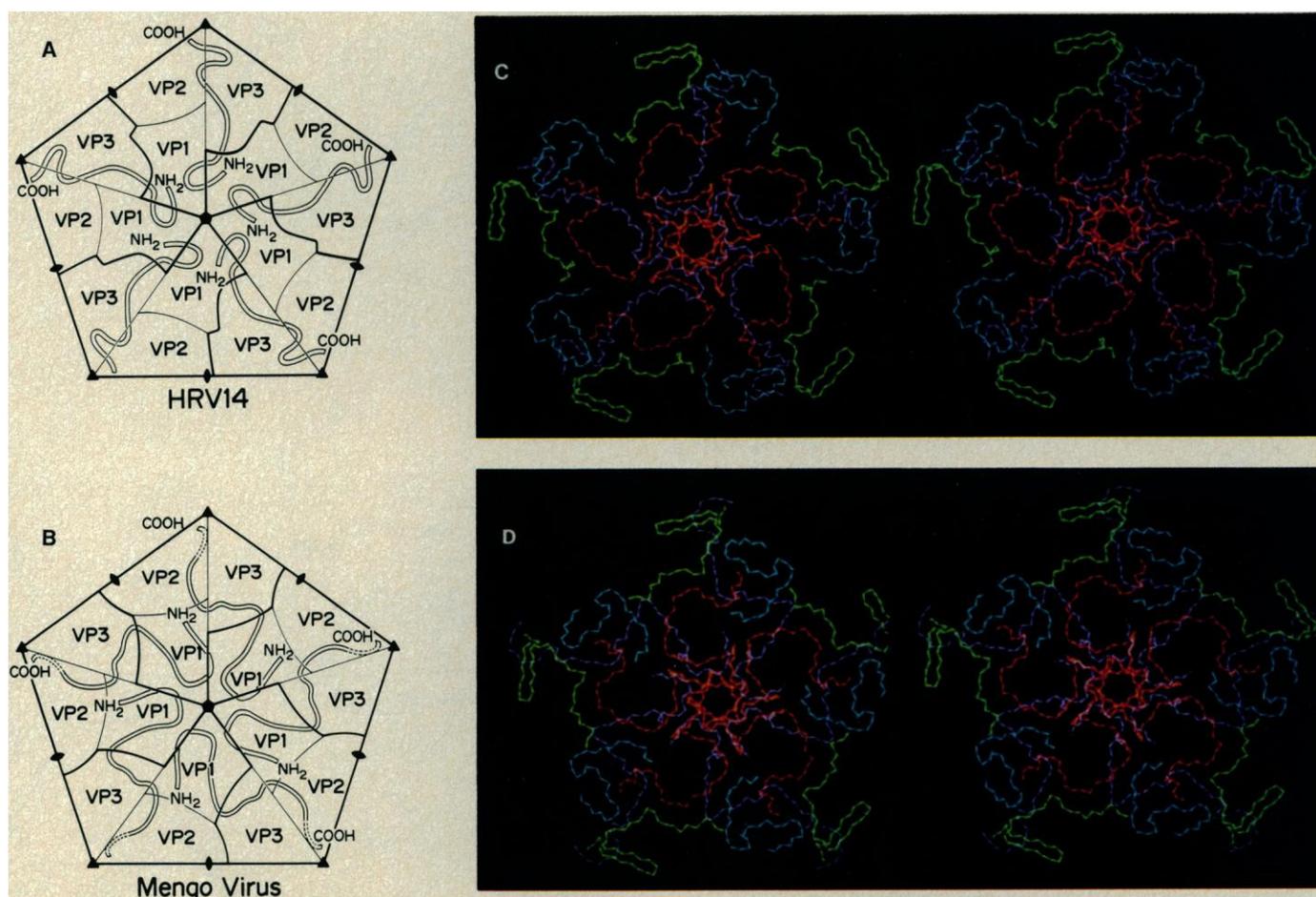
and deletions mostly at the ends of the  $\beta$  strands. The amino acid numbering and secondary structural nomenclature is also shown.

RNA by pentameric 14S assembly units (5-7). This residue is homologous with Ser<sup>11</sup> of Mengo virus VP2. Although Ser<sup>11</sup> of Mengo virus and Ser<sup>10</sup> of HRV14 are in essentially the same position, their side chains are oriented in different directions. The electron density in this region is not as well defined in Mengo virus as in HRV14, yet it is consistent with the carboxyl end of VP4 being close to Ser<sup>11</sup> as required for autoproteolysis of VP0 (5, 7). The first seven amino terminal residues of VP2 are partially disordered. There exists some electron density on and around the threefold axis, close to Ser<sup>11</sup> of VP2, which may be the position of amino terminal VP2 residues. The external VP2 puff, which is an insertion in the  $\beta$ -barrel between  $\beta$ E and  $\alpha$ B, is much shorter in Mengo virus than in HRV14. The second turn of the  $\alpha$ -helix within the puff is conserved in Mengo virus, as is the strand that extends the  $\beta$ -sheet next to  $\beta$ C. The neutralizing immunogen site NIm-II of HRV14, located on the puff, is deleted from Mengo virus, but there is an insertion immediately before  $\alpha$ B that interacts with loop II of VP1 and also helps to fill in the narrow part of the canyon as seen in HRV14.

The major difference in VP3 (Fig. 6) is that the "knob," an external insertion in  $\beta$ B, is mostly deleted in Mengo virus compared to HRV14. It forms part of the "south" side (see 5) of the canyon in HRV14; hence, its deletion in Mengo virus significantly alters part of the putative receptor binding site. The Cys<sup>86</sup> and Cys<sup>88</sup> in VP3 form a right-handed disulfide bond within the distorted helix  $\alpha$ A<sub>0</sub>, near the base of the pit, an unusual Cys-X-Cys situation (53). Residues 176 to 182, between  $\beta$ G and  $\beta$ H, also near the base of the pit are disordered in VP3. This is a hypervariable region in the comparison of Mengo virus, EMCV, Theiler's virus, and HRV14.

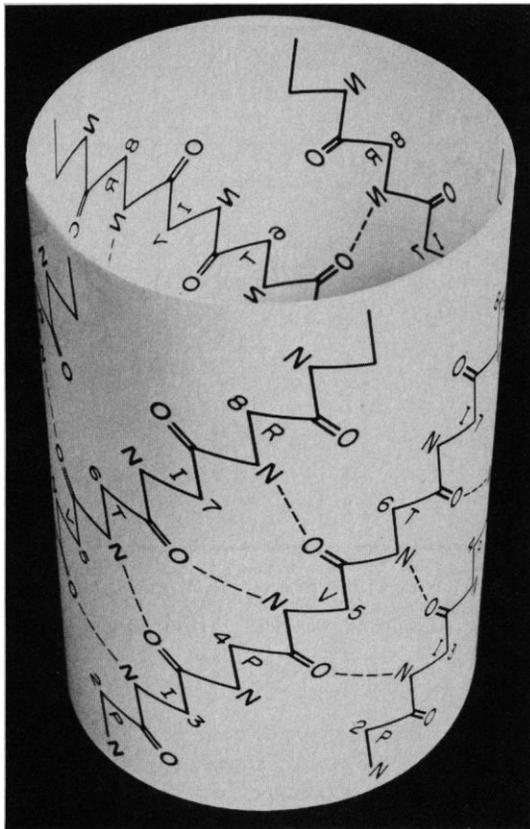
In HRV14, there are three residues deleted and the structure is ordered. They are in the base of the pit close to the disulfide bond. The four residues at the carboxyl end of VP3 are differently situated in Mengo virus compared to HRV14. Unlike HRV14, they do not form part of the lining of the putative receptor binding cavity.

The first 13 residues of VP4 in Mengo virus are mostly disordered as are residues 59 to 70. There is some density, close to Ser<sup>11</sup> of VP2, which could correspond to the four carboxyl terminal residues of VP4, but could be the four amino terminal residues of VP2. While the carboxyl termini of VP4 for HRV14 and Mengo have roughly the same interactions with VP2, the remainder of the chain exhibits an entirely different topological relation with respect to the other three major capsid proteins. In HRV14, VP4 is associated primarily with peptides from only one protomer, running essentially from the fivefold to the threefold axis. However, in Mengo virus VP4 begins near the fivefold axis of one protomer, interacts extensively with the amino terminus of VP3, and ends near the amino terminus of a VP2 in the neighboring, fivefold anticlockwise related, protomer (Fig. 7). VP4 is far more exposed to RNA in Mengo virus than in rhino- or poliovirus. Most parts of the amino ends of VP1 and VP3 are protected from RNA interaction. Because of (i) the amino terminal deletion of 34 residues in VP1 of Mengo virus relative to HRV14, (ii) the disorder of the carboxyl end of VP1, (iii) the change in direction of the carboxyl end of VP3, and (iv) the altered course of VP4, the interactions within a protomer of Mengo are somewhat less than in HRV14. Conversely, the interactions among protomers within a pentamer are enhanced relative to those in HRV14.

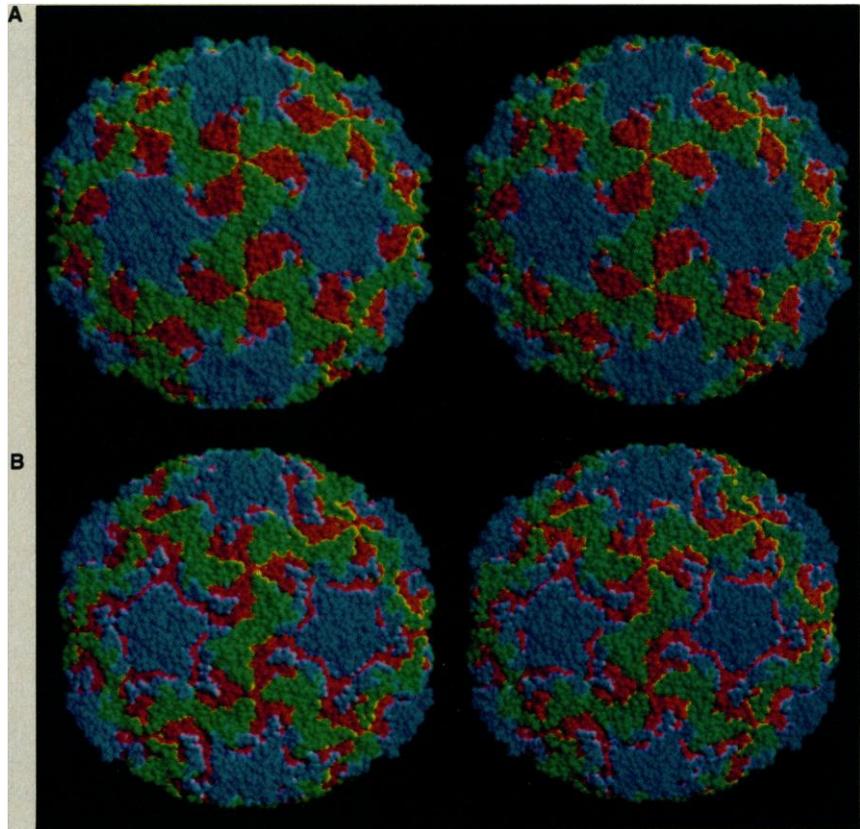


**Fig. 7.** Pentameric 14S assembly unit as seen from the outside, showing the course of VP4 represented by a ribbon in (A) HRV14 and (B) Mengo virus. The 6S protomer units within the pentamer are shown in thicker outline. An

accurate stereoscopic view is rendered on the computer graphics for (C) HRV14 and (D) Mengo virus by cutting away all the  $\beta$ -barrels, leaving only the amino ends of VP1 (blue), VP3 (red), and the whole of VP4 (purple).



**Fig. 8.** Hydrogen bonding arrangement in five-stranded parallel  $\beta$  cylinder made by the amino ends of VP3 about each icosahedral fivefold axis.



**Fig. 9.** Space-filling external view of (A) Mengo virus and (B) HRV14. Color code: VP1 (blue), VP2 (green), VP3 (red).

**Receptor binding site.** The canyon region of HRV14, which has been identified as the putative receptor binding site (5), is substantially different in Mengo virus. Half of the canyon is filled in by the additional two loops (loops I and II) in VP1 and the modified puff of VP2. This leaves a 22 Å deep and 30 Å wide pit (Fig. 6), corresponding roughly to the widest section of the canyon in HRV14. Because of the disorder of the carboxyl terminus of VP1, the change in position of the carboxyl terminus of VP3, and the deletion of part of the knob of VP3, the Mengo virus residues lining the pit are mostly different from those lining the canyon in HRV14. The pit also contains the seven disordered residues (176 to 182) of VP3. The residues lining the walls of the pit (Fig. 3) are donated about equally by VP1 and VP3, whereas the HRV14 canyon is lined primarily by residues from VP1 (Fig. 3). The base of the pit is formed almost exclusively by residues from VP3 whose structure is more conserved than VP1.

Although the structural motifs of many animal, plant, and insect RNA icosahedral viruses are based on the same pseudo  $T = 3$  organization of  $\beta$ -barrels, the surface features are quite different. The deep depressions on the HRV14 and Mengo virus surfaces are quite remarkable and unlike the surfaces of plant [tomato bushy stunt virus (TBSV) and SBMV] and insect (12) viruses. Viruses of vertebrates must contend with the immune defense system of the host. The canyon and pit structures of HRV14 and Mengo virus, respectively, may thus provide a viable receptor binding site inaccessible to host antibodies (Fig. 2). The difference between the HRV14 canyon and the Mengo virus pit may reflect differences in recognition and binding of their respective host cell receptors. Amino acid differences between EMCV and Mengo virus and, in particular, Theiler's virus and Mengo virus are on the surface of the virus, but with residues lining the pit (apart from the disordered section in VP3) remarkably well conserved. It is possible, therefore,

that the cardioviruses all use the same, or very similar, cellular receptors. Tissue specificity is perhaps controlled by differential stability of the virion, the flexible polypeptide in the pit or other factors.

**Potential immunogenic sites.** Neutralizing immunogenic (NIm) sites have been identified for HRV14 (54, 55) and polio (15) by screening for escape mutants to neutralizing monoclonal antibodies. The four types of NIm sites on HRV14 (60 of each type on the virion) occur at the most exposed and external parts of the capsid. They also correspond to hypervariable regions among rhino- and poliovirus sequences.

No specific information is available for Mengo virus about the positions of neutralizing immunogens, although Lund *et al.* (17) have shown, using polyclonal antibodies, that only isolated VP1 and VP2 induce the production of neutralizing antibodies. The structure shows only one type of extensive protrusion (or 60 per virion) consisting of the top corner (residues 58 to 64), loop I (residues 78 to 83), loop II (residues 93 to 105) in VP1, and the puff (residues 157 to 162) of VP2. These residues are positioned at a radius between 150 Å and 163 Å, whereas no other parts of the virus extend to a radius greater than 150 Å. Furthermore, these loops are in regions that correspond to the greatest frequency of changes between Mengo virus, EMCV, and Theiler's virus. Hence, this protrusion may represent the immunodominant NIm site on Mengo virus.

**The site of binding for antiviral compounds.** The binding site of two antiviral compounds, WIN 51711 and WIN 52084 (developed by the Sterling-Winthrop Research Institute), which inhibit uncoating after host cell membrane penetration has been determined on HRV14 (56). These compounds consist of an isoxazole group (I) connected by a seven-membered aliphatic chain to an oxazoliny phenol group (OP). The I group inserts itself into a hydrophobic

**Table 2.** Comparison of proteins.

Comparison	Protein 1*	Protein 2	Number of equivalenced residues	Percentage of equivalences		rms <sup>†</sup> (Å)	MBC/C‡	κ§ (deg)	Δ (Å)
				Protein 1	Protein 2				
Benchmarks	Hb(β)	Hb(α)	139	95	99	1.9	—		
	GAPDH(NAD)	LDH(NAD)	96	65	67	2.9	1.24		
	T4L	HEWL	78	48	60	4.1	1.53		
Plant viruses	SBMV(C)	TBSV(C)	176	84	93	2.2	1.13	3.1	0.4
HRV14 vs. Mengo	VP1 (HRV14)	VP1 (Mengo)	181 (150) <sup>  </sup>	63 (64)	65 (61)	2.6	1.22	1.7	1.4
	VP2 (HRV14)	VP2 (Mengo)	215 (173)	82 (83)	84 (81)	1.9	0.88	1.7	0.3
	VP3 (HRV14)	VP3 (Mengo)	219 (178)	93 (91)	95 (94)	1.6	1.10	1.5	0.5
HRV14	VP1 (HRV14)	VP3 (HRV14)	168	58	71	3.7	1.34		
	VP2 (HRV14)	VP3 (HRV14)	152	58	64	2.8	1.23		
	VP1 (HRV14)	VP2 (HRV14)	124	43	47	3.2	1.27		
Mengo	VP1 (Mengo)	VP3 (Mengo)	156	56	68	3.1	1.27		
	VP2 (Mengo)	VP3 (Mengo)	146	57	63	2.8	1.34		
	VP1 (Mengo)	VP2 (Mengo)	125	45	49	3.1	1.19		
HRV14 vs. SBMV	VP1 (HRV14)	SBMV(A)	123	43	59	3.3	1.42	3.5	2.5
	VP2 (HRV14)	SBMV(C)	134	51	64	2.8	1.40	19.1	1.1
	VP3 (HRV14)	SBMV(B)	136	58	65	2.6	1.34	15.3	0.6

\*Abbreviations: Hb(α) and Hb(β) are the α and β chains of horse hemoglobin. GAPDH(NAD) and LDH(NAD) are the NAD binding domains of glyceraldehyde-3-phosphate dehydrogenase and lactate dehydrogenase. T4L and HEWL are the T4 phage and hen egg white lysozyme. SBMV and TBSV are southern bean mosaic and tomato bushy stunt viruses. Letters in parentheses refer to the A, B, and C subunits. †rms (Å) is the root-mean-square deviation between equivalenced Cα atoms. ‡MBC/C is the minimum base change per codon between equivalenced residues. §κ° and Δ (Å) is the rotation and translation of the center of gravity of the protein subunit relative to the icosahedral axes to obtain the best superposition. ||Numbers in parentheses refer to the superposition of β-barrel domain only.

pocket in the interior of the VP1 β-barrel of HRV14 while the OP group blocks a “pore” in the base of the canyon.

The WIN antiviral agents do not bind to Mengo virus or EMCV and, hence, do not inhibit infection. Loops I and II of VP1 in Mengo virus fill the part of the canyon containing the pore, thus sterically hindering the entry of the antiviral compounds into the VP1 binding site. Although the pore is blocked in Mengo virus, the cavity within the VP1 β-barrel has sufficient space for the binding of the WIN compounds. The occurrence of such a large cavity lined with hydrophobic residues in VP1 of Mengo virus, rhinovirus, and poliovirus suggests that it might be a functional requirement for the disassembly of these viruses.

**Evolution.** Three-dimensional biological structures are generally conserved over far longer time spans than primary protein sequences and can, therefore, be used for alignment when amino acid sequences have diverged greatly, making their alignment by use of sequence alone difficult or impossible (57). The major variations in homologous structures usually correspond to deletions and insertions while maintaining the essential polypeptide folding motif. The Mengo virus and rhinovirus capsid structures were superimposed on one another by the techniques developed by Rossmann *et al.* (58–60). Structural comparisons also permitted the alignment of VP1, VP2, and VP3 to each other (Fig. 3 and Table 2).

The structural superpositions gave not only the amino acid alignments but also the relative orientation and position of the viral proteins in the capsid (Table 2). The alterations between the quaternary structure of Mengo virus and HRV14 are negligible, as are also those between the two  $T = 3$  plant viruses (TBSV and SBMV). However, there are significant changes between the plant and picornaviruses in the relative arrangement of their capsid proteins, amounting up to 19° rotation and 2.5 Å translation (Table 2).

Table 2 provides three benchmarks that show the degree of similarity of other related proteins (61). These have, like the viral proteins, functional and structural similarity to each other. The large number of similar characteristics of these proteins is now generally accepted as being, most probably, the result of divergent evolution (61). Similarities among the Mengo, rhino-, and plant virus capsid proteins are not as great as between the α and β chains of hemoglobin, but are equal to or greater than those of the other two

benchmark comparisons. Hence, it seems most probable that the various picornaviruses and  $T = 3$  RNA plant viruses evolved from a common ancestor. Argos *et al.* (62) suggest that such a precursor may have been related to an ancient receptor binding protein such as the lectin concanavalin A, which also has the same fold (62). The observation by Roberts *et al.* (63) that the hexon unit of the DNA-containing adenovirus includes two distinct β-barrels similar to the viral protein of the plant and animal RNA viruses may also have a bearing on the origin of these and other viruses, as may also the similarity of the globular component of the hemagglutinin spike of influenza virus (64).

In general, there is greater structural similarity between equivalent viral capsid proteins of Mengo virus and HRV14 than there is between the different viral capsid proteins of either Mengo virus or HRV14. Thus, VP1, VP2, and VP3 most likely diverged from each other (the primordial picornavirus) before the polyprotein of Mengo virus diverged from that of HRV14.

Figure 3 highlights the identical residues within the aligned Mengo and rhinovirus structures, as well as the identical residues of SBMV and TBSV. The conserved residues among the picornaviruses tend to be crowded into the interior of the β-barrels and away from the viral surface, whereas the conserved residues between the plant viruses are more uniformly distributed throughout the particle. This distribution might reflect selection pressure imposed on picornaviruses by the host's immune system, which demands that the viral surface be under constant flux for the virus to escape neutralization. Similarly, the greatest differences between Mengo and rhinovirus occur in VP1, the most external protein.

The alignments in Fig. 3 exhibit features presumably characteristic of picornaviruses and perhaps of β-barrel virus capsid proteins in general. The β-sheet regions tend to have particularly hydrophobic residues in alternating positions as these form the internal portion of the β-barrels. The corners at the wedges' ends are dominated by small-sized residues typical of turn regions and there is an especially well-conserved region near the end of βE. Since Mengo virus has greater similarity to FMDV and hepatitis A virus than does rhinovirus or poliovirus (9), the alignment of Mengo virus and rhinovirus sequences (Fig. 3) permits greater certainty in the alignment of other picornaviruses.

REFERENCES AND NOTES

1. R. R. Rueckert, in *Comprehensive Virology*, H. Fraenkel-Conrat and R. R. Wagner, Eds. (Plenum, New York, 1976), vol. 6, pp. 131-213.
2. T. Morishima, P. R. McClintock, G. S. Aulakh, L. C. Billups, A. L. Notkins, *Virology* **122**, 461 (1982).
3. T. W. Mak, J. S. Colter, D. G. Scraba, *ibid.* **57**, 543 (1974).
4. U. Boege, D. S. W. Ko, D. G. Scraba, *J. Virol.* **57**, 275 (1986).
5. M. G. Rossmann *et al.*, *Nature (London)* **317**, 145 (1985).
6. M. G. Rossmann *et al.*, in *Proceedings of the Meeting on Molecular Evolution of Life* (Södergarn, Lidingö, Sweden, 1986).
7. E. Arnold *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
8. A. C. Palmenberg *et al.*, *Nucleic Acids Res.* **12**, 2969 (1984).
9. A. C. Palmenberg, unpublished results.
10. H. Lipton, private communication.
11. S. Nitayaphan *et al.*, *Intervirology*, in press.
12. J. E. Johnson *et al.*, in preparation.
13. S. C. Harrison, A. J. Olson, C. E. Schutt, F. K. Winkler, G. Bricogne, *Nature (London)* **276**, 368 (1978).
14. C. Abad-Zapatero *et al.*, *ibid.* **286**, 33 (1980).
15. J. M. Hogle, M. Chow, D. J. Filman, *Science* **229**, 1358 (1985).
16. D. L. D. Caspar and A. Klug, *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1 (1962).
17. G. A. Lund, B. R. Ziola, A. Salmi, D. G. Scraba, *Virology* **78**, 35 (1977).
18. P. Talbot, D. J. Rowlands, J. N. Burroughs, D. V. Sangar, F. Brown, *J. Gen. Virol.* **19**, 369 (1973).
19. P. Carthew, *ibid.* **32**, 17 (1976).
20. K. Lonberg-Holm and B. E. Butterworth, *Virology* **71**, 207 (1976).
21. D. G. Scraba, in *The Molecular Biology of Picornaviruses*, R. Perez-Bercoff, Ed. (Plenum, New York, 1979), pp. 1-23.
22. J. S. Horder, J. D. Leonard, D. G. Scraba, *Virology* **97**, 131 (1979).
23. T. W. Mak, D. J. O'Callaghan, J. S. Colter, *ibid.* **40**, 565 (1970).
24. B. Baxt, D. O. Morgan, B. H. Robertson, C. A. Timpone, *J. Virol.* **51**, 298 (1984).
25. K. Strohmaier, R. Franze, K. H. Adam, *J. Gen. Virol.* **59**, 295 (1982).
26. P. D. Minor, P. A. Pipkin, D. Hockley, G. C. Schild, J. W. Almond, *Virus Res.* **1**, 203 (1984).
27. R. J. Colonna, P. L. Callahan, W. J. Long, *J. Virol.* **57**, 7 (1986).
28. D. L. Krah and R. L. Crowell, *ibid.* **53**, 867 (1985).
29. J. E. Mapoles, D. L. Krah, R. L. Crowell, *ibid.* **55**, 560 (1985).
30. J. E. Tomassini and R. J. Colonna, *ibid.* **58**, 290 (1986).
31. A. T. H. Burness and I. U. Pardoe, *J. Gen. Virol.* **64**, 1137 (1983).
32. ———, *ibid.* **55**, 275 (1981).
33. U. Boege, D. G. Scraba, K. Hayakawa, M. N. G. James, J. W. Erickson, *Virology* **138**, 162 (1984).
34. M. Luo *et al.*, *J. Mol. Biol.* **180**, 703 (1984).
35. E. Arnold *et al.*, *Acta Crystallogr.*, in press.
36. M. G. Rossmann, *The Molecular Replacement Method* (Gordon & Breach, New York, 1972).
37. A. C. Palmenberg and G. Duke, in preparation.
38. K. A. O. Ellem and J. S. Colter, *Virology* **15**, 340 (1961).
39. D. G. Scraba, C. M. Kay, J. S. Colter, *J. Mol. Biol.* **26**, 67 (1967).
40. G. Vriend and M. G. Rossmann, in preparation.
41. M. G. Rossmann and D. M. Blow, *Acta Crystallogr.* **15**, 24 (1962).
42. *International Tables for Crystallography*, T. Hahn, Ed. (Reidel, Dordrecht, 1983), vol. A.
43. G. Bricogne, *Acta Crystallogr. Sect. A* **30**, 395 (1974).
44. P. Argos, G. C. Ford, M. G. Rossmann, *ibid.* **31**, 499 (1975).
45. I. Rayment, T. S. Baker, D. L. D. Caspar, *Acta Crystallogr. Sect. B* **39**, 505 (1983).
46. T. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978).
47. W. P. J. Gaykema *et al.*, *Nature (London)* **309**, 23 (1984).
48. E. Arnold and M. G. Rossmann, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5489 (1986).
49. J. E. Johnson, T. Akimoto, D. Suck, I. Rayment, M. G. Rossmann, *Virology* **75**, 394 (1976).
50. I. Rayment, T. S. Baker, D. L. D. Caspar, W. T. Murakami, *Nature (London)* **295**, 110 (1982).
51. M. G. Rossmann *et al.*, *J. Mol. Biol.* **165**, 711 (1983).
52. B. R. Ziola and D. G. Scraba, *Virology* **71**, 111 (1976).
53. J. S. Richardson, *Adv. Protein Chem.* **34**, 167 (1979).
54. B. Sherry and R. Rueckert, *J. Virol.* **53**, 137 (1985).
55. B. Sherry, A. G. Mosser, R. J. Colonna, R. R. Rueckert, *ibid.* **57**, 246 (1986).
56. T. J. Smith *et al.*, *Science* **233**, 1286 (1986).
57. M. G. Rossmann, D. Moras, K. W. Olsen, *Nature (London)* **250**, 194 (1974).
58. S. T. Rao and M. G. Rossmann, *J. Mol. Biol.* **76**, 241 (1973).
59. M. G. Rossmann and P. Argos, *J. Biol. Chem.* **250**, 7525 (1975).
60. ———, *J. Mol. Biol.* **105**, 75 (1976).
61. B. W. Matthews and M. G. Rossmann, *Methods Enzymol.* **115**, 397 (1985).
62. P. Argos, T. Tsukihara, M. G. Rossmann, *J. Mol. Evol.* **15**, 169 (1980).
63. M. M. Roberts, J. L. White, M. G. Grütter, R. M. Burnett, *Science* **232**, 1148 (1986).
64. I. A. Wilson, J. J. Skehel, D. C. Wiley, *Nature (London)* **289**, 366 (1981).
65. We thank Keith Moffat, Wilfried Schildkamp, Brenda Smith, Karl Volz, Jun Tsao, and the Cornell High Energy Synchrotron Source (CHESS) operating staff for help, assistance, and advice in our data collection; Sharon Wilder and Kathy Shuster for help in the preparation of the manuscript; Tom Smith, James Griffith, Jack Johnson, Cyndy Stauffacher, Craig Tucker, M. V. Hosur, and Klaus Piontek (all of Purdue University) for many useful discussions; Howard Lipton (Northwestern University) for access to the Theiler's virus sequence before publication; and Richard M. Hilmer (E. I. Dupont de Nemours Inc.) for providing the program used to generate the space-filling figures. We also thank the staff of the Purdue University Computing Center for advice in programming and the helpful scheduling of computing time. Supported by NIH and NSF grants (M.G.R.), a Medical Research Council of Canada grant (D.G.S.), an NIH grant (A.C.P.), and an NIH postdoctoral fellowship (E.A.).

8 August 1986; accepted 24 November 1986

