HUBER, R. (1969). *Crystallographic Computing Procedures*, edited by F. R. AHMED, pp. 96-102. Copenhagen: Munksgaard.
JACK, A. & LEVITT, M. (1978). *Acta Cryst.* A**34**, 931-935.
JAMES, M. N. G., DELBAERE, L. T. J. & BRAYER, G. D. (1978). *Can. J. Biochem.* **56**, 396-402.
JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268-272.
KATANUMA, N., KOMINAMI, E., KOBAYASHI, K., BANNO, Y., SUZUKI, K., CHICHIBY, K., HAMAGUCHI, Y. & KATSUNUMA, T. (1975). *Eur. J. Biochem.* **52**, 37-50.
MATTHEWS, B. W. (1977). *The Proteins*, Vol. 3, 3rd ed., edited by H. NEURATH & R. L. HILL, pp. 403-590. New York: Academic Press.
MATTHEWS, B. W., KLOPFENSTEIN, C. K. & COLMAN, P. M. (1972). *J. Phys. E*, **5**, 353-359.
MATTHEWS, B. W., SIGLER, P. B., HENDERSON, R. & BLOW, D. M. (1967) *Nature (London)*, **214**, 652-656.
MOLNAR, C. E., BARRY, C. D. & ROSENBERGER, F. U. (1976). Tech. Mem. No. 229, Computer Systems Laboratory, Washington Univ., St Louis.
ROSSMANN, M. G. (1979). *J. Appl. Cryst.* **12**, 225-238.
ROSSMANN, M. G. & ARGOS, P. (1976). *J. Mol. Biol.* **105**, 76-96.
ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24-31.
SCHMID, M. F., WEAVER, L. H., HOLMES, M. A., GRÜTTER, M. G., OHLENDORF, D. H., REYNOLDS, R. A., REMINGTON, S. J. & MATTHEWS, B. W. (1981). *Acta Cryst.* A**37**, 701-710.
SUSSMAN, J. L., HOLBROOK, S. R., CHURCH, G. M. & KIM, S. (1977). *Acta Cryst.* A**33**, 800-804.
WOODBURY, R. G., GRUZENSKI, G. M. & LAGUNOFF, O. (1978). *Proc. Natl Acad. Sci. USA*, **75**, 2786-2789.
WOODBURY, R. G., KATANUMA, N., KOBAYASHI, K., TITANI, K. & NEURATH, H. (1978). *Biochemistry*, **17**, 811-819.
YOSHIDA, N., EVERITT, M. T., NEURATH, H., WOODBURY, R. G. & POWERS, J. (1980). *Biochemistry*, **19**, 5799-5804.

Acta Cryst. (1985). B**41**, 147-157

The Refinement of Southern Bean Mosaic Virus in Reciprocal Space

BY ABELARDO M. SILVA* AND MICHAEL G. ROSSMANN

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

(*Received* 26 *June* 1984; *accepted* 5 *October* 1984)

Abstract

The restrained least-squares refinement procedure of Konnert & Hendrickson [*Acta Cryst.* (1980), A**36**, 344-350] has been applied to 2·8 Å resolution, native, southern bean mosaic virus diffraction data. The initial model, based on a multiple-isomorphous-replacement map which had been improved by three cycles of real-space molecular-replacement averaging, gave an overall R factor of 40·7%. The R factor decreased to 25·5% after 65 refinement cycles interspersed with model checks on a computer graphics system using the *FRODO* program [Jones (1978). *J. Appl. Cryst.* **11**, 268-272]. The refinement was made possible by the vectorization of the program code, the use of non-crystallographic symmetry as a constraint and the selection of partial data sets. The latter was possible as a consequence of the non-crystallographic symmetry. A total of 34 water molecules per icosahedral asymmetric unit were identified. A region of relatively low, but continuous, electron density can be seen near a series of basic residues on the internal viral surface. This might be a small portion of RNA with partial icosahedral symmetry.

I. Introduction

(*A*) *The virus*

Southern bean mosaic virus (SBMV) is a small icosahedral plant virus of diameter varying from 328 Å at the fivefold axes to about 285 Å at the 'quasi'-threefold axes. The triangulation number of the capsid is $T = 3$, according to the nomenclature of Caspar & Klug (1962). There is only one type of coat protein (molecular weight 28 200), consisting of 260 amino acids. The amino acid sequence was determined by Hermodson, Abad-Zapatero, Abdel-Meguid, Pundak, Rossmann & Tremaine (1982). The encapsidated RNA has an approximate molecular mass of $1{\cdot}4 \times 10^6$ daltons which accounts for about 21% of the total molecular mass of the particle. Measurements of the Ca^{2+} and Mg^{2+} content vary from 80 to 280 and from 120 to 380 ions per virion, respectively (Abdel-Meguid, Yamane, Fukuyama & Rossmann, 1981; Hsu, Sehgal & Pickett, 1976; Hull, 1978).

The molecular structure of the cowpea strain of SBMV has been determined by X-ray diffraction techniques at a resolution of 2·8 Å (Abad-Zapatero, Abdel-Meguid, Johnson, Leslie, Rayment, Rossmann, Suck & Tsukihara, 1980, 1981; Rossmann, Abad-Zapatero, Hermodson & Erickson, 1983). The space group of the crystals is $R32$ with hexagonal

* Present address: Departamento de Física, Facultad de Ciencias Exactas, U.N. de La Plata, C. C. No. 67, 1900 La Plata, Argentina.

0108-7681/85/020147-11$01.50

cell constants $a = 334{\cdot}30$ and $c = 757{\cdot}95$ Å and three particles in the hexagonal unit cell. The particles sit on special positions incorporating the 32 space-group symmetry as part of their icosahedral symmetry. Hence, out of the 180 coat protein subunits that constitute the capsid, only 30 are crystallographically independent. The non-crystallographic icosahedral fivefold and twofold axes further reduce the number of independent subunits to three (subunits *A*, *B* and *C*) per non-crystallographic asymmetric unit. These three subunits are related by a quasi-threefold axis. Fig. 1 shows the disposition of the subunits with respect to the symmetry axes.

The crystal packing induces alignment of the viruses along their fivefold axes, where there is interdigitation of neighboring particles (Fig. 2). Hence, the rhombohedral cell axes are coincident with a set of fivefold icosahedral axes. The rhombohedral cell angle was, therefore, adjusted by less than 0·2° to a value of 63·4349° as is required by a perfect icosahedron. There are two kinds of fivefold axes in the crystal, those immediately around the crystallographic threefold axis which are involved in interparticle contacts and the equatorial ones where the particle makes no contacts.

Quasi-equivalent subunits *A*, *B* and *C* show the same β-roll folding pattern (Fig. 3) - a double layer of beta sheets consisting of eight (*A* and *B* subunits) or nine (*C* subunits) antiparallel strands connected by turns and five helices. The initial model extended from residue 63 to residue 260 in subunit *A*, from 65 to 260 in subunit *B*, and from 39 to 260 in subunit *C*, together with four metal ions tentatively assigned as Ca^{2+} (Rossmann *et al.*, 1983). While all three subunits showed crystallographic disorder in the basic amino-terminal section of the polypeptide, the strand βA was observed only in the *C* subunit. The extended βA strand is involved in the CC_2 (Fig. 1) subunit contacts where there is additional space compared to the AB_5 contact region. Lower electron density in the RNA region could not be interpreted. This region contains not only the RNA but also the amino-terminal segments of the polypeptides. Either there is no unique structure or the RNA and associated protein has no icosahedral symmetry. While the space group imposes only 32 symmetry averaging on the contents of the particle, nevertheless the packing (as seen above) is controlled by icosahedral symmetry. This guarantees at least 30-fold averaging of the capsid content.

There were, therefore, a total of 616 amino acids and 4 Ca atoms per icosahedral asymmetric unit in the original model, leading to 46 670 atoms in the crystallographic asymmetric unit. The number of independent reflections with $F > \sigma$ for the native crystals was 298 615 within the 2·8 Å resolution sphere, or about 90% of all possible reflections. This large number of atoms and reflections required special procedures for refinement as described here. The results are examined here only to show conver-

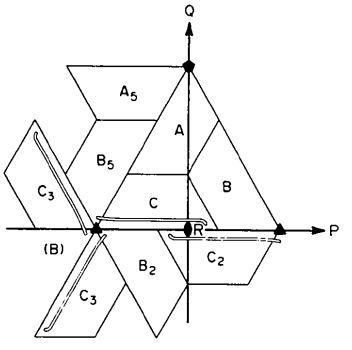

Fig. 1. Arrangement of the quasi-equivalent *A*, *B* and *C* subunits within one icosahedral asymmetric unit and the right-handed coordinate system *P*, *Q*, *R* to which all atomic coordinates have been referred. The origin is at the virus center. The diagram also shows the rough outline of subunits and the nomenclature used here to reference them. The view is from outside the virus looking toward the center. Hence, *R* is coming up, out of the page. The reference subunits *A*, *B* and *C* are unsubscripted. The subscript 5 implies an anticlockwise rotation of 72° about the shown fivefold axis when viewed from the origin. The subscript 2 implies a 180° rotation about the twofold axis (*R*) and the subscript 3 refers to an anticlockwise rotation about the threefold axis on the left half of the diagram. The parenthetic (*B*) refers to a subunit which is generated by a threefold rotation on B_5 or B_2. The subunit contact CC_2 implies contacts between the two subunits related by the *R* axis. Similarly, the AB_5 contacts are a consequence of the quasi-twofold axis.

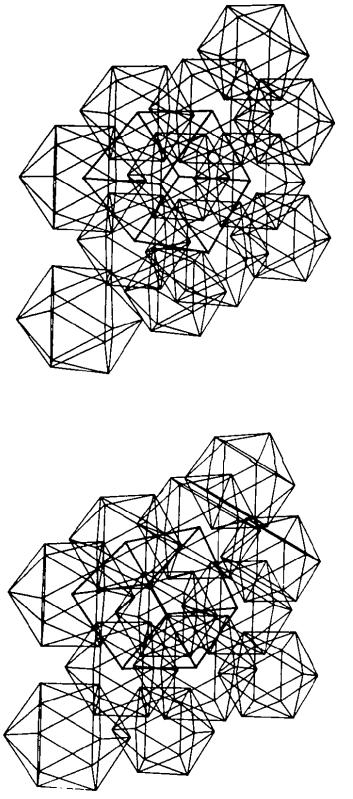

Fig. 2. View down the threefold axis of the type II crystals displaying the precise alignment of the fivefold particle axes.

gence and quality of the stereochemistry; an analysis of the refined structure will be given elsewhere (Silva & Rossmann, work in progress).

(*B*) *Quality of the diffraction data*

Two X-ray diffraction data sets had been collected for native crystals on oscillation photographs (Abad-Zapatero *et al.*, 1981). The first set was collected on 1·0° oscillation photographs and processed to 3·5 Å resolution, while the second was collected on 0·5° oscillation photographs and processed to 2·8 Å resolution. The two data sets were scaled and combined (Abad-Zapatero *et al.*, 1981) and the standard errors were estimated as described by Rossmann, Leslie, Abdel-Meguid & Tsukihara (1979). The $|F_o|$ values of the merged data were compared with their expected mean values, which were assumed to be proportional to $f_C e^{-Bs^2/4}$, by means of an R factor defined as

$$R = \frac{\sum |(|F_o| - Cf_C\, e^{-Bs^2/4})|}{\sum |F_o|},$$

where f_C is the atomic scattering factor for carbon. The constant C was obtained by least squares. The temperature factor B was an overall value for the structure, taken as 20 Å^2. The R factor, plotted as a function of resolution (Fig. 4), showed a discontinuity at 3·5 Å resolution, the point of junction between the two data sets. A similar discontinuity at the same resolution was seen in the distribution of estimated error (σ), although no discontinuity was visible in the distribution of mean $|F_o|$. It was therefore apparent that the observed data contained a systematic error, probably introduced when the two data sets were scaled and merged. Hence, the two data sets were re-scaled by scaling both to a single exponential 'scattering-factor' curve. Neither the R-factor plot (Fig. 4) nor the σ distribution showed a discontinuity after this readjustment. Furthermore, the mean values of the adjusted $|F_o|$'s showed better agreement with their expected mean values as a function of the resolution. These data were then used in the refinement procedure.

One other problem was evident: the data beyond 3·0 Å resolution showed a progressive departure from the anticipated distribution function. Clearly, major errors are present in these high-resolution data, probably due to program faults in the processing procedure such as the inclusion of pseudo reflections beyond the edge of the film. To avoid these spurious data, the observed reflection list was terminated at 2·88 Å resolution.

(*C*) *Crystallographic refinement*

The refinement of a structure in reciprocal space is performed by minimizing the residual

$$R = \sum_{\mathbf{h}} [|F_o(\mathbf{h})| - |F_c(\mathbf{h})|]^2 \qquad (1)$$

as a function of the atomic parameters.

The residual in (1) can be re-written as

$$R = \sum_{\mathbf{h}} |\Delta F_{\mathbf{h}}|^2 = \sum_{\mathbf{h}} \Delta F_{\mathbf{h}}\, e^{i\varphi_c^{\mathbf{h}}} \Delta F_{\mathbf{h}}\, e^{-i\varphi_c^{\mathbf{h}}}$$

$$= \sum_{\mathbf{h}} |[|F_o(\mathbf{h})|\, e^{i\varphi_c^{\mathbf{h}}} - F_c(\mathbf{h})]|^2, \qquad (2)$$

where $\varphi_c^{\mathbf{h}}$ is the phase associated with the calculated structure factor $F_c(\mathbf{h})$. Applying Parseval's theorem to (2), it follows that

$$R = V \int_V [\rho_o(\mathbf{r}) - \rho_c(\mathbf{r})]^2\, \mathrm{d}V, \qquad (3)$$

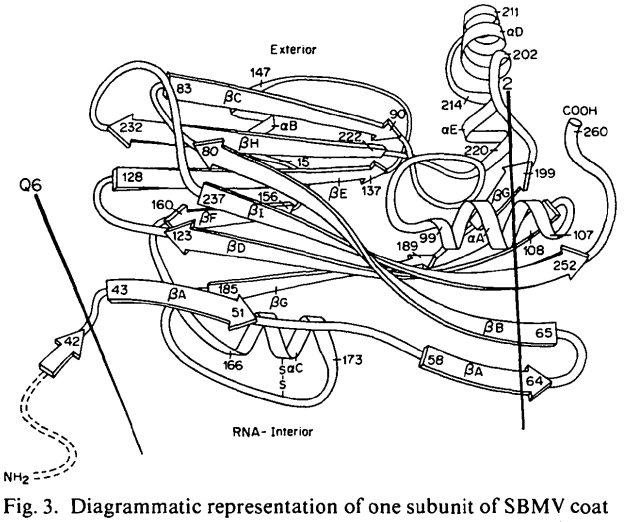

Fig. 3. Diagrammatic representation of one subunit of SBMV coat protein. Secondary structural elements and approximate sequence numbers are shown. (The diagram is based on an original drawing by Jane Richardson.)

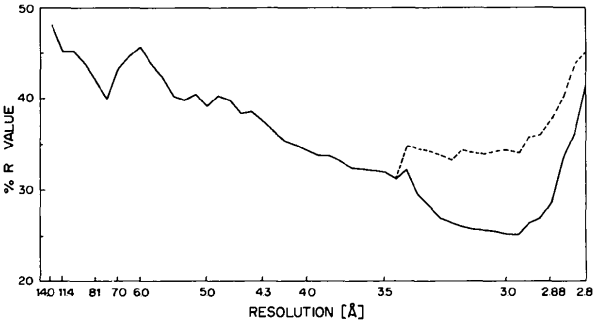

Fig. 4. R value between the observed structure factors and their expected mean values $Cf_C\, e^{-Bs^2/4}$, where C is an overall constant adjusted by least squares, f_C is the scattering factor for carbon and B is the overall temperature factor. The decreasing trend of the curve with resolution reflects the smaller dispersion of the intensities at higher resolution. The dashed line represents the values of R between 3·5 and 2·80 Å before correcting for the systematic error introduced when the two data sets were merged. The curve also indicates the poor quality of the data near the 2·80 Å limit.

where V is the volume of the unit cell, $\rho_o(\mathbf{r})$ is the electron density computed with the observed amplitudes and calculated phases and $\rho_c(\mathbf{r})$ is the electron density associated to the model. This means that the parameter shifts, which tend to minimize the residual (1) in reciprocal space, can be interpreted in direct space as moving the model to maximize its fitting with the 'observed' electron density. During refinement, ρ_o is implicitly being updated through the phases associated to the changing molecular model. In other words, reciprocal-space refinement produces a simultaneous convergence of model and map toward a maximum agreement.

There are four reciprocal-space least-squares procedures which are frequently used for macromolecular refinement. Hendrickson & Konnert (1980) introduce restraints to the geometry of the model dependent on known stereochemical information as a way to increase the overdeterminancy of the system. Sussman, Holbrook, Church & Kim (1977) applied restraints and constraints to increase the overdeterminancy and to reduce the number of variables. Agarwal (1978) used FFT methods to compute the elements of the normal matrix and the gradient vector in alternate cycles of geometry regularization and structure factor residual minimization. Jack & Levitt (1978) also used FFT techniques but introduced an energy-minimization algorithm to attain a normal geometry and increase overdeterminancy. Their method is, therefore, identical to the Konnert-Hendrickson procedure in principle, except that the energy terms are continuous functions.

Refinement in direct space is based on the minimization of (3) as a function of ρ_c, usually with constraints on the stereochemistry (Diamond, 1971; Jones & Liljas, 1984*a*, *b*). In this case, ρ_o is kept fixed while the model is being moved. The behavior of the refinement critically depends on the quality of the initial map and therefore on the starting phases. After several cycles the map is recomputed with the phases derived from the new model. The behavior of reciprocal- and direct-space refinement procedures should be similar if in the latter ρ_o were updated at each cycle with the phases associated with the new model. The great advantage of direct-space methods is that the integral in (3) can be broken up into the sum of integrals. Each of these is then a function of different subsets of atomic parameters. Thus, the minimization of R can be achieved by successive local minimizations. In (1) the sum cannot be dissociated since each structure factor is a function of all atomic parameters.

In reciprocal-space refinement, non-crystallographic symmetry is generally used as a restraint (Hendrickson & Konnert, 1980). Thus, all atoms in the asymmetric unit are considered as independent, although extra observational equations are introduced to link atoms by the non-crystallographic symmetry. However, if the non-crystallographic symmetry is assumed to be exact and accurately known, the problem then becomes overdimensioned in terms of the number of variables to be refined.

In direct space the refinement can be done against an electron density map already averaged over the different non-crystallographically related parts. Jones & Liljas (1984*a*) refined satellite tobacco necrosis virus (STNV) using this principle. They computed averaged $2F_o - F_c$, α_c, electron density maps, and refined these maps by use of molecular-replacement real-space averaging. This map then played the role of ρ_o in (3) which was minimized by rigid-body automatic fitting alternating with cycles of geometry regularization.

For SBMV, we used the well documented Konnert-Hendrickson reciprocal-space refinement package with modifications to take explicit advantage of the non-crystallographic symmetry of the viral capsid thereby reducing the size of the problem.

II. Methods

In the initial stages of this project, we had available a computer system dependent on two CDC 6500's and one CDC 6600. Our initial efforts were thus to prepare programs for that system. Later, we also had access to the more powerful Cyber 205 computer which allowed the refinement of SBMV in reciprocal space.

The refinement followed an essentially standard path. The initial model was subjected to a number of refinement cycles. Structure factors were then computed and a $(2F_o - F_c)$, α_c, map was calculated and averaged over the ten non-crystallographic asymmetric units. The model was corrected against the new map on a graphics system (MMS-X; Barry, Bosshard, Ellis & Marshall, 1975) using the program *FRODO* (Jones, 1978). Further cycles of refinement were then performed and the whole process was repeated several times. In later stages the model was checked against the unaveraged map in order to save time-consuming calculations. While this is standard practice for small and medium size proteins, it is a very complex and time-consuming operation for a problem of the size of a virus. The most relevant changes introduced in the overall process are discussed next.

(*A*) *Structure factor calculations*

We developed a fast structure factor calculation technique (Ten Eyck, 1973, 1977) on the 6600 system which was used in the initial stages of this work. We report some of our observations here primarily to show the advantages and problems with the process.

The structure factors were calculated by Fourier transformation of a model map (Ten Eyck, 1977)

built by using a two-Gaussian approximation for the atomic density (Agarwal, 1978). Computer limitations required that the sampling interval was only slightly better than one-half of the resolution limit of the diffraction data. The total c.p.u. time was about 8000 s on a CDC 6400 for about 3×10^5 reflections. The c.p.u. time taken by the Cyber program was 15 000 s for the complete calculations. We estimate that a vectorized FFT method on the Cyber 205 using a higher sampling density (thereby improving the quality of the approximation) would have taken a c.p.u. time of less than 1000 s for the same set of structure factors.

The quality of the FFT computed structure factors was later assessed by comparison with those calculated by the vectorized summation of trigonometric functions in the Cyber 205 computer. The two types of calculated structure factors were compared by means of an R factor whose overall value was 9·34% for 298 615 reflections (Fig. 5). The high R factor at low resolution is probably a consequence of the limited atomic radius used in building the model density. The overall height and increasing R factor with resolution is due to 'aliasing'.

(*B*) *Electron density maps and model examination*

All Fourier transformations were performed in the 6600 system using an optimized *P*1 routine. In the earlier stages of the refinement process (up to cycle 36–see Table 3), the maps were skewed and averaged (Johnson, 1978) over the 10 non-crystallographic asymmetric units. For this purpose, the maps were computed at a spacing of one-fifth of the resolution limit of the data (Bricogne, 1974) because linear interpolation was used in the averaging process. The c.p.u. time required to calculate one crystallographic asymmetric unit sampled at one-fifth of the resolution was 90 000 s on a CDC 6400. In the final stages of refinement, however, unaveraged maps were used for checking the model. The c.p.u. time could then be reduced to 6000 s by calculating only one non-crystallographic asymmetric unit on a grid of one-third of the resolution.

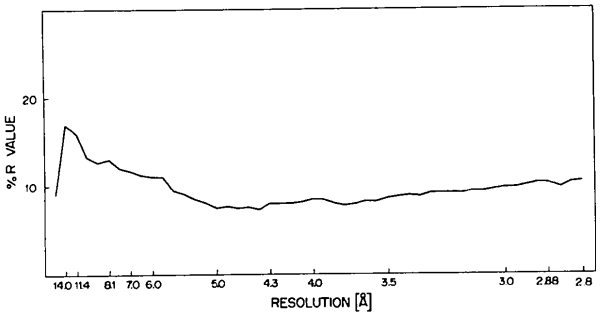

Fig. 5. Comparison of structure factors calculated using an *R*3-specific FFT program *versus* a standard algorithm based on the summation of trigonometric functions. The former was obtained by building an electron density map from atoms on which had been applied a dummy temperature factor of 40 Å^2 in order to produce a low-resolution effect. The contribution of each atom was terminated beyond a radius of 2·83 Å. The grid on which the map was computed had an interval of 1·33 Å, which was only slightly better than one-half of the highest resolution terms (2·8 Å) of the observed data.

Points of tension in a model (regions which deviated from idealized lengths and angles) to be examined on the graphics system were magnified in an additional refinement cycle for which the structure factor terms were overweighted. This tends to produce distortions in the geometry of the model due to inconsistencies between the conformational restraints and the electron density map. Several criteria were then used to determine which parts of the model needed re-inspection:

(*a*) short van der Waals contacts (less than 3·0 Å);

(*b*) unusual main-chain torsion angles (a positive value for the Ramachandran angle φ except for glycines);

(*c*) short H bonds (less than 2·5 Å);

(*d*) large differences between the independent *A*, *B* and *C* subunits (greater than 2 Å);

(*e*) a poor 'figure of merit' ($f > 0{\cdot}3$).

The main-chain figure of merit was derived by defining the three vectors:

$$\mathbf{V}_1 = \mathbf{r}_{C_{\alpha_i}} - \mathbf{r}_{C_i},$$

$$\mathbf{V}_2 = \mathbf{r}_{C_i} - \mathbf{r}_{N_{i+1}},$$

$$\mathbf{V}_3 = \mathbf{r}_{N_{i+1}} - \mathbf{r}_{C_{\alpha_{i+1}}},$$

where i is the residue number. The matrix

$$M_{ij} = \mathbf{V}_i \cdot \mathbf{V}_j$$

contains all information about the geometry of the peptide bond. The difference between the matrix for an ideal model M^I and an actual model M^M can be used to judge the overall quality of the latter. Therefore, the Euclidean norm of the difference of the matrices was used as figure of merit, f,

$$f = \| M^I - M^M \| = \sum_{i,j} (M^I_{i,j} - M^M_{i,j})^2.$$

A few examples of the sensitivity of this indicator are shown in Table 1. In general, acceptable geometry occurs when $f < 0{\cdot}3$. The overall quality of the chain geometry can be judged by the averaged value of f.

Maxima and minima of $F_o - F_c$ maps were located with a peak-search routine. Their positions with respect to the model were determined to discriminate between errors in the model, missing atoms, and in particular to locate water molecules. No peaks were interpreted as water molecules when located within the RNA region as these might be partially ordered phosphate groups.

(*C*) *Introduction of non-crystallographic constraints*

The non-crystallographic symmetry of the viral capsid was introduced as a constraint to reduce the

Table 1. *Figure of merit for main-chain geometry*

Example number	Distances C_α-C_i (Å)	C_i-N_{i+1} (Å)	N_{i+1}-$C_{\alpha_{i+1}}$ (Å)	C_{α_i}-$C_{\alpha_{i+1}}$ (Å)	Angles C_{α_i}-C_i-N_{i+1} (°)	C_i-N_{i+1}-$C_{\alpha_{i+1}}$ (°)	Torsion angle ω (°)	Figure of merit
1	1·51	1·33	1·46	3·79	116·0	122·0	180·0	0·00
2	1·54	1·32	1·48	3·82	114·1	123·9	175·0	0·165
3	1·49	1·33	1·46	3·74	111·2	124·0	179·9	0·262
4	1·50	1·23	1·47	3·71	112·6	125·9	175·3	0·356
5	1·54	1·32	1·40	3·71	108·9	128·1	173·7	0·450
6	1·46	1·27	1·40	3·58	112·4	118·6	171·2	0·506
7	1·45	1·36	1·45	3·61	100·9	130·3	176·4	0·960

The first example is an idealized peptide bond. Further examples show a decreasing quality of agreement with an idealized peptide bond.
Subscripts i and $i+1$ refer to successive amino acids along the polypeptide.
The torsion angle ω is about the C_iN_{i+1} bond.

number of independent parameters. The 60 icosahedral symmetry operators, $\hat{A}_\alpha$, were used explicitly in the computation of structure factors and their derivatives. For instance,

$$\frac{\partial F_c(\mathbf{h})}{\partial \mathbf{r}_j} = 2\pi i f_j(s)\, e^{-B_j s^2/4} \sum_{\alpha=1}^{60} (\mathbf{h}\hat{A}_\alpha)\, e^{2\pi i \mathbf{h}\hat{A}_\alpha \mathbf{r}_j} \quad (4)$$

and

$$\frac{\partial F_c(\mathbf{h})}{\partial B_j} = -\frac{s^2}{4} f(s)\, e^{-B_j s^2/4} \sum_{\alpha=1}^{60} e^{2\pi i \mathbf{h}\hat{A}_\alpha \mathbf{r}_j} \quad (5)$$

(Rossmann, 1976). These expressions account for all 60 equivalent atoms in the capsid without distinguishing between crystallographic and non-crystallographic symmetry. Hence, the initial number of independent atoms to be refined reduces to 4667, that is those belonging to the subunits *A*, *B* and *C* in one non-crystallographic asymmetric unit.

(*D*) *Selection of partial data sets*

Despite the reduction in the number of variables, imposed by the non-crystallographic constraints, the problem remained a formidable one if all 298 615 crystallographically independent observed reflections were to be used in the refinement. However, the reduction of size of the asymmetric unit in real space should be equivalent to a corresponding reduction in reciprocal space. Hence, one-tenth of the independent data might suffice for refinement.

Reduction of the number of observations due to the imposition of crystallographic symmetry is, however, not entirely equivalent to a reduction due to non-crystallographic symmetry. While the first establishes very simple relations between pairs of structure factors, the latter leads to a linear system where all structure factors are involved (Main & Rossmann, 1966). Only if $|F_o|\, e^{i\alpha_c}$ satisfy that system would a reduction of observations be valid. Hence convergence, while using a subset of data, will be slower the greater the departure of non-crystallographic symmetry in the $F_o\, e^{i\alpha_c}$ map. While refinement against one-tenth of the data might cause slow or even divergent behavior, yet refinement against less than the full data set decreases the time of computations, while retaining some part of the non-crystallographic restraint among structure factors. Similarly, retention of more than the data in one non-crystallographic asymmetric unit of reciprocal space will improve the quality of the data set due to non-crystallographic redundancy. In practice, data sets between one-seventh and one-half of the observed data set were used as a compromise between accuracy and c.p.u. time.

All reflections in the data set were sorted into increasing resolution order to select a subset of reflections. Subsets were generated by picking up reflections at equal intervals (*i.e.* every *n*th reflection) from this list. The behavior of the refinement was checked at different stages by using various subsets and then comparing shifts and trends of *R* factors* *vs* resolution.

Even by reducing the number of variables and observations as described above, one cycle of refinement still required about 12 h of Cyber 205 c.p.u. time when using only one-seventh of the data (about 45 000 reflections) and an unvectorized program. Further decrease of computing time was achieved by vectorizing the program code. The only important change involved was to re-write the routine for computation of structure factors and derivatives using explicit vector instructions. This reduced the c.p.u. time requirements by a further factor of about 15 and, hence, made feasible the refinement process.

III. Progress in refinement

(*A*) *The initial model*

The starting point of the refinement (cycle 1) was the model built in the graphics system with respect to the map produced by three cycles of molecular-replacement averaging of the original multiple-isomorphous-replacement map (Rossmann *et al.*, 1983). Structure factors were calculated with the coordinates of this model and used in the calculation of

* An overall scale factor, calculated by least squares, was used to scale $|F_o|$ and $|F_c|$ in all *R* factors reported in this paper. The *R* factor is defined as $R = \sum |(|F_o| - |F_c|)| / \sum |F_o|$.

a $(2F_o - F_c)$ map on a grid of 0·56 Å intervals. This, therefore, was the first electron density map based on the assumption of individual atoms, and thus showed remarkable improvement. Extensive modifications could now be made on the original model. The most frequent changes involved the reorientation of peptide bonds. There were a few cases where there was density for the side chain but the model was outside the density. There was no linking density between residues 177 and 178 in the internal helix αC of subunit *A*. There were 38 amino acids, out of the total of 661, where the quality of the map gave concern that the correct interpretation had been achieved. In all three subunits, the greatest difficulties occurred in the poorly defined αC helix and at the carboxy-terminal end where the main chain interacts with a Ca^{2+} ion site. The two upper curves in Fig. 6 show the *R* values as a function of resolution for the initial model and after rebuilding in the averaged $(2F_o - F_c)$ map.

(*B*) *Least-squares refinement*

The restraints applied to the model included interatomic distances and angles, planarity, chiral volume, non-bonded contacts and torsion angles. The general form of the contribution to the residual for a given type of restraint is

$$\sum_i [w(d_T - d_M)/\sigma_T]^2,$$

where the sum is taken over all restraint parameters of a given type. The weight for each term in the sum has the form w/σ_T, where σ_T is the target standard deviation for a particular restraint and w is the weight given to that term in the residual; d_T is the target value for the restraint and d_M is its actual value derived from the model. The values for w and σ_T used, as well as the observed *R* values, are given in Table 2 for successive stages of the refinement. Contacts between subunits *A*, *B* and *C* were restrained to avoid short non-bonded contacts. However, contacts between subunits related by icosahedral symmetry were not restrained. H bonds were treated as non-bonded contacts.

Only atomic coordinates were refined in cycles 1 to 26. Subsequently, individual restrained temperature factors (Konnert & Hendrickson, 1980) were also included (Table 3). The resolution range and the

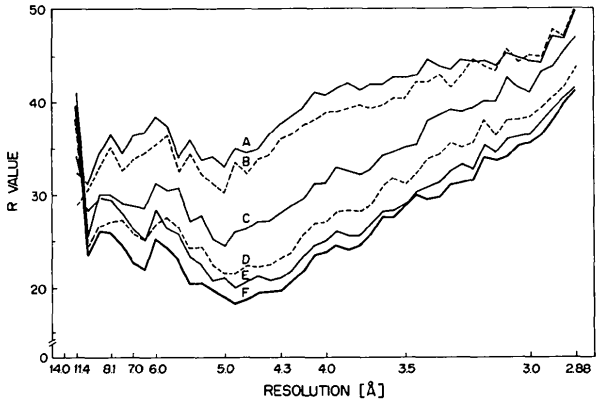

Fig. 6. *R* value *versus* resolution at different stages in the refinement. Curve *A*: model built on multiple-isomorphous-replacement plus molecular-replacement map. Curve *B*: model recorrected with respect to the first $(2F_o - F_c)$ map. Curve *C*: model at cycle 17. Curve *D*: model at cycle 36. Curve *E*; model at cycle 54. Curve *F*: final model after cycle 65.

Table 2. *Progressive adjustment of refinement restraints*

Model after cycle number	0	17	36	44	54	65
Distances						
w	1·00	1·00	1·20	1·25	1·25	1·35
#	12 962	12 962	12 962	13 060	13 060	13 042
$>4\sigma$	451	173	187	184	89	84
σ_1(0·030)	0·031	0·033	0·029	0·031	0·034	0·026
σ_2(0·040)	0·075	0·065	0·063	0·065	0·064	0·056
σ_3(0·050)	0·118	0·065	0·065	0·069	0·072	0·064
Planes						
w	1·00	1·00	1·00	1·00	1·00	1·30
#	782	782	782	788	788	789
$>2\sigma$	110	2	14	24	28	5
σ(0·020)	0·030	0·013	0·018	0·020	0·020	0·014
Chiral volume						
w	2·00	2·00	1·50	2·30	2·30	2·30
#	797	797	797	802	802	801
$>2\sigma$	30	0	40	2	3	5
σ(0·150)	0·203	0·069	0·150	0·083	0·078	0·085
Non-bonded contacts						
w	1·00	1·00	1·25	2·00	2·00	2·00
σ_1(0·500)	0·302	0·302	0·302	0·283	0·271	0·249
σ_2(0·500)	0·500	0·403	0·374	0·335	0·315	0·288
σ_3(0·500)	0·409	0·358	0·292	0·290	0·272	0·243
Torsion angles						
w	1·00	1·00	1·00	1·00	1·00	1·00
#	616	616	616	616	616	620
$0 \le \varphi < 180°$	61	61	26	7	7	12
σ(5)	5·7	2·7	4·0	4·4	4·5	4·0
r.m.s. shifts (Å)	—	0·625	0·255	0·266	0·207	0·317
R	38·9	34·1	28·2	Not computed	26·8	25·5
Resolution (Å)	∞-2·8	∞-2·8	14-2·9	14-2·9	14-2·9	14-2·9

The table shows actual r.m.s. values of various types of restrained distances, angles and volumes. Parenthetic values for the σ values show target values used in weights w/σ_T.

Distances: σ_1 is the r.m.s. deviation (Å) for bond lengths, between atoms i and $i+1$, from their idealized values; σ_2 is the r.m.s. deviation (Å) for bond-angle distances, between atoms i and $i+2$, from their idealized values; σ_3 is the r.m.s. deviation (Å) in planar groups, between atoms i and $i+3$, from their idealized values.

Planes: σ is the r.m.s. deviation (Å) from the mean plane.

Chiral volume: σ is the r.m.s. deviation (Å) from the idealized volume (Å^3) inscribed by the bonded atoms.

Non-bonded contacts: σ_1 is the r.m.s. deviation (Å) of single torsion contacts between atoms i and $i+3$; σ_2 is the r.m.s. deviation (Å) of multiple torsion contacts between atoms i and $j > i+3$; σ_3 is the r.m.s. deviation (Å) of possible H bonds.

Torsion angles is the r.m.s. deviation of torsion angles (°) from their idealized values. φ is the Ramachandran angle which should in general lie between −180° and 0° for non-glycine atoms.

The *R* value is for the full data set between the resolution limits shown.

# indicates the number of distances, planes, *etc.* which were restrained.

$>n\sigma$ indicates the number of restraints which remained bigger than $n\sigma_T$ (target) at the end of the refinement stage.

Table 3. *Progress of refinement*

Cycles (inclusive)	Subset of reflections: Resolution (Å)	Density	Number	Atoms: Number	Parameters refined	Terminal ⟨B⟩ (1/Å²)	R value for subset: from (%)	to (%)	Comments
1-3	14-5·00	1/7	9315	4661	*xyz*	20	32·5	27·2	Convergence checked using other subsets at cycle 3.
4-9	14-3·50	1/7	26 428	4661	*xyz*	20	33·2	27·9	Convergence checked using other subsets at cycle 8.
10-13	14-2·80	1/7	42 180	4661	*xyz*	20	32·3	30·1	Convergence checked using other subsets at cycle 13.
15-17	14-2·80	1/4	73 814	4661	*xyz*	20	33·2	31·3	Model corrected on graphic system after cycle 17 with respect to averaged $(2F_o - F_c)$ map.
18-23	14-2·80	1/3	98 420	4661	*xyz*	20	32·8	32·1	
24-26	14-2·88	1/3	94 117	4661	*xyz*	20	31·2	30·9	Erroneous high-resolution data eliminated.
27-36	14-2·88	1/3	94 117	4661	*xyzB*	22·2	30·9	26·6	Model corrected on graphic system after cycle 36 with respect to averaged $(2F_o - F_c)$ map.
37-43	14-2·88	1/3	94 117	4661	*xyzB*	24·6	28·2	25·9	Model corrected on graphic system after cycle 43 with respect to an unaveraged $(2F_o - F_c)$ map after cycle 36. Five amino acids added.
45-47	14-2·88	1/3	94 117	4698	*xyzB*	24·3	26·6	25·7	
49-50	14-2·88	1/3	94 117	4738	*xyzB*		25·7		40 water molecules added at cycle 49.
51-54	14-2·88	1/2	141 223	4738	*xyzB*	25·9	26·8	25·3	Model corrected on graphic system after cycle 54 with respect to an unaveraged $(2F_o - F_c)$ map.
55-59	14-2·88	1/3	94 615	4728	*xyzB*	24·8	27·2	24·7	43 water molecules present.
60-65	14-2·88	1/2	141 979	4728	*xyzB*	24·7	25·9	24·6	

Notes: (i) The 'density' of a subset represents the fraction of reflections of the complete data set used in the refinement; (ii) The *R*-value range gives the *R* value prior to cycle *i* and after cycle *j* where the refinement is for cycles *i* to *j*.

density of the subset for the observed amplitudes were increased as the refinement progressed. The *R* values for the subset used in the refinement and for a disjoint one of the same density are shown as a function of resolution for cycles 3 and 8 in Fig. 7. This demonstrates the non-independent character of the data. Fig. 8 shows a similar plot for the subset used in cycle 36 compared to the *R*-value distribution for the whole data set, demonstrating how closely the full data set follows the behavior of its subset.

The model was checked and corrected on the graphics system after cycles 17, 36, 44 and 54 on $(2F_o - F_c)$ maps. Details about the quality of the model after these cycles are given in Table 2. Fig. 6 shows the *R* values as a function of resolution for these cycles and Table 3 gives a detailed account of the subsets and resolution ranges for the data used in each cycle.

After cycle 26 a total of 40 water molecules were located on an unaveraged $(F_o - F_c)$ map, by searching two icosahedral asymmetric units. Not all of these were consistently present in both units, although usually those which were well defined obeyed both icosahedral and quasi-symmetry requirements. Nevertheless, many of these could not be seen in the averaged map. These water molecules were included from cycle 49 onward (considered to be oxygen atoms with full occupancies) and were monitored through the following refinement cycles. Ten of these molecules were eliminated after cycle 54 as these had shown large positional and temperature factor shifts. Furthermore, they were absent in the subsequent

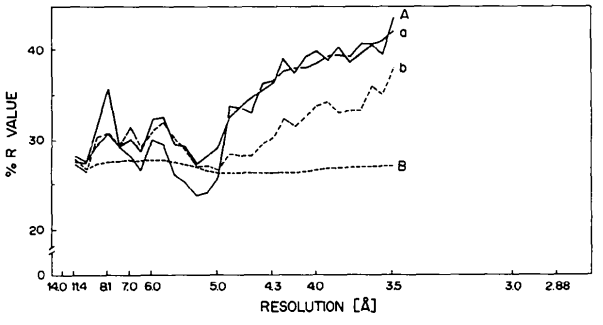

Fig. 7. Distribution of *R* factors to demonstrate that refinement with respect to a subset of the data (*A* and *B*) produced improvement of the *R* value in another disjoint subset (*a* and *b*, respectively). Curves *A* and *B* correspond to the refinement after cycles 3 and 8 with 5 Å and 3·5 Å data, respectively. Subsets *A* and *a* had 9315 reflections while subsets *B* and *b* had 26 428 reflections.

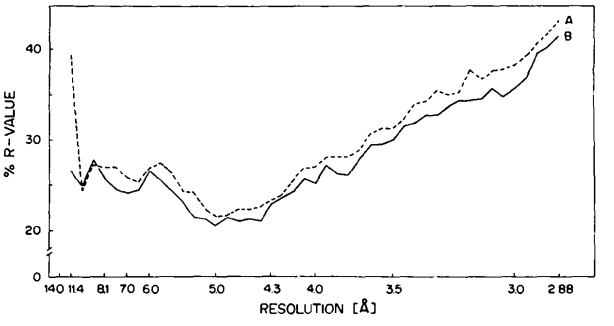

Fig. 8. *R* value *vs* resolution. Curve *A*: model at cycle 36 against a subset of 94 117 reflections used for the refinement. Curve *B*: The same model but against the full data set.

unaveraged map. However, 13 new water molecules were added at this stage.

There are two possible explanations why many of the water molecules could not be seen on the averaged map. First, they may not obey strictly the icosahedral symmetry. Secondly, the linear interpolation approximation used in averaging the maps may be introducing error comparable with the usually weak signal associated with a water molecule. Most probably both factors contribute in a negative sense.

The data in the range 2·88–2·80 Å were eliminated after cycle 24 for the reasons discussed above and because there was no improvement in the R factor in this range. The latter observation gave further confirmation that there were large errors in the observed data in this resolution range.

Some major alterations were made to the model after cycle 54. The residues 255 to 260 in the C subunit were totally rebuilt, thus clearing up difficulties in the interpretation of the density with respect to the Ca site in the CB interface. Subsequent to this operation, the C-terminal section of all three subunits and the coordinates of the Ca^{2+} were found to have closely similar conformations. The other major change was the extension of the ordered amino-terminal section in subunits A and B. Interpretable density was found for residues 62 and upward in both these subunits, albeit with mean temperature factors approaching 80 $Å^2$ instead of the normal mean of 25 $Å^2$. Hence, the last disordered residue in the amino-terminal random domain of the A and B subunits is Arg 61. This residue is involved in hydrogen bonding across the icosahedral twofold axis between the C subunits.

Refinement was concluded at cycle 65 after it was not possible to reduce the R value without distortion of the model's idealized geometry. The final overall R value was 25·55% for 283 959 reflections in the 14·0 to 2·88 Å resolution interval. The final distribution of R with resolution is shown in Fig. 6. The increasing trend of R is probably a consequence of the decreasing quality of the observed data with resolution.*

IV. The final model

The r.m.s. deviation between the initial and final model was 1·37 Å for subunit A, 1·23 Å for subunit B and 1·51 Å for subunit C. Table 4 shows the number of atoms *versus* displacement for the three subunits. The larger r.m.s. deviation for subunit C was due to the reinterpretation of the C-terminus region at cycle 54. Contacts between subunits related by icosahedral symmetry were not restrained, nevertheless they behave normally (Table 5).

Table 4. *Distribution of the number of atomic shifts between initial and final model*

Displacement (Å)		Subunits			Displacement (Å)		Subunits		
		A	*B*	*C*			*A*	*B*	*C*
0·0	0·5	274	317	287	4·0	4·5	7	0	9
0·5	1·0	612	643	722	4·5	5·0	3	0	1
1·0	1·5	311	262	346	5·0	5·5	3	0	3
1·5	2·0	141	125	131	5·5	6·0	3	0	3
2·0	2·5	67	66	65	6·0	6·5	0	0	3
2·5	3·0	37	33	52	6·5	7·0	0	0	3
3·0	3·5	28	26	21	Greater than 7·0		2	0	7
3·5	4·0	11	9	19					

Table 5. *Distribution of non-bonded contacts that were unrestrained during refinement*

Interval (Å)		Number of atoms		
		H-bonds	Ionic	Non-bonded
2·0	2·4			
2·4	2·6	4		
2·6	2·8	4	2	2
2·8	3·0	18		7
3·0	3·2	16	1	32
3·2	3·4	7	1	47
3·4	3·6			40

The greater the redundancy of information in the observed data due to the presence of non-crystallographic symmetry, the greater will be the accuracy of the final model although the R factor might even increase as the number of equivalences increases. Indeed, at present there is no truly satisfactory method of estimating error. The conjugate gradient method, used to solve systems of normal equations, does not provide the information necessary to determine error on the parameters refined by the least-squares procedure. We attempt here, however, to give some upper bounds of the atomic parameter errors.

The r.m.s. deviations between the main-chain atoms of pairwise superimposed subunits were 0·60, 0·88 and 0·74 Å. A similar comparison prior to the least-squares refinement gave 0·73, 1·17 and 1·11 Å for the A–B, B–C and C–A comparisons, respectively, suggesting convergence to progressively greater similarity. Whether or not the subunits show any significant structural difference will be examined in a subsequent paper. However, these values give an upper bound of $0·60/\sqrt{2} = 0·42$ Å for the r.m.s. error of the atomic positions. [The larger differences of 0·88 and 0·74 Å reflect real structural differences (Rossmann *et al.*, 1983).] Luzzati statistics (Luzzati, 1952) suggest an error of about 0·3 Å, but such an estimate is based on assuming no errors in the $|F_o|$ values.

The quality of the geometry of the final model is described in Table 2. The r.m.s. difference in bond length between the final model and 'ideal' distance

* The final coordinates of the A, B and C subunits relative to a P, Q, R orthogonal system (Rossmann *et al.*, 1983), as well as the associated water and Ca atoms, in addition to the F_o and final F_c, α_c values have all been deposited with the Brookhaven Protein Data Bank, Brookhaven National Laboratory (Reference: 3SBV, R3SBVSF), and are available in machine-readable form from the Protein Data Bank at Brookhaven or one of the affiliated centers at Cambridge, Melbourne or Osaka. The data have also been deposited with the British Library Lending Division as Supplementary Publication No. SUP 37013 (10 microfiche). Free copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

is 0·026 Å. Mean temperature factors per residue are shown in Fig. 9 as a function of the residue number. The mean temperature factor is 26 Å^2 for subunit *A*, 24 Å^2 for subunit *B* and 22 Å^2 for subunit *C*.

The correlation between temperature factors in different subunits is clear (Fig. 9) and demonstrates their physical significance. Large temperature factors are in general associated with elements connecting the β-strands of the subunit. For instance, the internal helix αC and its connection with βG (an insertion in SBMV relative to TBSV) and the external loop connecting βH and βI have unusually large temperature factors in all three subunits. The αC helix was particularly 'fuzzy' in the original multiple-isomorphous and molecular-replacement map. The helix αC is surrounded by RNA while the βH-βI loop is the most external region of the viral shell involved in the interparticle contacts in subunit *A*,

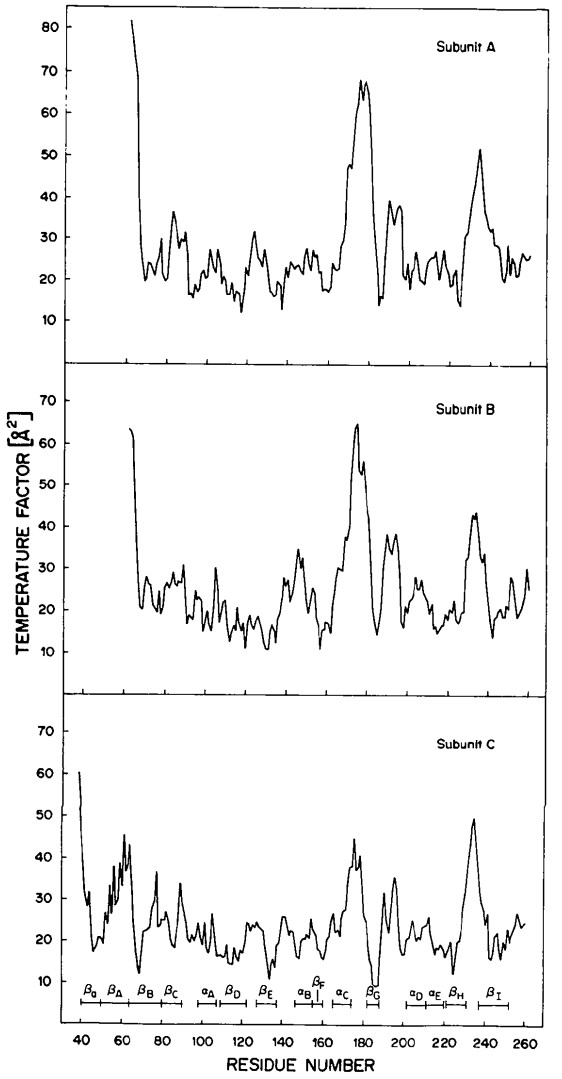

Fig. 9. Mean temperature factor for each residue as a function of the residue number of subunits *A*, *B* and *C*.

Table 6. *Mean temperature factors for basic residues*

Subunit	*A*		*B*		*C*		
Residue	Main chain (Å^2)	Side chain (Å^2)	Main chain (Å^2)	Side chain (Å^2)	Main chain (Å^2)	Side chain (Å^2)	Location
Lys 54	—	—	—	—	27·0	36·4	Internal surface
Arg 56	—	—	—	—	29·1	41·6	Internal surface
Arg 61	—	—	—	—	36·9	49·7	Internal surface
Arg 101	22·4	27·4	14·5	21·0	18·6	25·9	External surface
Lys 109	17·8	19·0	19·9	19·4	15·2	15·8	Subunit interface
Arg 116	15·6	15·3	16·5	21·7	16·5	17·4	Internal surface
Lys 154	22·9	26·0	20·6	24·6	22·8	26·6	Subunit interface
Lys 175	64·4	69·5	62·3	67·4	36·9	45·2	Internal surface
Arg 181	47·7	52·0	35·2	44·4	22·4	26·2	Fivefold or quasi-sixfold contacts
Lys 195	33·3	41·4	31·8	36·0	33·1	36·0	Internal surface
Arg 196	27·4	34·1	23·1	33·3	29·5	32·0	Internal surface
Lys 200	19·2	16·3	18·7	17·4	16·0	15·8	Subunit interface
Arg 223	18·6	23·7	17·8	21·8	15·6	20·6	External surface
Lys 234	47·5	52·5	41·0	46·2	45·7	50·5	External surface
Arg 241	25·1	27·6	16·9	16·4	20·7	27·9	Subunit interface
Arg 249	19·8	23·9	18·4	23·3	15·6	19·3	Subunit interface

possibly producing deviations from non-crystallographic symmetry. Other regions of high temperature factors are loops which make intersubunit contacts along fivefold and quasi-sixfold axes, and also the basic region between βG and αD in contact with RNA. Table 6 shows temperature factors for all basic residues. These again show correlation between subunits. They also show that basic residues in contact with RNA are in general no more disordered than residues on the external surface. Those residues associated with the αC region have greater positional displacement in the *A* and *B* subunits compared to those in the *C* subunit.

The final model includes 34 water molecules associated with the three subunits. Their temperature factors vary between 17 and 88 Å^2. Most of the water molecules are on the external viral surface while a few are in subunit interfaces. Particle contacts, along the fivefold axes, are mediated by water. Several electron density peaks can be seen in these contact regions as well as in the cavity formed between two interdigitating particles. However, some waters placed in these densities neither refine well nor make acceptable H-bonding associations. Some electron density, which might also be interpreted as partially ordered water, is observed inside the channel formed by the fivefold-related subunits. This density penetrates as far as Arg 181. At the resolution of the current data, there is no evidence for an extended water layer on the internal surface of the viral capsid as proposed for STNV (Jones & Liljas, 1984*b*).

A metal ion on the quasi-threefold axis, as had been proposed by Abdel-Meguid *et al.* (1981) and Rossmann *et al.* (1983), gave rise to very short contacts with Glu 194 of all three subunits, distorted the geometry in this region and displaced several atoms outside the electron density. Nor was there a well defined electron density peak at this site. Consequently, the Ca atom at this site was omitted after cycle 54.

The only well defined metal-ion sites are those mediating subunit contacts between Asp 138 and Asp 141 of one subunit and O 199, O 260, O 259 and $O_{\delta 1}$ Asn 259 of the other. This means that two charges are provided by one subunit and a third by the terminal carboxyl of the other. Temperature factors for the Ca^{2+} atoms are 27·2, 16·6 and 20·9 $Å^2$ at the *A–B*, *B–C* and *C–A* interface, respectively.

The electron density inside the viral capsid is in general very low and discontinuous. A continuous region of low density can be observed at the internal surface near the quasi-threefold axis. It is close to a series of basic residues (Lys 195, Arg 196, Arg 249 of all three subunits). This density could be interpreted as an RNA strand of at least six nucleotides, but these were not included in the present model.

This is the first time that the structure of a complete virus has been refined by a reciprocal-space method. The explicit use of non-crystallographic symmetry in reciprocal and direct space made possible this process, which considerably simplifies the refinement of large macromolecular assemblies. Nevertheless, the procedure used here taxed the limits of the computing power of the Cyber 205. Further improvements as indicated above would be to use FFT techniques throughout for the calculations of structure factors and their derivatives. However, the most demanding part of the process is the recurring visual inspection of the model in the graphics system. This is caused by the intrinsic weakness in the least-squares method which cannot surmount local minima. A possible solution would be the systematic exploration of different local minima by rotating side chains and peptide bonds while checking contacts and normal stereochemistry.

We thank Ignacio Fita for help in the implementation of the Konnert–Hendrickson least-squares program, and Wayne Hendrickson for discussions relating to the identification of water molecules in the solvent. We have enjoyed discussions with Lars Liljas and Alwyn Jones as to the best refinement procedures throughout the progress of this work. We are grateful to Sharon Wilder for assistance in the preparation of this manuscript. The work was supported by the National Science Foundation and the National Institutes of Health.

References

ABAD-ZAPATERO, C., ABDEL-MEGUID, S. S., JOHNSON, J. E., LESLIE, A. G. W., RAYMENT, I., ROSSMANN, M. G., SUCK, D. & TSUKIHARA, T. (1980). *Nature (London)*, **286**, 33-39.

ABAD-ZAPATERO, C., ABDEL-MEGUID, S. S., JOHNSON, J. E., LESLIE, A. G. W., RAYMENT, I., ROSSMANN, M. G., SUCK, D. & TSUKIHARA, T. (1981). *Acta Cryst.* B**37**, 2002-2018.

ABDEL-MEGUID, S. S., YAMANE, T., FUKUYAMA, K. & ROSSMANN, M. G. (1981). *Virology*, **114**, 81-85.

AGARWAL, R. C. (1978). *Acta Cryst.* A**34**, 791-809.

BARRY, C. D., BOSSHARD, H. E., ELLIS, R. A. & MARSHALL, G. R. (1975). *Computers in Life Science Research*, edited by W. SILER & D. A. B. LINDBERG, pp. 137-147. New York: Plenum.

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

CASPAR, D. L. D. & KLUG, A. (1962). *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1-24.

DIAMOND, R. (1971). *Acta Cryst.* A**27**, 436-452.

HENDRICKSON, W. A. & KONNERT, J. H. (1980). *Computing in Crystallography*, edited by R. DIAMOND, S. RAMASESHAN & K. VENKATESAN, pp. 13·01-13·25. Indian Academy of Science: Bangalore.

HERMODSON, M. A., ABAD-ZAPATERO, C., ABDEL-MEGUID, S. S., PUNDAK, S., ROSSMANN, M. G. & TREMAINE, J. H. (1982). *Virology*, **119**, 133-149.

HSU, C. H., SEHGAL, O. P. & PICKETT, E. E. (1976). *Virology*, **69**, 587-595.

HULL, R. (1978). *Virology*, **89**, 418-422.

JACK, A. & LEVITT, M. (1978). *Acta Cryst.* A**34**, 931-935.

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

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

JONES, T. A. & LILJAS, L. (1984*a*). *Acta Cryst.* A**40**, 50-57.

JONES, T. A. & LILJAS, L. (1984*b*). *J. Mol. Biol.* **177**, 735-768.

KONNERT, J. H. & HENDRICKSON, W. A. (1980). *Acta Cryst.* A**36**, 344-350.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802-810.

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

ROSSMANN, M. G. (1976). *Acta Cryst.* A**32**, 774-777.

ROSSMANN, M. G., ABAD-ZAPATERO, C., HERMODSON, M. A. & ERICKSON, J. W. (1983). *J. Mol. Biol.* **166**, 37-83.

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

SUSSMAN, J. L., HOLBROOK, S. R., CHURCH, G. M. & KIM, S. H. (1977). *Acta Cryst.* A**33**, 800-804.

TEN EYCK, L. F. (1973). *Acta Cryst.* A**29**, 183-191.

TEN EYCK, L. F. (1977). *Acta Cryst.* A**33**, 486-492.

International Union of Crystallography

Acta Cryst. (1985). B**41**, 157

Acta Crystallographica Indexes

The indexes to Volume 39 (1983) of *Acta Crystallographica* have just been distributed to subscribers. The International Union of Crystallography regrets the delay in publishing these indexes, which is due to the introduction of a computerized index-production system. The system will be used to produce the next five-year index to Volumes 39-43. The indexes to Volume 40 (1984) are expected to be distributed on time.

A ten-year compilation of the indexes for Volumes 29-38 (1973-1982) was distributed to subscribers in mid-1984. Further copies are available at a price of Dkr 150 (Dkr 75 for scientists who give an undertaking that the index is for their own personal use).