We wish to thank Mr Hilding Karlsson, Institute of Chemistry, University of Uppsala, for technical assistance during the measurements. This work was partly supported by project RPBP 01·6 (Polish Academy of Sciences).

References

BARTOSZAK, E. & JASKÓLSKI, M. (1988). *Acta Cryst.* Submitted.
CHANDRASEKHAR, K. & PATTABHI, V. (1980). *Acta Cryst.* B**36**, 2486–2488.
FURBERG, S. & SOLBAKK, J. (1972). *Acta Chem. Scand.* **26**, 2855–2862.
International Tables for X-ray Crystallography (1974). Vol. IV. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)
JASKÓLSKI, M. (1982). *Collected Abstracts of the Fourth Symposium on Organic Crystal Chemistry*, Poznań, September 1982, edited by Z. KAŁUSKI, pp. 70–71. Poznań: UAM.
JASKÓLSKI, M. (1987). *Acta Cryst.* C**43**, 1761–1763.
JASKÓLSKI, M., ALEJSKA, M. & WIEWIÓROWSKI, M. (1986). *J. Crystallogr. Spectrosc. Res.* **16**, 31–39.
JOHNSON, C. K. (1976). *ORTEP*II. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.
LEHMANN, M. S. & LARSEN, F. K. (1974). *Acta Cryst.* A**30**, 580–584.
LUNDGREN, J.-O. (1982). *Crystallographic Computer Programs.* Report UUIC-B13-4-05. Institute of Chemistry, Univ. of Uppsala, Sweden.
MOTHERWELL, W. D. S. & CLEGG, W. (1978). *PLUTO.* Program for plotting molecular and crystal structures. Univ. of Cambridge, England.
SHELDRICK, G. M. (1976). *SHELX*76. Program for crystal structure determination. Univ. of Cambridge, England.
TAKUSAGAWA, F. & KOETZLE, T. F. (1978). *Acta Cryst.* B**34**, 1910–1915.
TAKUSAGAWA, F. & KOETZLE, T. F. (1979). *Acta Cryst.* B**35**, 867–877.
WOO, N. H., SEEMAN, N. C. & RICH, A. (1979). *Biopolymers*, **18**, 539–552.

Acta Cryst. (1989). B**45**, 85–92

Structure Determination of Mengo Virus

BY MING LUO,* GERRIT VRIEND,† GREG KAMER AND MICHAEL G. ROSSMANN

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, *USA*

(*Received* 9 *May* 1988; *accepted* 16 *September* 1988)

Abstract

The structure of Mengo virus was determined to 3·0 Å resolution using human rhinovirus 14 as an initial phasing model at 8·0 Å resolution. Oscillation diffraction photographs were collected at the Cornell High Energy Synchrotron Source using orthorhombic Mengo virus crystals. The crystal space group was $P2_12_12_1$, $a = 441{\cdot}4$, $b = 427{\cdot}3$ and $c = 421{\cdot}9$ Å, with one icosahedral particle per asymmetric unit, giving 60-fold noncrystallographic redundancy. The orientations of the four viral particles in the unit cell were determined with a rotation function. Their positions relative to the crystallographic symmetry axes were found by a combination of Patterson-function analysis and a subsequent R-factor search using human rhinovirus 14 atomic coordinates as a model. The initial phases to 8·0 Å resolution were then computed by placing human rhinovirus 14 particles in the orientations and positions of Mengo virus particles. These phases were improved by ten cycles of real-space molecular replacement averaging. Phases between 8·0 and 3·0 Å resolution were obtained by molecular replacement phase extension. One or two reciprocal-space lattice points were used for each extension followed by two cycles of averaging.

Introduction

Mengo virus is a small icosahedral RNA animal virus with a molecular weight of $8{\cdot}5 \times 10^6$. About 30% of its mass is RNA. It belongs to the family of picornaviruses (Rueckert, 1986) and the genus cardiovirus. There are 60 copies of each of four different viral proteins (VP1, VP2, VP3 and VP4) in its coat. The course of the structure determination was similar to that of human rhinovirus 14 (HRV14) (Rossmann, Arnold, Erickson, Frankenberger, Griffith, Hecht, Johnson, Kamer, Luo, Mosser, Rueckert, Sherry & Vriend, 1985) and poliovirus (Hogle, Chow & Filman, 1985), although the structure itself is substantially different. VP1, VP2 and VP3 of Mengo virus have 12, 30 and 23% identical amino acids to HRV14; there are 47, 22 and 7 amino acids of Mengo virus that are insertions with respect to HRV14 and 62, 28 and 13 amino acids of HRV14 that are insertions with respect to Mengo virus, respectively. The Mengo virus structure has been

* Present address: Center for Macromolecular Crystallography, Department of Microbiology, University Station, University of Alabama, Birmingham, Alabama 35294, USA.
† Present address: Department of Structural Chemistry, University of Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands.

0108-7681/89/010085-08$03.00

described (Luo, Vriend, Kamer, Minor, Arnold, Rossmann, Boege, Scraba, Duke & Palmenberg, 1987) but few details were given of the crystallographic techniques. The latter are particularly interesting because phases were extended from 8 Å resolution to 3 Å resolution starting with a crude approximation derived from the structure of HRV14.

There is now quite an extensive history of phase extension by the use of noncrystallographic symmetry. However, it is only in the last three years that this technique has been used to bridge the gap from low resolution (6 Å or worse) to high resolution (3·5 Å or better). The Mengo virus determination has taken this procedure one step further than what had previously been accomplished.

The first suggestion that it might be possible to use noncrystallographic symmetry for *ab initio* phase determination was presented in reciprocal space (Rossmann & Blow, 1962, 1963; Crowther, 1969; Main & Rossmann, 1966; Main, 1967). However, the procedure was more readily understood and followed in real space. Initially, real-space averaging among similar noncrystallographically related objects was used only to improve phases. Gradually, however, phase extension was considered (Argos, Ford & Rossmann, 1975; Rossmann & Henderson, 1982; Rayment, 1983). The equivalence of the reciprocal- and real-space procedures was also established (Main, 1967; Bricogne, 1974; Argos & Rossmann, 1980; Arnold & Rossmann, 1986; Colman, 1974). *Ab initio* phase determination had been achieved at very low resolution (25 Å) for polyoma virus (Rayment, Baker, Caspar & Murakami, 1982) and southern bean mosaic virus (Johnson, Akimoto, Suck, Rayment & Rossmann, 1976). However, the work by Gaykema, Hol, Vereijken, Soeter, Bak & Beintema (1984), on the structure determination of hemocyanin, was particularly encouraging because phases were extended from 3·5 to 3·0 Å resolution. The first time phases were extended to high resolution from 6 Å was in the determination of HRV14, followed shortly afterward by the structure determination of poliovirus (Hogle *et al.*, 1985). In both cases the initial phasing start was provided by a series of poor low-resolution isomorphous heavy-atom derivatives.

The power, P, of the noncrystallographic phasing constraints is dependent on a variety of factors, but of particular importance are the noncrystallographic redundancy N, the fraction of observed data f in a specific resolution shell, the percentage accuracy of the structure amplitudes R (roughly equal to R_{merge}) and the volume (U) within which the noncrystallographic relationship holds compared to the crystal unit-cell volume V. It has been shown (Arnold & Rossmann, 1986) that these factors are related by the expression

$$P \propto \{(Nf)^{1/2}/[(U/V)R]\}.$$

The relative value of P is a crude measure of the quality of phase extension for any particular shell of reciprocal space. The larger P, the greater will be the precision of the derived phases.

In the case of the Mengo virus structure determination reported here, no isomorphous replacement was used. The initial structure approximation to 8 Å resolution was obtained by using the structure of HRV14 in light of its anticipated similarity to Mengo virus (Luo, Arnold, Erickson, Rossmann, Boege & Scraba, 1984). The orientations and positions of the four Mengo virions in the unit cell were determined with a rotation function, Patterson function and R-factor search. The phases between 8·0 and 3·0 Å resolution were then obtained by phase extensions using a similar procedure as had been used in the HRV14 structure determination.

Crystals and X-ray diffraction data

Orthorhombic crystals of Mengo virus were prepared as described by Luo *et al.* (1987). The $P2_12_12_1$ unit cell has dimensions of $a = 441{\cdot}4$, $b = 427{\cdot}3$, $c = 421{\cdot}9$ Å as determined by post-refinement (Schutt & Winkler, 1977; Rossmann, Leslie, Abdel-Meguid & Tsukihara, 1979). The volume per molecular weight, V_M (Matthews, 1968), was 2·34 Å^3 dalton^{-1} for four particles in the unit cell. Thus, there was one virus particle per asymmetric unit and the noncrystallographic redundancy was 60-fold. Data collection was by oscillation photography at the Cornell High Energy Synchrotron Source (CHESS). The crystals diffracted at least to 2·6 Å resolution in exposures lasting 1 to 2 min using a 0·3° oscillation angle. However, they were highly radiation sensitive. Furthermore, once irradiated even for only a few seconds, the crystals continued to deteriorate within minutes. Thus, conventional crystal setting procedures could not be used. Instead, the crystals were merely centered in the X-ray beam and used for a single data exposure. A procedure, generally known as the 'American Method' because it requires shooting first and asking questions later, was used in the work on HRV14 (Rossmann & Erickson, 1983) and developed further for the current work (Vriend, Rossmann, Arnold, Luo, Griffith & Moffat, 1986).* Once a reasonable crystal orientation had been derived, further processing was accomplished with the procedures described by Rossmann (1979) and Rossmann *et al.* (1979). An analysis of the final data set used in the structure determination is shown in Table 1. The data include reflections with partiality greater than 0·5. Reflections with a variety of problems (overloaded, overlapped, poor agreement between integrated and profile-fitted results, *etc.*) were flagged

* This technique has been recently superseded by an autoindexing method (Kim, 1989).

Table 1. *Data used in structure determination*

The number of unique observations [$F^2 > 1\sigma(F^2)$] and the percentage of total possible data are given.

Resolution range (Å)	Number observed	Percentage of possible data
∞–30	837	54
30–15	7684	71
15–10	21 239	72
10–7·5	40 196	70
7·5–5·0	157 716	67
5·0–3·5	396 425	62
3·5–3·0	263 549	46
3·0–2·75	87 801	19
2·75–2·6	19 634	3
Total	995 081	42
Number of film packs	143	
R factor (%)*	11·21	

* $R = [\sum_h \sum_l (I_h - I_{hl}) / \sum_h \sum_l I_{hl}] \times 100$, where I_h is the mean of the I_{hl} observations of reflection h.

and included in the final data set by reducing the weight on such reflections in averaging with better quality reflections [for further details on the procedure, see Arnold, Vriend, Luo, Griffith, Kamer, Erickson, Johnson & Rossmann (1987)].* Thus, the data shown in Table 1 essentially includes most of the reflections observed with $F^2 > 1\sigma(F)$. On average, a particular film has about 10% (at 5 Å resolution) or 30% (at 3 Å resolution) full reflections that are too weak to observe above 1σ. Thus, comparison with Table 1 shows that about 20% of the data was never recorded.

The crystal system was established as orthorhombic by observing the good agreement between orthorhombic symmetry-related structure amplitudes on individual films. Tetragonal symmetry (two of the cell dimensions are almost equal) was ruled out because of disagreement between fourfold-related reflections. The space group was presumed as $P2_12_12_1$, both because of the systematically small measured axial reflection amplitudes and because of the good close-packed arrangement of spheres. The structure determination later confirmed these observations.

Particle orientation and position

The orientations of Mengo virus particles in the unit cell were determined with a self-rotation function (Rossmann & Blow, 1962). Initially data were used between 20 and 13 Å resolution to search for twofold ($\kappa = 180°$), threefold ($\kappa = 120°$) and fivefold ($\kappa = 72$ and 144°) axes (Fig. 1). Two sets of icosahedrally related peaks were observed, each with a twofold axis roughly parallel to the crystallographic *b* axis.

When the self-rotation function calculation was repeated with data between 10 and 5 Å resolution, the peaks were split into two parts. This resulted in four sets of icosahedrally related peaks corresponding to the four virus particles in the unit cell (Fig. 2). Those icosahedral twofold axes that are roughly parallel to *b* were inclined to *b* by 1°.

The precise orientation of each viral particle in the crystal unit cell was obtained by making a least-squares fit between the directions of the observed self-rotation function peaks corresponding to one of the four icosahedra and the axes of an icosahedron placed initially into a standard orientation (Fig. 3). It was found that this standard icosahedron had to be rotated by $\kappa = 50{\cdot}090°$ about an axis given by $\psi = 1{\cdot}015°$ and $\varphi = 8{\cdot}800°$ [see Rossmann & Blow (1962) for definitions of κ, ψ, φ]. The r.m.s. deviation between the observed peak positions (two out of three independent fivefold positions averaged between $\kappa = 72$ and 144°, two out of four independent threefold positions and four out of six independent twofold positions) and the standard icosahedron was 0·34°. A stereoscopic view of the four virus particles in the orthorhombic unit cell is shown in Fig. 4. Table 2 gives the orientation of one of the particles in the unit cell.

As one of the icosahedral twofold axes was almost parallel (at least for data with resolution lower than 13 Å) to the crystallographic *b* axis, it was possible to compute a low-resolution Patterson in which the Harker section perpendicular to *b* would show the particle position. If the virus particle center is at *X,Y,Z* then there will be a crystallographically related virus at $-X,Y+\frac{1}{2},\frac{1}{2}-Z$ [using the origin as defined in *International Tables for Crystallography* (1983)]. Similarly, for an atom at *x,y,z* relative to the virus center there will be another identical atom at $\bar{x},y,\bar{z}$. Thus, considering vectors between the atoms at $X+x,Y+y,Z+z$, $X-x,Y+y,Z-z$, $-X+x,Y+y+\frac{1}{2},\frac{1}{2}-Z+z$ and $-X-x,Y+y+\frac{1}{2},\frac{1}{2}-Z-z$, it is easy to see that there will be a piling up of vectors at $2X,\frac{1}{2},2Z-\frac{1}{2}$. Indeed, a strong Harker peak was found at $\frac{1}{2},\frac{1}{2},0$ when a Patterson was calculated with data between 20 and 13 Å resolution (Fig. 5). This peak was 58% of the height of the origin peak. The next highest peak was only 9% of the origin peak. Thus, the reference particle center was close to $\frac{1}{4},Y,\frac{1}{4}$.

In order to obtain maximum separation of the particle centers within the unit cell, it is necessary to place them as far as possible from any of the 2_1 axes. Hence, the probable particle position had to be near $\frac{1}{4},\frac{1}{4},\frac{1}{4}$, which would produce a closest approach of 300·2 Å, comparable with the known 300 Å diameter of the particle. However, the reference particle center is not

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1MEV, R1MEVSF), and are available in machine readable form from the Protein Data Bank at Brookhaven or one of the affiliated centres at Melbourne or Osaka. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 37026 (31 microfiche). Free copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

restricted to any special position. Indeed, in a Patterson function at 3 Å resolution, the Harker peak was lower and split into two parts.

The precise positions of the virus particle centers were determined by an *R*-factor search using the homologous HRV14 atomic coordinates as a model. Four HRV14 particles were placed into the Mengo unit cell with appropriate orientations. Structure factors were then calculated for different particle centers (Crowther & Blow, 1967; Lattman, 1985). These calculations were performed rapidly by factoring out the particle center and performing the summation over all the atoms only once for each structure factor. An *R* factor was then calculated between the observed and calculated structure amplitudes.

The *R* factor was explored within a box defined by $\frac{a}{4} \pm 5$, $\frac{b}{4} \pm 10$, $\frac{c}{4} \pm 5$ Å. Although different resolution limits were tested, a plot of the *R* factor for all the data between 20 and 5 Å was used to best determine the particle center position. In retrospect, a rather lower resolution cutoff might have been better as data beyond 8 Å resolution showed essentially random *R* factor values. The background area gave an *R* factor of 58% or higher. A minimum of 54·2% in the *R* factor was found at $x = 0{\cdot}2517$, $y = 0{\cdot}2535$, $z = 0{\cdot}2474$ (Fig. 6). A plot of the *R* factor *versus* resolution when the

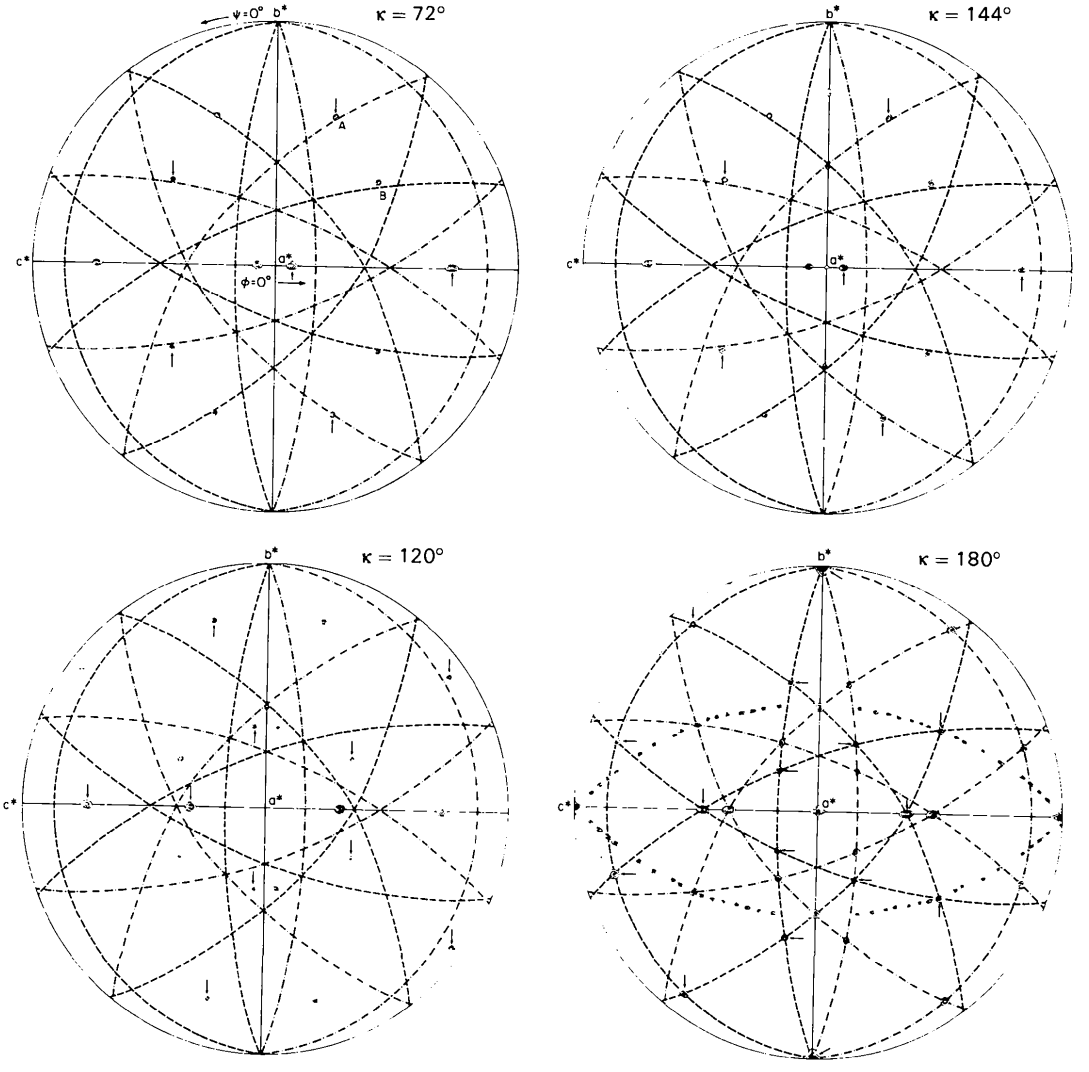

Fig. 1. Interpretation of the self-rotation function for data between 20 and 13 Å resolution when $\kappa = 180$, 120, 72 and 144°. Contours are drawn at levels 11, 12, 13, . . ., choosing an arbitrary scale. Mean background is zero. The r.m.s. deviation from background is two arbitrary units. There are two sets of icosahedrally organized twofold axes related by 222 crystallographic symmetry. Great circles connecting the twofold axes are marked with – – – for one of the icosahedra and with –·–·–·– for the other. Peaks corresponding to the former are marked with arrows. For instance, the marked peaks on the $\kappa = 72$ and 144° are surrounded by five arcs of great circles belonging to the icosahedron identified with – – – lines. The rotation function was computed using only the large terms to represent the second crystal. The number of large terms was 7·4% of the total observed set. The radius of integration was 300 Å. The peaks marked *A* and *B* on $\kappa = 72°$ are shown in greater detail at higher resolution in Fig. 2.

particle was at this position is shown in Fig. 7, which suggested that HRV14 was a useful model around 11 Å resolution.

Phase extension by molecular replacement averaging

The procedure followed for phase extension was similar to that already described in the structure determination of HRV14 (Arnold *et al.*, 1987). In brief, the electron density was averaged within the viral envelope over the 60 noncrystallographic asymmetric units. The solvent was flattened to the arbitrary (but convenient) value of zero in the solvent region outside the particle and in the RNA region inside the capsid protein. The choice of zero as solvent level was not entirely coincidental as the Fourier summation did not include an $F(000)$ term. The averaged electron density was then back-transformed to give a new set of calculated structure factors. The calculated structure factors were locally scaled to the observed data in 20 resolution shells. The mean amplitude of calculated reflections fell sharply at the edge of the current resolution limit until better phasing had been obtained. The quality and progress of the phase determination was ascertained primarily by the correlation coefficients, C, and R factors between observed F_o and calculated F_c structure amplitudes. These were defined as

$$C = \frac{\sum(F_o - \langle F_o\rangle)(F_c - \langle F_c\rangle)}{[\sum(F_o - \langle F_o\rangle)^2 \sum(F_c - \langle F_c\rangle)^2]^{1/2}}$$

$$R = \frac{\sum |(|F_o| - k|F_c|)|}{\sum |F_o|} \times 100.$$

The correlation coefficients appeared to be particularly useful as these were independent of the scale factor.

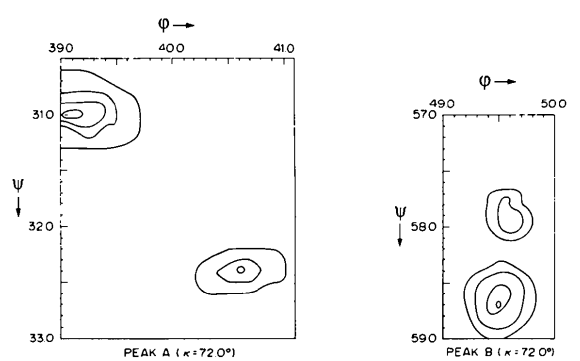

Fig. 2. An example of peak splitting in the higher-resolution self-rotation function for data between 10 and 5 Å resolution for $\kappa = 72°$ using a radius of integration of 300 Å. The number of large terms was 37% of the total observed set. The plots correspond to the peaks marked *A* and *B* in the low-resolution rotation function (for $\kappa = 72°$) shown in Fig. 1. Contours are drawn at levels 10, 11, 12, ... (arbitrary units). The mean background was zero. Note that the peak splitting is not evident at lower resolution.

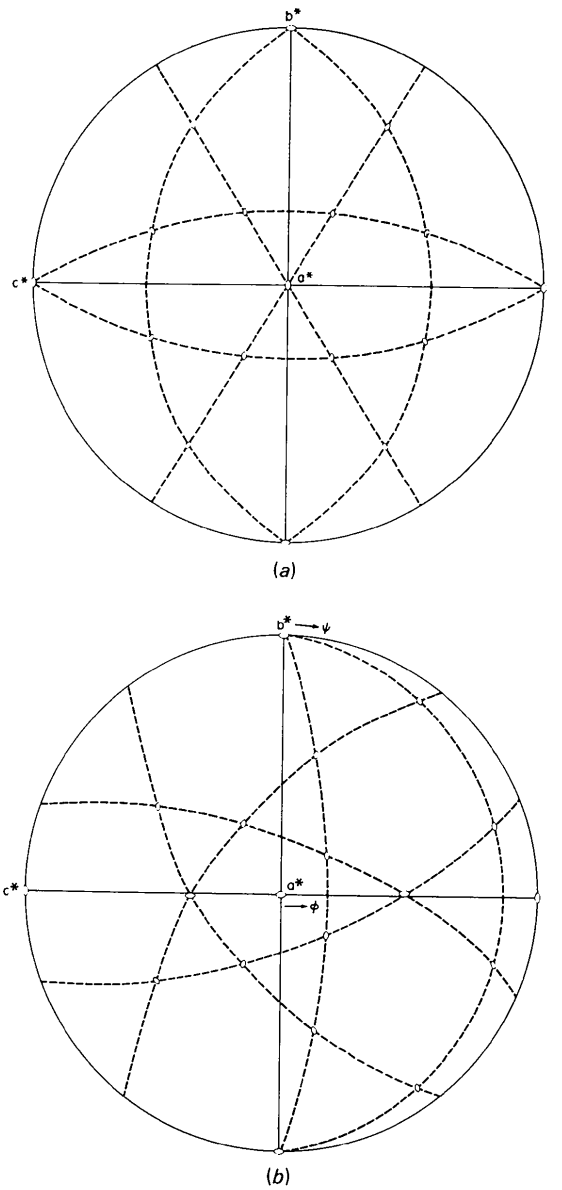

Fig. 3. Stereographic representation of the twofold axes of an icosahedron (*a*) in the defined standard orientation and (*b*) for a Mengo virion in the orthorhombic cell positioned at approximately $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}$.

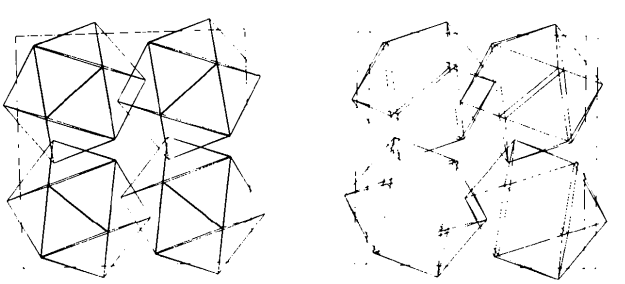

Fig. 4. A stereoscopic view of four virus particles in the unit cell of the orthorhombic space group.

Table 2. *Direction cosines of icosahedral axes with respect to crystal axes for particle at* $x = 0{\cdot}2517$, $y = 0{\cdot}2535$, $z = 0{\cdot}2474$

Application of the symmetry operators in the order shown here produces all the icosahedral symmetry elements.

(1) Threefold axis	0·873117	0·003357	−0·487499
(2) Twofold axis	0·766950	−0·012457	0·641586
(3) Fivefold axis	0·989765	−0·006205	0·142574
(4) Twofold axis	−0·004195	−0·999888	−0·014399

In subsequent cycles, F_c values were used in place of F_o amplitudes where no F_o measurements were available (Rayment, 1983). This greatly improved the correlation coefficients and R factors between the other reflections where F_o and F_c were both available. There is also a sound theoretical reason for such an inclusion (Main & Rossmann, 1966; Arnold & Rossmann, 1986). The F_o values used to compute electron density maps were weighted according to the scheme discussed by Arnold *et al.* (1987). Each phase extension was limited to an increase of only two reciprocal lattice points outward. The phase extension was, therefore, always within the first positive portion of a spherical diffraction function [the G function of Rossmann & Blow (1962)]. That is, $Hr < 0{\cdot}7$ where H is the phase extension in reciprocal space and r is the external particle radius.

During the initial averaging procedure, the electron density was computed on a grid using about one-fifth of the resolution with an eight-point linear interpolation to determine electron density at nonintegral grid points. This required a great deal of space for storing both the electron density and the list of noncrystallographically related points for subsequent sorting. To reduce the amount of computing, the density was averaged over a subset of noncrystallographically related points (Table 3). The precise subset was varied from cycle to cycle to assure that each part of the crystallographic asymmetric unit was treated equivalently. However, this procedure was superseded at higher resolution when the larger number of grid points became unmanageable. Beyond 4·1 Å resolution the electron density was calculated on a grid corresponding to less than one-third of the resolution. Interpolation to find the density at a nonintegral grid point was then done with a vectorized adaptation of the 11-point approximation technique described by Nordman (1980). This procedure reduces, by an order of magnitude, the number of electron density values that need to be stored in comparison to the linear interpolation procedures recommended by Bricogne (1974). It also permitted averaging over all 60 noncrystallographically equivalent positions as opposed to the earlier approximation. A similar procedure was used by Hogle *et al.* (1985) in the structure determination of poliovirus. With the decreased number of stored grid densities and the larger available memory of modern computers, it was possible to avoid the time-consuming double sorting used by

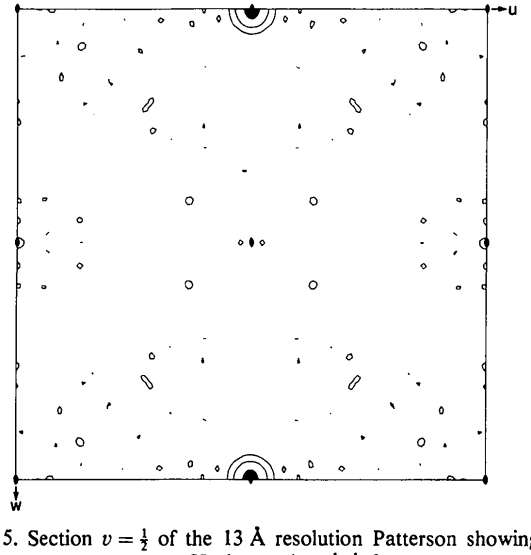

Fig. 5. Section $v = \frac{1}{2}$ of the 13 Å resolution Patterson showing a strong Harker peak at $\frac{1}{2}$, $\frac{1}{2}$, 0.

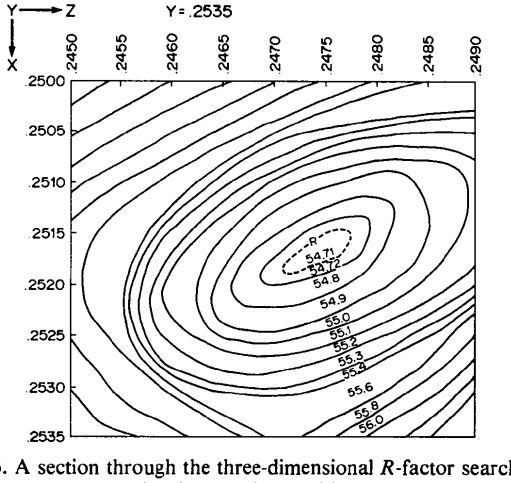

Fig. 6. A section through the three-dimensional R-factor search for the virus particle position.

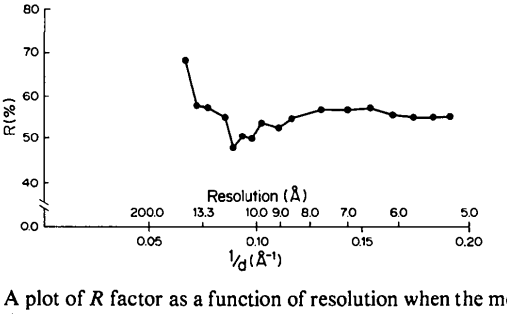

Fig. 7. A plot of R factor as a function of resolution when the model virus particle was positioned at 0·2517,0·2535,0·2474.

Table 3. *Conditions for phase extension and phase improvements*

Cycle No.	Outer resolution limit	No. of incremental steps in extension from previous resolution limit	No. of reciprocal lattice points extended per step	Electron density map sampling as fraction of resolution	No. of cycles of averaging at each resolution step	No. of noncrystallographic asymmetric units used for averaging
0–10	8·00	0	0	1/5	10	20
11–20	6·7726	5	2	1/5	2	20
21–24	6·3810	4	1	1/5	1	20
25–28	6·0322	2	2	1/5	2	20
29–38	6·0322	0	0	1/5	10	20
39–72	4·1186	17	2	1/5	2	10
73–90	3·5264	9	2	1/3·0	2	60
91–100	3·5264	0	0	1/2·8	10	60
101–140	3·000	10	2	1/2·5	2 or more	60
141–150	3·000	0	0	1/2·2	10	60

Bricogne (1974) and Johnson (1978) at a cost of using more time in the nonlinear interpolation. Either the whole density was stored in packed form in the earlier stages of the determination or no less than one-half of the asymmetric unit of the cell was kept in random-access memory. In the final stages of phase extension two different passes were made, each time with a different stored density, in order to associate all the noncrystallographically related points for averaging. The final grid used for averaging the 3 Å resolution map was 320 × 320 × 320, corresponding to 1·38, 1·34, 1·32 Å per grid step in the *a*, *b* and *c* directions.

The initial phases between 15 and 8 Å resolution were obtained from the model based on the HRV14 structure. It was improved by ten cycles of molecular replacement real-space averaging. An envelope containing only the protein capsid was used to exclude the solvent and RNA region in the unit cell from the volume of electron density being averaged. The envelope was composed of an outer sphere with radius of 170 Å and an inner sphere with radius of 80 Å (the inner radius was changed to 100 Å in extending phases from 4·1 to 3·0 Å resolution). A set of 12 planes was placed perpendicular to the virus radius vector, which limited the spherical envelope. The planes were positioned at the midpoint of the interparticle vectors connecting the closest neighboring particles. The calculated structure factors from the Fourier back transform were not included during the phase improvement at 8 Å resolution, in order to avoid the bias of the HRV14 structure model. The *R* factor dropped from 54% before averaging to 23% after averaging and the correlation coefficient increased from 0·54 before averaging up to 0·87 after averaging (Fig. 8).

The phases beyond 8 Å resolution were obtained by gradual extension of the phases from low resolution to high resolution. The outside resolution of each step is shown in Table 3. A total of 150 cycles of averaging were computed in extending phases from 8 Å resolution to 3 Å resolution. Immediately following each phase extension, one or two (but more beyond 3·5 Å resolution) cycles of averaging were performed to refine the newly obtained phases before further extension. The correlation coefficient and *R* factor continuously improved at lower resolution following cycles of further phase extension (Fig. 8). It is important to be conservative in the use of an initial phasing set. If too high a resolution is used, where phases may be largely random, problems can arise in the relative definition of hand, as was the case for an initial abortive phase determination of HRV14 (Arnold *et al.*, 1987) and for an earlier determination of turnip crinkle virus using the tomato bushy stunt virus structure (Hogle, Maeda & Harrison, 1986). The number of cycles required for the refinement of phases at a given resolution is a function of the phasing power (see above). For instance, in the structure determination of HRV1A (unpublished results), where there was only tenfold redundancy and only 55% of the data had been observed, it was necessary to do at least 12 cycles of phase improvement at each resolution step. As is always the case, the correlation coefficients decrease near the edge of the current resolution as a consequence of the absence of phase information just outside the current cutoff. This effect is worse at higher resolution because of the poor quality and fewer available data. Phase extension must stop when the phasing power becomes too small. Phase extension could then be resumed when more or better data are measured.

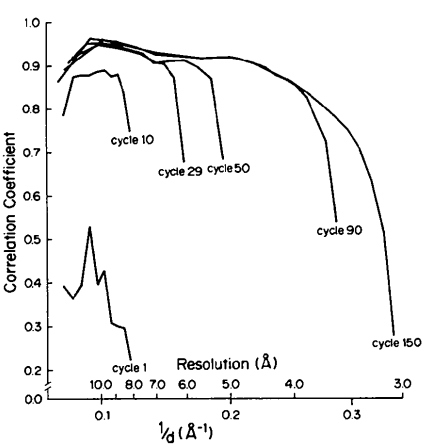

Fig. 8. Plot of correlation coefficient as the phases were extended from 8 to 3 Å resolution.

The main chain could be traced at 6 Å resolution in light of the anticipated similarity of the structure to HRV14 and other icosahedral RNA viruses. However, this tracing included, for instance, recognition of the two insertions (loop I and loop II) which are peculiar to the VP1's of cardioviruses. The complete main-chain tracing and side-chain identification were achieved with the aid of a minimap at 3·5 Å resolution. The atomic model was initially built into the 3·5 Å electron density map and later improved with respect to the far superior 3 Å resolution map (Fig. 9), in which carbonyl oxygens were mostly evident, with an Evans and Sutherland PS300 computer graphics system and *FRODO* program (Jones, 1978, 1982). Map 'quality' is largely a subjective assessment. However, the map at 3 Å resolution was comparable to the corresponding map of HRV14 (Arnold *et al.*, 1987).

The samples of virus for crystallization, which made this whole project possible, were supplied to us by Drs Ulrike Boege and Doug Scraba of the University of Alberta. We are greatly appreciative of their collaboration, interest and support. We wish to thank Drs Ann Palmenberg and Greg Duke of the University of Wisconsin for the determination of the Mengo virus RNA sequence, which allowed us to confidently interpret the electron density map; Edward Arnold for useful discussions and help with data collection; and Iwona Minor, Jun Tsao and S. Krishnaswamy for help in data collection, film processing and production of Fig. 9, respectively. We are grateful to the staff and operators of the Cornell High Energy Synchrotron Source for their kind help and support. We thank Sharon Fateley and Sharon Wilder for help in the preparation of the manuscript. The work was supported by grants from the National Institutes of Health and the National Science Foundation to MGR.

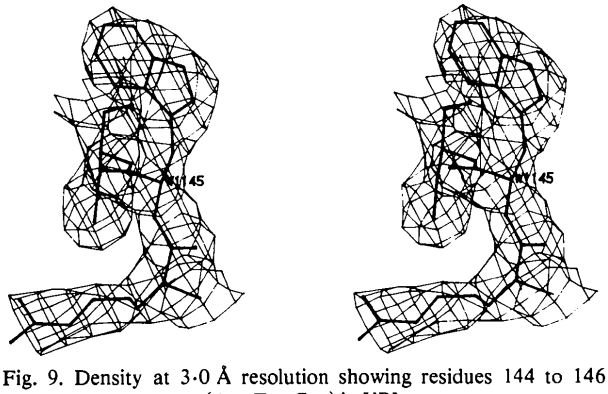

Fig. 9. Density at 3·0 Å resolution showing residues 144 to 146 (Arg-Trp-Cys) in VP1.

References

ARGOS, P., FORD, G. C. & ROSSMANN, M. G. (1975). *Acta Cryst.* A**31**, 499–506.

ARGOS, P. & ROSSMANN, M. G. (1980). *Theory and Practice of Direct Methods in Crystallography*, edited by M. F. C. LADD & R. A. PALMER, pp. 361–417. New York: Plenum Press.

ARNOLD, E. & ROSSMANN, M. G. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 5489–5493.

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.

BRICOGNE, G. (1974). *Acta Cryst.* A**30**, 395–405.

COLMAN, P. M. (1974). *Z. Kristallogr.* **140**, 344–349.

CROWTHER, R. A. (1969). *Acta Cryst.* B**25**, 2571–2580.

CROWTHER, R. A. & BLOW, D. M. (1967). *Acta Cryst.* **23**, 544–548.

GAYKEMA, W. P. J., HOL, W. G. J., VEREIJKEN, J. M., SOETER, N. M., BAK, H. J. & BEINTEMA, J. J. (1984). *Nature (London)*, **309**, 23–29.

HOGLE, J. M., CHOW, M. & FILMAN, D. J. (1985). *Science*, **229**, 1358–1365.

HOGLE, J. M., MAEDA, A. & HARRISON, S. C. (1986). *J. Mol. Biol.* **191**, 625–638.

International Tables for Crystallography (1983). Vol. A. Dordrecht: D. Reidel.

JOHNSON, J. E. (1978). *Acta Cryst.* B**34**, 576–577.

JOHNSON, J. E., AKIMOTO, T., SUCK, D., RAYMENT, I. & ROSSMANN, M. G. (1976). *Virology*, **75**, 394–400.

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

JONES, T. A. (1982). *Computational Crystallography*, edited by D. SAYRE, pp. 303–317. Oxford: Clarendon Press.

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

LATTMAN, E. (1985). *Methods Enzymol.* **115**, 55–77.

LUO, M., ARNOLD, E., ERICKSON, J. W., ROSSMANN, M. G., BOEGE, U. & SCRABA, D. G. (1984). *J. Mol. Biol.* **180**, 703–714.

LUO, M., VRIEND, G., KAMER, G., MINOR, I., ARNOLD, E., ROSSMANN, M. G., BOEGE, U., SCRABA, D. G., DUKE, G. M. & PALMENBERG, A. C. (1987). *Science*, **235**, 182–191.

MAIN, P. (1967). *Acta Cryst.* **23**, 50–54.

MAIN, P. & ROSSMANN, M. G. (1966). *Acta Cryst.* **21**, 67–72.

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

NORDMAN, C. E. (1980). *Acta Cryst.* A**36**, 747–754.

RAYMENT, I. (1983). *Acta Cryst.* A**39**, 102–116.

RAYMENT, I., BAKER, T. S., CASPAR, D. L. D. & MURAKAMI, W. T. (1982). *Nature (London)*, **295**, 110–115.

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

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). *Nature (London)*, **317**, 145–153.

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

ROSSMANN, M. G. & BLOW, D. M. (1963). *Acta Cryst.* **16**, 39–45.

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

ROSSMANN, M. G. & HENDERSON, R. (1982). *Acta Cryst.* A**38**, 13–20.

ROSSMANN, M. G., LESLIE, A. G. W., ABDEL-MEGUID, S. S. & TSUKIHARA, T. (1979). *J. Appl. Cryst.* **12**, 570–581.

RUECKERT, R. R. (1986). *Fundamental Virology*, edited by B. N. FIELDS & D. M. KNIPE, pp. 357–390. New York: Raven Press.

SCHUTT, C. & WINKLER, F. K. (1977). *The Rotation Method in Crystallography*, edited by U. W. ARNDT & A. J. WONACOTT, pp. 173–186. Amsterdam: North-Holland.

VRIEND, G., ROSSMANN, M. G., ARNOLD, E., LUO, M., GRIFFITH, J. P. & MOFFAT, K. (1986). *J. Appl. Cryst.* **19**, 134–139.