Acta Cryst. (1997). D53, 93–102

Structure Determination of Minute Virus of Mice

ANTONIO L. LLAMAS-SAIZ,[a]† MAVIS AGBANDJE-MCKENNA,[a]‡ WILLIAM R. WIKOFF,[a]§ JESSICA BRATTON,[b] PETER TATTERSALL[b] AND MICHAEL G. ROSSMANN[a]*

[a]*Department of Biological Sciences, Purdue University, West Lafayette, IN 47907-1392, USA, and* [b]*Departments of Laboratory Medicine and Genetics, School of Medicine, Yale University, 333 Cedar Street, New Haven, CT 06510, USA. E-mail: mgr@indiana.bio.purdue.edu*

(*Received 24 April 1996; accepted 12 August 1996*)

Abstract

The three-dimensional crystal structure of the single-stranded DNA-containing ('full') parvovirus, minute virus of mice (MVM), has been determined to 3.5 Å resolution. Both full and empty particles of MVM were crystallized in the monoclinic space group $C2$ with cell dimensions of $a = 448.7$, $b = 416.7$, $c = 305.3$ Å and $\beta = 95.8°$. Diffraction data were collected at the Cornell High Energy Synchrotron Source using an oscillation camera. The crystals have a pseudo higher $R32$ space group in which the particles are situated at two special positions with 32 point symmetry, separated by $\frac{1}{2}c$ in the hexagonal setting. The self-rotation function showed that the particles are rotated with respect to each other by 60° around the pseudo threefold axis. Subsequently, a more detailed analysis of the structure amplitudes demonstrated that the correct space-group symmetry is $C2$ as given above. Only one of the three twofold axes perpendicular to the threefold axis in the pseudo $R32$ space group is a 'true' crystallographic twofold axis; the other two are only 'local' non-crystallographic symmetry axes. The known canine parvovirus (CPV) structure was used as a phasing model to initiate real-space electron-density averaging phase improvement. The electron density was easily interpretable and clearly showed the amino-acid differences between MVM and CPV, although the final overall correlation coefficient was only 0.63. The structure of MVM has a large amount of icosahedrally ordered DNA, amounting to 22% of the viral genome, which is significantly more than that found in CPV.

1. Introduction

Parvoviruses are a family of small animal viruses which package their single-stranded (ss) DNA genomes within non-enveloped $T = 1$ icosahedral capsids. They are widespread and infect many different animal species from moths to man. Parvoviral particles have an approximate external radius of 140 Å and a molecular weight of $5.5–6.2 \times 10^3$ kDa. Particles of the murine parvovirus, minute virus of mice (MVM), contain a total of 60 subunits, consisting of three viral proteins designated VP1, VP2 and VP3. There are about nine VP1 subunits per particle, with the balance of the subunits being made up of VP2 in empty (lacking DNA) capsids or a mixture of VP2 and VP3 in full virions (Tattersall, Cawte, Shatkin & Ward, 1976). The dominant capsid protein, VP2, is the C-terminal 64 kDa region of the 83 kDa VP1 polypeptide. In DNA-containing particles, many VP2 proteins undergo post-assembly cleavage, removing about 20 residues from their amino termini, to generate VP3 (Tattersall, Shatkin & Ward, 1977).

The atomic resolution structures of full and empty canine parvovirus (CPV) (Tsao *et al.*, 1991; Wu & Rossmann, 1993) and feline parvovirus (FPV) (Agbandje, McKenna, Rossmann, Strassheim & Parrish, 1993) have been reported. The structural proteins have an eight-stranded antiparallel β-barrel topology, frequently found in viral capsid proteins (Rossmann & Johnson, 1989). The first 37 residues of VP2, containing a predominantly poly-Gly conserved sequence, are not visible in the structures, and probably one in five amino termini are externalized along the fivefold axes. Unlike most other virus capsids, but like the capsid of the ssDNA phage φX174 (McKenna, McKenna, Xia, Willingmann, Ilag, Krishnaswamy *et al.*, 1992), there are large insertions between the β-strands of the β-barrel. These form the principal surface features and adapt the virus to its specific host. One of the surface features of CPV and FPV is a small spike around each of the threefold axes.

There are extensive sequence differences between the six parvovirus genera (Murphy *et al.*, 1995) comparable to those between picornavirus genera. MVM and CPV belong to the *Parvovirus* genus of parvoviridae. Sequence comparisons between MVM and CPV (Chapman & Rossmann, 1993) show that the loops which dominate the viral surface, particularly the threefold spikes, are quite dissimilar. Of the 587 amino acids in VP2 of MVM, 52% are identical to CPV.

MVMp, the prototype strain, infects fibroblast cells, while MVMi is an immunosuppressive strain which

† Present address: Departamento de Cristalografía, Instituto de Química-Física 'Rocasolano', CSIC, Serrano 119, 28006 Madrid, Spain. ‡ Present address: Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, England. § Present address: Department of Molecular Biology MB 13, Scripps Research Institute, 10666 N. Torrey Pines Road, La Jolla, CA 92037, USA.

infects lymphocyte T cells (Tattersall & Bratton, 1983); mutations that switch tropism have been shown to map into the VP2 capsid gene (Gardiner & Tattersall, 1988). Most of the MVM host range mutations in VP2 correspond to amino acids that are on the surface of CPV near the edges of the threefold spikes.

We report here the rather difficult structure determination of MVMi on account of the pseudo higher order space-group symmetry. Variations of such problems are encountered fairly frequently in the determination of virus crystal structure (Zlotnick *et al.*, 1993; Muckelbauer, Kremer, Minor, Diana *et al.*, 1995), because the crystal lattice is unable to accommodate the complete icosahedral symmetry of the virus particles. Hence, our account of how some of the problems were solved may be helpful in similar situations that are likely to occur in other virus structure determinations. A discussion of the biological implications will be given elsewhere (Agbandje-McKenna, Llamas-Saiz, Wang, Tattersall & Rossmann, 1997).

2. Methods

2.1. *Crystallization and data collection*

MVMi virus was recovered by transfection from the infectious clone pMVMi (Gardiner & Tattersall, 1988) and propagated at low passage in S49 murine lymphoma cells. Infected cells were extracted by freezing and thawing three times in TE 8.7 (50 m*M* Tris, 0.5 m*M* EDTA, pH 8.7). Extracts were cleared by low-speed centrifugation and the virus sedimented to equilibrium through a 60% sucrose cushion over CsCl (1.40 g cm^{-3}), each in the same buffer. For some preparations, this procedure was modified by performing a chloroform extraction step before layering the virus supernatant onto the step gradient. Gradients were fractionated and analyzed by hemagglutination assay (Tattersall *et al.*, 1976). Fractions containing either full virions or empty capsids were separately re-banded in CsCl, dialyzed and concentrated prior to crystallization.

Crystals were grown using the hanging-drop vapor-diffusion method with conditions similar to those used for CPV (Luo, Tsao, Rossmann, Basak & Compans, 1988). The reservoir solution contained 0.75%(*w*/*v*) PEG 8000 and 8 m*M* $CaCl_2.2H_2O$ in 10 m*M* Tris–HCl (pH 7.5), over which was suspended a 10 µl hanging drop produced by mixing 5 ml of virus solution (10 mg ml^{-1}) in 10 m*M* Tris–HCl at pH 7.5 with 5 ml of reservoir solution. Crystals grew to a maximum dimension of 0.4 mm in about 4 to 8 weeks. Crystals of full and empty particles were isomorphous.

X-ray diffraction data of both full and empty particles were collected at the F1 station of the Cornell High Energy Synchrotron Source (CHESS) using Fuji imaging plates as detectors. The crystals were cooled to 277 K to prevent excessive radiation damage. Oscillation angles of 0.3 and 0.4° were used with an X-ray wavelength of approximately 0.91 Å and a crystal-to-detector distance of 280 or 300 mm. Exposure times varied from 10 to 40 s depending on the size of the crystal. The imaging plates were scanned using a BAS 2000 Fuji scanner with a 100 µm raster step size. To minimize radiation damage, no setting photographs were taken of the randomly oriented crystals (Rossmann & Erickson, 1983). This produced four to 18 useful images per crystal, with diffraction extending to 3.0 Å resolution in most cases. The final data set contained 280 images from 35 crystals of full MVMi particles. The data for the empty particles were less complete, consisting only of 45 images from five crystals.

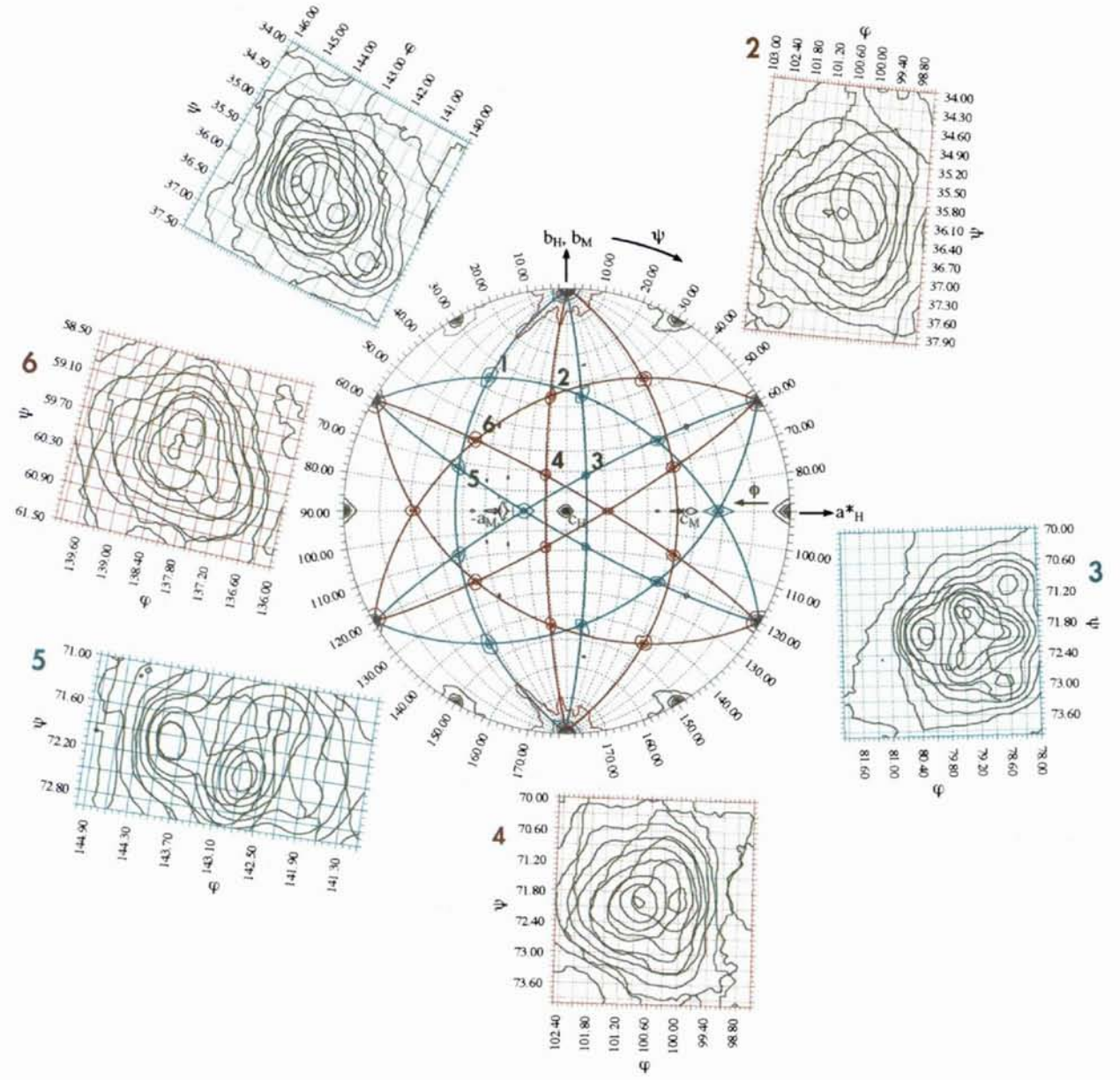
2.2. *Data processing*

The full particle diffraction data were processed and analyzed first. The *OSC*123 (Kim, 1989) and *DENZO* (Otwinowski, 1993) programs were used for initial indexing. Both *R*32 and *C*2 were possible space groups, consistent with the crystal symmetry and systematic absences. Initially, the data were processed using the higher symmetry space group *R*32 in the hexagonal setting with $a_H = 414.2$ and $c_H = 569.5$ Å (the subscript *H* denotes the hexagonal setting of *R*32). However, the scaling of the data was not particularly satisfactory and phase determination appeared to have problems. Thus, the data were reprocessed in the lower symmetry space group, *C*2 (see below for details). The relationship between the rhombohedral's hexagonal setting and the chosen monoclinic system are shown in Fig. 1 and are given by,

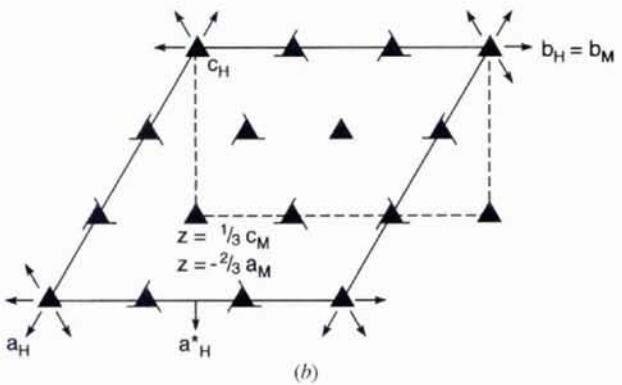
$$a_M^2 = \tfrac{1}{3}a_H^2 + \tfrac{4}{9}c_H^2$$
$$b_M = b_H = a_H$$
$$c_M^2 = \tfrac{1}{3}a_H^2 + \tfrac{1}{9}c_H^2$$
$$\cos\beta_M = (a_H^2 - \tfrac{2}{3}c_H^2)/(a_H^4 + \tfrac{4}{9}c_H^4 + \tfrac{5}{3}a_H^2c_H^2)^{1/2},$$

where the subscript *M* denotes the monoclinic cell parameters. The crystal orientation, mosaicity and cell dimensions were further refined (after sorting out the orientations as described below) using the Purdue data processing package (Rossmann, 1979; Rossmann, Leslie, Abdel-Meguid & Tsukihara, 1979). The final monoclinic cell dimensions were $a_M = 448.7$, $b_M = 416.7$, $c_M = 305.3$ Å and $\beta_M = 95.8°$.

When reprocessing in space group *C*2, data from each crystal were processed using each of the three different twofold axes of the *R*32 system in turn as the unique monoclinic twofold axis. The images belonging to the same crystal were scaled together for each of the three possible orientations. That orientation which used the true crystallographic twofold axis should have the lowest $R_{\rm merge}$. The best discrimination was obtained for crystal 17.2 (crystal 17 of the second CHESS trip), for which 18 good images were available (Table 1). This crystal was used as an initial standard to which



(*a*)



(*b*)

Fig. 1. (*a*) Self-rotation functions for $\kappa = 180^{\circ}$ using data between 10 and 6 Å resolution in the *C*2 space group. Axial directions for the rhombohedral system's hexagonal setting are shown as a_H^*, b_H, and c_H. The axial directions for the monoclinic system a_M, b_M and c_M are also shown. The orientation is chosen to correspond to the rhombohedral space-group setting looking down the threefold axis. Great circles are shown connecting adjacent twofold axes. Contours and grid circles for the two independent particles are differentiated by using red and blue. Insets show details of selected peaks using data between 4.0 and 3.0 Å resolution. The black contours show the rotation function calculated with the data processed in space group *R*32, whereas the green contours show the rotation function calculated for space group *C*2. (*b*) The relationship between the pseudo higher order space group *R*32 and the actual space group *C*2. Axes are labeled as indicated above.

Table 1. *Discrimination for selecting the correct crystallographic axis of crystal 17.2*

Orientation	R_{merge}*	Number of reflections Measured	Unique	Common
1	14.62	11076	10994	82
2	26.19	11000	10894	106
3	9.32	11241	11054	187

* $R_{merge} = [\sum_h \sum_i |(\langle I_h \rangle - I_{hi})|]/(\sum_h \sum_i \langle I_h \rangle)$, for $I_h > 5\sigma(I)$.

Table 2. *Discrimination in scaling crystal 9.3 to crystal 17.2, with the latter in orientation 3*

Orientation of crystal 9.3	R_{merge}*	Number of reflections Measured	Unique	Common
1	16.62	23380	21953	1427
2	12.13	23372	22618	754
3	16.31	23360	22609	751

* R_{merge} is defined in Table 1.

Table 3. *Data merge cycles*

Cycle	Resolution for F_{calc} computation (Å)	No. of crystals	No. of reflections Measured	Unique	R_{merge}* (15–3.5 Å)
0	15.0–5.0	15	456449	353192	12.60
1	15.0–5.0	15	457790	356205	12.87
2	15.0–5.0	30	706345	478493	14.03
3	15.0–3.5	30	705882	486044	13.67
4	15.0–3.5	38	933857	551094	15.51

* R_{merge} is defined as in Table 1, but using $I > 3\sigma(I)$.

other crystal data could be scaled. Crystal 9.3 was found to have a good discrimination (Table 2). The combined data set of crystal 17.2 and 9.3 was used as a new standard to select the correct orientation for the third crystal, and so forth. The gradual increase in data permitted better discrimination, on account of better overlap between the current standard and the data of a crystal with an as yet unknown orientation. This procedure permitted the combination of 15 different crystals. A self-rotation function using these data showed that there was a 2.2° deviation from the exact orientation of the particles in space group $R32$, demonstrating that the data combination had been successful. However, there remained over 30 crystals whose orientation could not be established with confidence.

The 15-crystal data set was used to obtain an initial MVMi structure determination (see below), which was used to compute F_{calc}'s from the Fourier back-transform of the averaged map. These F_{calc}'s included data for the previously unobserved reflections. By scaling to the F_{calc}'s, it was possible to establish a few more crystal orientations, thus enlarging the standard data set. This data set was used for further cycles of electron-density averaging, yielding an improved set of F_{calc}'s that was then used for further crystal orientation selection. A total of four such cycles were completed to finally include 35

Table 4. *Final data statistics*

Resolution (Å)	R_{merge}*	% Available data	No. unique reflections
(*a*) Pseudo $R32$ space group (full particles)			
25.00–12.50	35.13	87	3862
12.50–8.33	26.99	91	11028
8.33–6.25	11.39	93	21900
6.25–5.00	12.81	91	35276
5.00–4.17	13.90	87	49820
4.17–3.57	14.19	76	61374
3.57–3.12	17.02	50	53995
3.12–3.00	20.22	9	11985
(*b*) Monoclinic $C2$ space group (full particles)			
25.00–12.50	23.58	71	9436
12.50–8.33	15.19	77	27833
8.33–6.25	12.21	78	54782
6.25–5.00	13.70	79	85507
5.00–4.17	15.01	69	118958
4.17–3.57	16.04	55	133070
3.57–3.12	18.99	30	98130
3.12–3.00	21.06	6	22920
(*c*) Monoclinic $C2$ space group (empty particles)			
25.00–12.50	9.51	18	2395
12.50–8.33	11.35	20	7244
8.33–6.25	10.24	19	13427
6.25–5.00	13.29	17	19162
5.00–4.17	16.21	15	25288
4.17–3.57	16.61	8	18814
3.57–3.12	19.84	2	5752

* R_{merge} is defined as in Table 1, but using $I > 3\sigma(I)$.

Table 5. *Wavelength adjustment for separate synchrotron data trips*

Trip No.	Trip date	No. of images	Adjusted wavelength
1	Jan. 1993	45	0.9087
2	Sept. 1993	84	0.9071
3	Apr. 1994	81	0.9095
4	June 1994	15	0.9124
5	Nov. 1994	55	0.9157

out of 45 possible individual crystal data sets (Table 3). The final data statistics accounted for 64.6% of the data to 3.5 Å resolution, with an R_{merge} of 15.5% and 1.7 measurements of each reflection on average (Table 4). The empty particle data were oriented with respect to the F_{calc}'s of the completed full particle data set; the final R_{merge} was 15.1%, representing 12.7% of the theoretically possible number of reflections.

The diffraction data for the full particles were collected over a series of five different visits to CHESS. As the precise wavelength used on each trip was not necessarily identical, the cell dimensions associated with data from each trip were separately post-refined. Some systematic differences were found in the cell dimensions derived from each trip. All data had been processed assuming a wavelength of 0.908 Å. This was adjusted separately for each data trip, relative to post-refined cell dimensions using only the data from trips 2 and 3 (Table 5). The original wavelength was based on a comparison

with data collected for human rhinovirus 14, for which the cubic cell dimension had been determined using a Cu $K\alpha$ source.

The final R_{merge} values were substantially worse than for other viral diffraction data sets collected under similar conditions (Arnold *et al.*, 1987; McKenna, Xia, Willingmann, Ilag & Rossmann, 1992). In these latter cases, the effective mosaic spread was found to be around 0.05° by post-refinement, while for MVMi a limit of 0.25° was set for both the horizontal and vertical spread to avoid a breakdown of the post-refinement and scaling processes. A separate mosaic spread was post-refined for each crystal, but not each image. Of these, 31 crystals hit either the horizontal, vertical or both mosaic spread limits. A similar proportion hit the limit in the post-refinement of the empty particle data. Clearly, there must have been a large number of partial reflections which were considered to be full, causing the poor R_{merge} values.

A possible cause of the poor crystal quality may have been that the crystals were extremely temperature sensitive. Extensive precautions were required to stabilize the temperature of the crystal environment, particularly in transportation to CHESS. Changes of few degrees from room temperature caused the crystals to dissolve.

2.3. *Structure determination*

The Matthews coefficient (Matthews, 1968), V_M, for the pseudo-$R32$ space group is 2.4 $Å^3\,Da^{-1}$, assuming two virions in the rhombohedral cell (six in the equivalent hexagonal setting). This requires that each particle is situated on a special 32 position, separated by $\frac{1}{2}$ of the hexagonal c_H axis. The particles must be rotated relative to each other by 60°, otherwise the length of the hexagonal c_H axis would be halved. This was verified by a self-rotation function (Fig. 1), which clearly shows the two different orientations.

In the monoclinic $C2$ space group, one of the rhombohedral twofold axes becomes the monoclinic twofold axis along b_M. Since the rhombohedral threefold axis is perpendicular to the twofold axes, the pseudo-threefold axis in $C2$ will be in the ac plane, as seen in the rotation function (Fig. 1). In $C2$, the two independent particles are each situated on a crystallographic twofold axis, giving a total non-crystallographic redundancy of 60. The adjustable parameters are the rotation of each particle about its crystallographic twofold axis and the relative translation of the particles along these axes. Thus, there are three parameters to be determined, with the particle positions being at (0,0,0) and $(0,\sim\frac{1}{2},\frac{1}{2})$. The high resolution self-rotation function showed that the particles have moved about the crystallographic twofold by 2.2° relative to their average position in space group $R32$. A more accurate orientation of each of the two particles was determined by making a least-squares fit to the high-resolution rotation-function peaks. The r.m.s. deviations of the ends of the rotation vectors to idealized icosahedral symmetry are 0.17 and 0.55 Å, at a radius of 140 Å, for the 'blue' and for the 'red' particles, respectively (see Fig. 1 for color definitions). The $[P]$ matrix defines the rotation necessary to orient a standard icosahedron in the 'h cell' to its orientation in the unknown 'p cell', such that $\mathbf{X} = [P]\mathbf{Y}$, where $\mathbf{X}$ and $\mathbf{Y}$ are Cartesian coordinates in the h and p cells, respectively (Rossmann *et al.*, 1992). The $[P]$ matrices for each of the icosahedra were found to be,

$$\begin{pmatrix} 0.8686 & 0.0 & 0.4955 \\ 0.0000 & 1.0 & 0.0000 \\ -0.4955 & 0.0 & 0.8686 \end{pmatrix}$$ for the first ('blue') particle,

and

$$\begin{pmatrix} -0.2840 & 0.0 & -0.9588 \\ 0.0000 & 1.0 & 0.0000 \\ 0.9588 & 0.0 & -0.2840 \end{pmatrix}$$ for the second ('red') particle,

where $\mathbf{Y}$ is an orthogonal system in the $C2$ MVMi cell as defined by Rossmann & Blow (1962).

The structure of monoclinic CPV was used as an initial phasing model. An electron-density map of CPV was calculated in its $P2_1$ cell, using the observed CPV amplitudes and the refined phases that had been obtained by electron-density averaging (Tsao *et al.*, 1991, 1992). The resultant averaged map was transformed to the standard orientation in the h cell. This also defined the envelope of the viral particles, as the 60-fold redundancy will have removed other neighboring particles that lie outside the reach of the local non-crystallographic symmetry. The h cell contained a structure of a full CPV particle appropriate for interpolation into the MVMi cell.

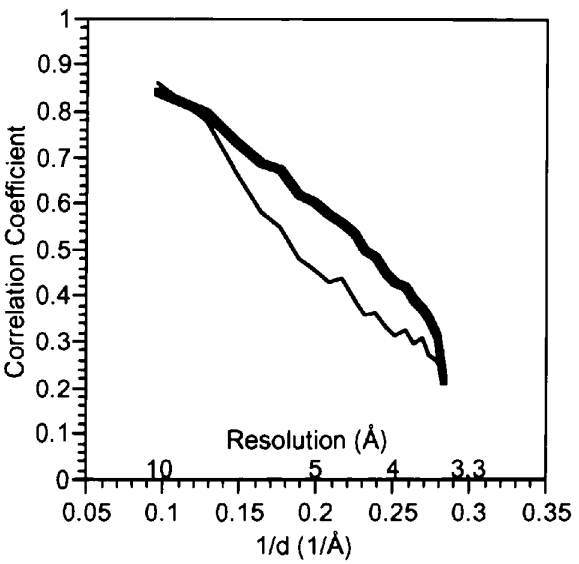
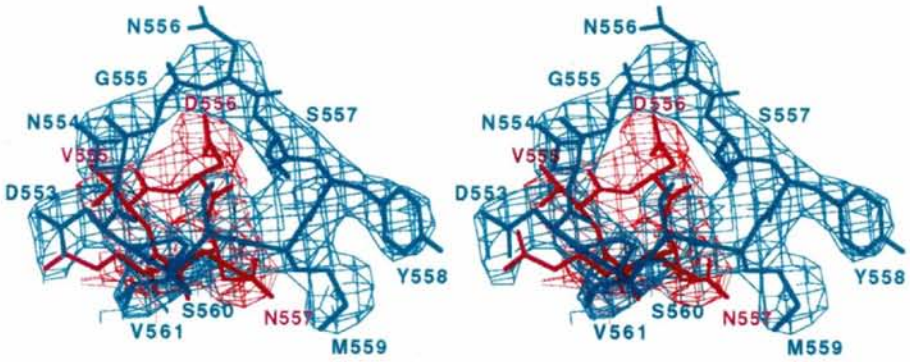
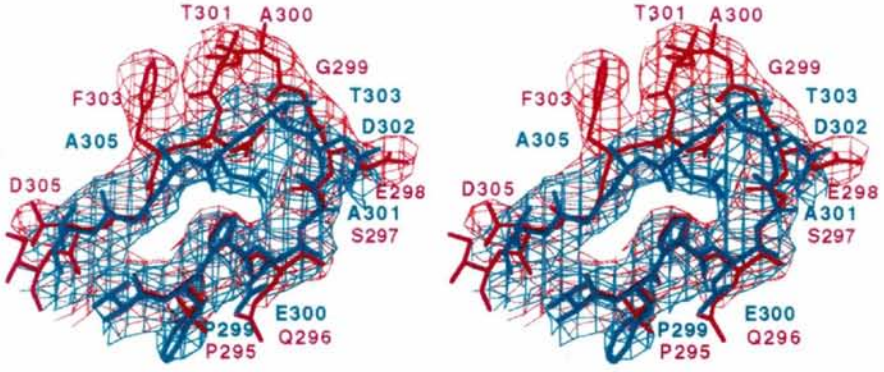


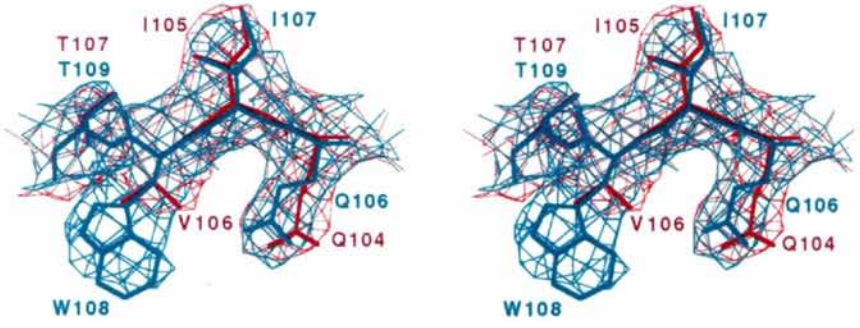
Fig. 2. Distribution of correlation coefficients with resolution, showing the original refinement in space group $R32$ (thin line) and the final results in space group $C2$ (thick line).



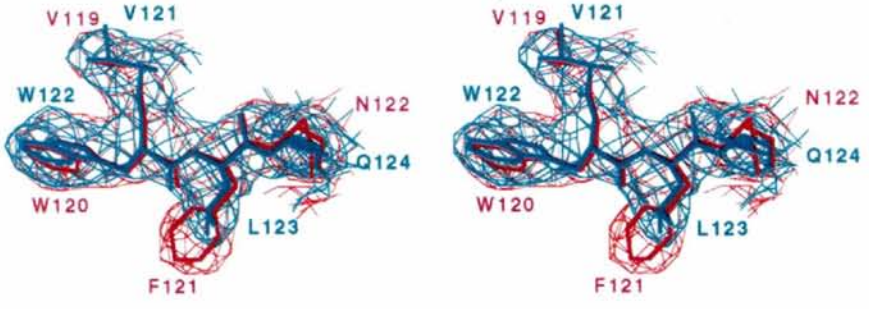
(*a*)



(*b*)



(*c*)



(*d*)

Fig. 3. Stereoviews of the MVMi electron density (blue) superimposed on the original CPV density (red). (*a*) Insertion of six residues (554–559) in MVMi, relative to CPV. (*b*) Deletion of three residues in MVMi after residue 303, relative to CPV. (*c*) Substitution of CPV residue Val106 with Trp108 in MVMi. (*d*) Substitution of CPV residue Phe121 with Leu123 in MVMi.

Table 6. *Refinement of particle positions and orientation using the 'climb' procedure*

Cycle	Particle 1 ('blue') y	Particle 1 ('blue') κ (°)	Particle 2 ('red') y	Particle 2 ('red') κ (°)	No. of crystals in data set	Resolution range (Å)
0	0.0	29.70	0.5000	−106.50	15	15.0–5.0
9	0.0	29.73	0.5000	−106.52	15	15.0–5.0
14	0.0	29.78	0.5000	−106.50	15	15.0–5.0
18	0.0	29.78	0.5000	−106.50	15	15.0–5.0
10	0.0	29.77	0.5000	−106.53	30	15.0–5.0
13	0.0	29.83	0.5000	−106.54	30	15.0–5.0
16	0.0	29.85	0.5000	−106.55	30	15.0–5.0
19	0.0	29.87	0.5000	−106.56	30	15.0–5.0
9	0.0	29.87	0.5000	−106.54	30	15.0–3.5
25	0.0	29.86	0.5009	−106.54	35	15.0–3.5
33	0.0	29.86	0.5000	−106.53	35	15.0–3.5

The electron density in the MVMi cell, based on the CPV structure, was Fourier back-transformed. The R factor** and correlation coefficient † (CC) were 40.4 and 0.40, respectively. The first five cycles of molecular replacement electron-density averaging (Rossmann *et al.*, 1992) employed a partial data set which included only the reflections from the original 15 crystals. Unit weights were applied to the Fourier terms and no use was made of the F_{calc} values for unobserved reflections. The density was averaged within a spherical shell with inner and outer radii of 70 and 140 Å, respectively, using planes to separate overlapping spheres. The orientations and relative position of the particles were further refined on 11 different occasions by searching for the parameters that gave minimal scatter between non-crystallographically equivalent density (the 'climb' procedure) (Muckelbauer, Kremer, Minor, Tong *et al.*, 1995) (Table 6). Subsequently, more observed data were introduced and weighted F_{calc} values were used where there were no observed reflections. The R factor decreased to 30.8%, and the mean correlation coefficient improved to 0.63. Inspection of the electron density after this phase-improvement procedure, using the interactive graphics program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991), showed that some surface loops of the capsid had been truncated. Thus, the outer radius was increased to 148 Å. Nine more cycles of averaging with a mask based on the averaged h-cell density and using the final data set resulted in the map used for modeling the MVMi structure. The final R factor and correlation coefficients were 30.6 and 0.63, respectively, for all merged data (Table 4, Fig. 2).

The non-crystallographic redundancy is only 20 in space group $R32$, as opposed to 60 in space group $C2$. The lower redundancy in $R32$ should produce better correlation coefficients if everything else were equal, but the actual results showed that $C2$ phase refinement was distinctly better (Fig. 2), verifying that $C2$ was the correct choice of space group. Nevertheless, the correlation coefficients were not as good as those found for other virus determinations (Arnold *et al.*, 1987).

2.4. *Map interpretation*

Although the phase refinement was not as good as had been hoped, the electron-density map calculated at 3.5 Å resolution was of high quality. The map was interpreted with respect to the known amino-acid sequence of MVMi, guided by the known CPV structure. In those places where there was an insertion, deletion or a structurally recognizable amino-acid difference, the MVMi structure was invariably obvious, showing that the initial CPV phase bias had been eliminated.

The largest difference between MVMi and CPV occurs at the inserted MVMi residues 554–559 (Fig. 3*a*). The inserted residues are clearly visible on the particle surface. An example of a deletion of three residues of CPV relative to MVMi occurs after MVMi residue 303 (Fig. 3*b*); a small CPV residue (valine) changed to a larger residue in MVMi (tryptophan) is shown in Fig. 3(*c*), while the converse, namely a phenylalanine in CPV changed to a leucine in MVMi, is shown in Fig. 3(*d*).

The CPV model was completely rebuilt to fit the MVMi electron density and amino-acid sequence. The polypeptide chain of VP2 could be traced from residue 39 to the carboxy-terminal residue 587, although there were some regions of weaker density. Comparison of the mean radial distance of Cα atoms between MVMi and CPV showed a small, but definite, linear trend (Fig. 4). At 140 Å radius, the average radial distance of MVMi atoms was 0.5 Å greater than in CPV. This

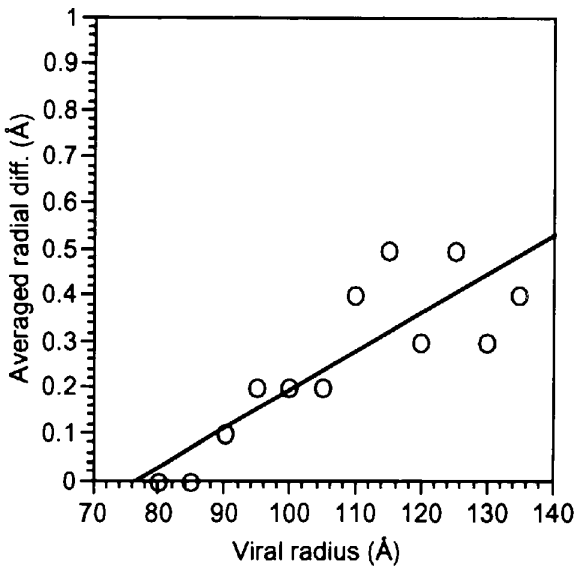


Fig. 4. Comparison of averaged radial difference between MVMi and CPV Cα positions as a function of radial distance from the viral center.

* $R = [\sum|(F_{obs} - kF_{calc})|/\sum|F_{obs}|] \times 100$, where $k = \sum F_{obs}/\sum F_{calc}$.

† $CC = \dfrac{[\sum(\langle F_{obs}\rangle - F_{obs})(\langle F_{calc}\rangle - F_{calc})]}{[\sum(\langle F_{obs}\rangle - F_{obs})^2 \sum(\langle F_{calc}\rangle - F_{calc})^2]^{1/2}}$.

might be the result of a true change in the size of the MVMi particle or as a result of poor calibration of wavelength. Since the latter seems more likely, the MVMi coordinates have been adjusted appropriately. The r.m.s. deviation for equivalent Cα atoms between MVMi and CPV was 0.96 Å based on superposition of the icosahedral symmetry axes and 0.94 Å based on the best superposition of the VP2 structure. This comparison determined an alignment (Fig. 5) of MVMi to CPV which differed in some details around insertions and deletions to that given by Chapman & Rossmann (1993).

After completion of the model building, a substantial amount of internal electron density remained uninterpreted. A significant amount of this density corresponded to the 11 DNA nucleotides found in the analysis of CPV (Tsao *et al.*, 1992; Chapman & Rossmann, 1995). Some of the remaining density was interpreted as eight additional DNA nucleotides (Agbandje-McKenna *et al.*,

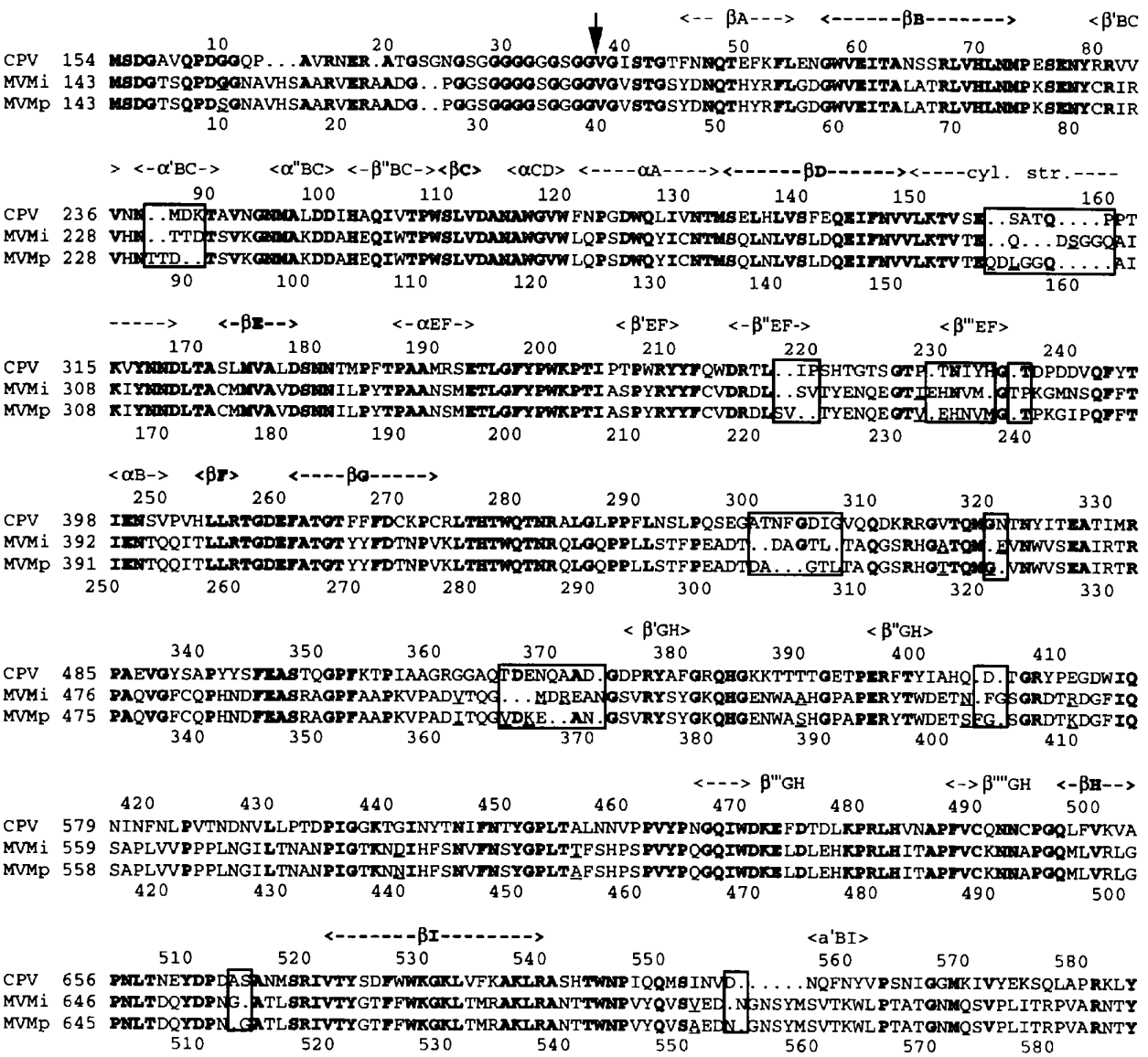


Fig. 5. Alignment of MVMi and CPV based on their respective structures. Also shown is the alignment of MVMp and CPV based on sequence (Chapman & Rossmann, 1993). Regions where the Chapman and Rossmann sequence alignment differ to the current structural alignment are boxed. Conserved residues are in bold. The 14 differences between the two MVM strains are underlined. VP2 residue numbering for CPV is given above and for MVMi below the aligned sequences. VP1 numbers are given at the beginning of each new line in the figure. The arrow points to the beginning of the ordered structure seen in the electron density. The sequences for MVMp and MVMi are those published by Astell *et al.* (Astell, Thomson, Merchlinsky & Ward, 1983; Astell, Gardiner & Tattersall, 1986, respectively). Secondary-structure elements for MVMi are also shown.

1997). The identification of this density as DNA was confirmed in an averaged difference map with respect to the empty particle data. The 19 bases seen in one icosahedral asymmetric unit implies that 60 × 19 bases, or approximately 22% of the total genome, display icosahedral symmetry.

Refinement of the structure has not yet been completed. A discussion of the structure–function relationships will be given elsewhere (Agbandje-McKenna *et al.*, 1997). The present coordinates have been deposited with the Protein Data Bank.*

3. Discussion

Although the electron density was easy to interpret, the indications of quality, such as $R_{\rm merge}$ on the observed data and correlation coefficients, remained poor. It was surprising to find that the space group was not exactly $R32$, but was really $C2$. There are several possible explanations for the difficulties in this structure determination. The identification of the monoclinic twofold axis might be in error for some crystals or some of the domains within the same crystal might correspond to $R32$, while others correspond to $C2$ symmetry. This could account for the high mosaicity of the crystals, because the different domains do not match exactly. A third possibility would be that the virions themselves lack exact icosahedral symmetry. However, any such deviation would be small and, in many cases, the crystallization process would be unable to select the preferred orientation. Nevertheless, such deviation would create some disorder, accounting for poor data and high mosaicity. Lack of icosahedral symmetry might also be caused by variable composition of the virions with respect to their content of VP1, VP2 and VP3. Such asymmetry might also be caused by interactions between the capsid protein and the unusually large amounts of icosahedrally ordered DNA, which cannot be identical in the various asymmetric units.

In general, it would be expected that strictly icosahedral particles would easily crystallize into close-packed arrangements with cubic or trigonal space groups. Usually there is good close packing of particles, but in a surprising number of instances there are small, but irritating, departures, as for example in the crystal packing of coxsackievirus B3 (Muckelbauer, Kremer, Minor, Diana *et al.*, 1995) and nodamura virus (Zlotnick *et al.*, 1993). Like MVMi, both these structures contain two independent particles per crystallographic asymmetric unit and a large amount of nucleic acid folded with icosahedral symmetry.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1MVM, R1MVMSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0639).

We are grateful to our many Purdue colleagues who participated in the data collection trips to CHESS and to the ever helpful CHESS staff. We are grateful to Cheryl Towell and Sharon Wilder for help in preparation of this paper. The work was supported by National Institutes of Health grants to MGR (AI 11219), to Colin Parrish and MGR (AI 33468) and to PT (CA 29303); a Spanish Ministry of Education Fellowship to ALL-S; and a National Institutes of Health Program Project Grant (AI 35212) and a Lucille P. Markey Foundation Award.

References

Agbandje, M., McKenna, R., Rossmann, M. G., Strassheim, M. L. & Parrish, C. R. (1993). *Proteins*, **16**, 155–171.

Agbandje-McKenna, M., Llamas-Saiz, A., Wang, F., Tattersall, P. & Rossmann, M. G. (1997). In preparation.

Arnold, E., Vriend, G., Luo, M., Griffith, J. P., Kamer, G., Erickson, J. W., Johnson, J. E. & Rossmann, M. G. (1987). *Acta Cryst.* A**43**, 346–361.

Astell, C. R., Gardiner, E. M. & Tattersall, P. (1986). *J. Virol.* **57**, 656–669.

Astell, C. R., Thomson, M., Merchlinsky, M. & Ward, D. C. (1983). *Nucleic Acids Res.* **11**, 999–1018.

Chapman, M. S. & Rossmann, M. G. (1993). *Virology*, **194**, 491–508.

Chapman, M. S. & Rossmann, M. G. (1995). *Structure*, **3**, 151–162.

Gardiner, E. M. & Tattersall, P. (1988). *J. Virol.* **62**, 2605–2613.

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

Kim, S. (1989). *J. Appl. Cryst.* **22**, 53–60.

Luo, M., Tsao, J., Rossmann, M. G., Basak, S. & Compans, R. W. (1988). *J. Mol. Biol.* **200**, 209–211.

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

McKenna, R., Xia, D., Willingmann, P., Ilag, L. L. & Rossmann, M. G. (1992). *Acta Cryst.* B**48**, 499–511.

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

Muckelbauer, J. K., Kremer, M., Minor, I., Diana, G., Dutko, F. J., Groarke, J., Pevear, D. C. & Rossmann, M. G. (1995). *Structure*, **3**, 653–667.

Muckelbauer, J. K., Kremer, M., Minor, I., Tong, L., Zlotnick, A., Johnson, J. E. & Rossmann, M. G. (1995). *Acta Cryst.* D**51**, 653–667.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A. & Summers, M. D. (1995). Editors. *Virus Taxonomy. Classification and Nomenclature of Viruses*. Vienna: Springer-Verlag.

Otwinowski, Z. (1993). *Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Rossmann, M. G. (1979). *J. Appl. Cryst.* **12**, 225–238.

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

Rossmann, M. G. & Erickson, J. W. (1983). *J. Appl. Cryst.* **16**, 629–636.

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

Rossmann, M. G., Leslie, A. G. W., Abel-Meguid, S. S. & Tsukihara, T. (1979). *J. Appl. Cryst.* **12**, 570–581.

Rossmann, M. G., McKenna, R., Tong, L., Xia, D., Dai, J., Wu, H., Choi, H. K. & Lynch, R. E. (1992). *J. Appl. Cryst.* **25**, 166–180.

Tattersall, P. & Bratton, J. (1983). *J. Virol.* **46**, 944–955.

Tattersall, P., Cawte, P. J., Shatkin, A. J. & Ward, D. C. (1976). *J. Virol.* **20**, 273–289.

Tattersall, P., Shatkin, A. J. & Ward, D. C. (1977). *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). *Science*, **251**, 1456–1464.

Tsao, J., Chapman, M. S., Wu, H., Agbandje, M., Keller, W. & Rossmann, M. G. (1992). *Acta Cryst.* B**48**, 75–88.

Wu, H. & Rossmann, M. G. (1993). *J. Mol. Biol.* **233**, 231–244.

Zlotnick, A., McKinney, B. R., Munshi, S., Bibler, J., Rossmann, M. G. & Johnson, J. E. (1993). *Acta Cryst.* D**49**, 580–587.