References

ALEXANDROU, N. E. & ARGYROPOULOS, N. (1977). Sixth International Congress of Heterocyclic Chemistry, Teheran.

CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104–109.

DOMENICANO, A., VACIAGO, A. & COULSON, C. A. (1975). *Acta Cryst.* B**31**, 1630–1641.

International Tables for X-ray Crystallography (1968). Vol. III, 2nd edition, pp. 215–216. Birmingham: Kynoch Press.

MAIN, P. (1970). Private communication.

MAIN, P., WOOLFSON, M., LESSINGER, L., GERMAIN, G. & DECLERCQ, J.-P. (1974). *MULTAN, A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.* Univs. of York, England, and Louvain-la-Neuve, Belgium.

STEWART, J. M., KRUGER, G. J., AMMON, H. L., DICKINSON, C. & HALL, S. R. (1976). The XRAY 76 system. Tech. Rep. 446. Computer Science Center, Univ. of Maryland, College Park, Maryland.

STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.

Acta Cryst. (1978). B**34**, 567–578

An 11 Å-Resolution Electron Density Map of Southern Bean Mosaic Virus

BY IVAN RAYMENT, JOHN E. JOHNSON, DIETRICH SUCK,* TOSHIO AKIMOTO† AND MICHAEL G. ROSSMANN

Departments of Biological Sciences and Biochemistry, Purdue University, West Lafayette, Indiana 47907, *USA*

AND IN PART BY KARL LONBERG-HOLM AND BRUCE D. KORANT

Central Research and Development Department, E. I. Dupont de Nemours and Co. Inc., 1007 *Market Street, Wilmington, Delaware* 19898, *USA*

AND JOHN E. JOHNSON

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

(*Received* 5 *May* 1977; *accepted* 26 *August* 1977)

Screened precession photography was used to collect 11 Å-resolution data for native and K_2HgI_4 derivative crystals of southern bean mosaic virus. Rotation functions showed that not only the virus but also the sites of HgI_4^{2-} attachment had icosahedral symmetry. A difference Patterson synthesis was systematically searched for icosahedral distributions of vectors between heavy atoms. The resultant search function was examined for sets of solutions consistent with $T = 3$ symmetry. The major set of consistent peaks was then refined by a least-squares procedure which assumed the non-crystallographic symmetry. The resultant single isomorphous replacement phases were used to compute an electron density distribution which was icosahedrally averaged. The averaged map was the basis for an improved set of phases. Two further cycles of molecular replacement produced an electron density map with features similar to an earlier 22·5 Å-resolution structure which had been determined without the benefit of any heavy-atom derivatives. The new 11 Å-resolution map showed additional details consistent with the $T = 3$ symmetry.

Introduction

Particles of southern bean mosaic virus (SBMV) are roughly spherical with a molecular weight of $6{\cdot}6 \times 10^6$ (Miller & Price, 1946; Yphantis, 1964) (Appendix I). Their mean diameter is 284 Å, although the distance between opposing fivefold vertices is 318 Å (Johnson, Akimoto, Suck, Rayment & Rossmann, 1976). There are 180 identical protein subunits, each of molecular weight 28 250 (Tremaine, 1966; Hill & Shepherd, 1971) (Appendix I) consistent with $T = 3$ quasi-symmetry (Caspar & Klug, 1962). Four different crystal forms have been reported (Akimoto, Wagner, Johnson & Rossmann, 1975), of which the type II crystals are the most suitable for detailed structural investigations.

* Present address: European Molecular Biology Laboratory, Postfach 10 22 09, 69 Heidelberg 1, Federal Republic of Germany.

† Present address: Faculty of Pharmaceutical Sciences, University of Tokyo, Bunkyo-ku, Tokyo, Japan.

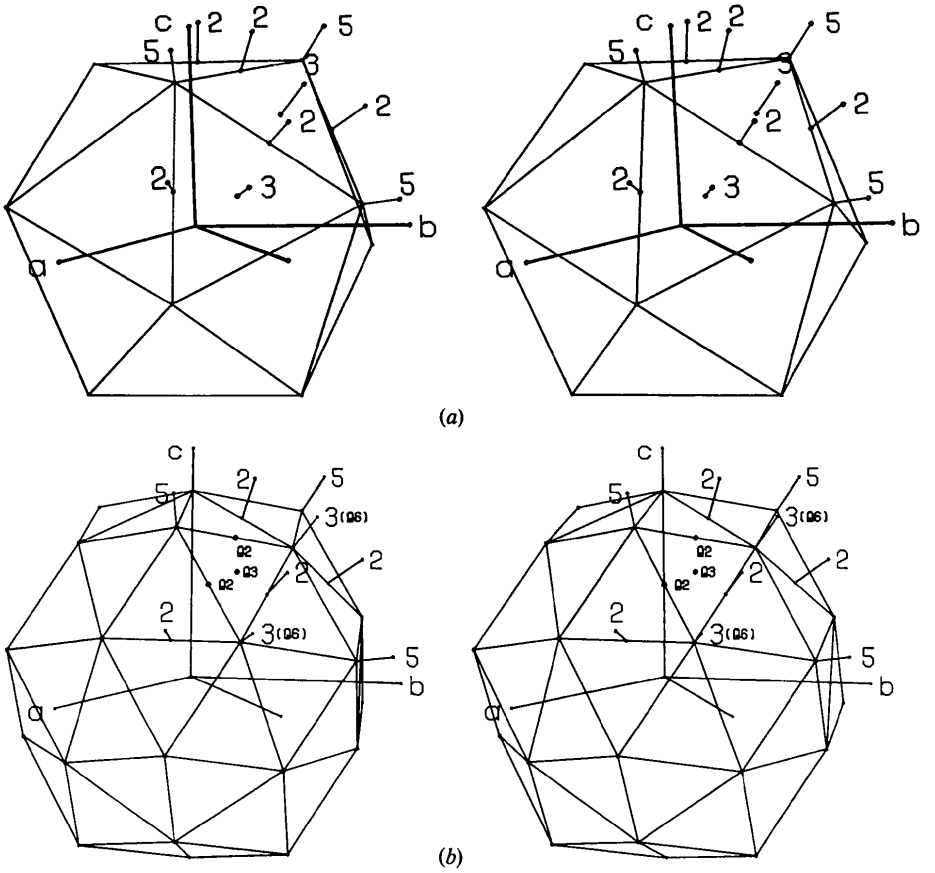


Fig. 1. The icosahedral virus in the $R32$ crystal cell; (*a*) an icosahedron showing crystallographic and icosahedral symmetry elements and (*b*) the same icosahedron triangulated to demonstrate the quasi $T = 3$ symmetry. Note that the quasi-twofold axes relate the adjacent icosahedral fivefold and threefold (quasi-sixfold) axes.

These belong to space group $R32$ where $a = 318$ Å and $\alpha = 64° 0'$ (or $a = 337$, $c = 756$ Å in the hexagonal setting) with one virus particle in the rhombohedral unit cell located at the intersection of one threefold and three twofold axes (Fig. 1*a*). Thus, the crystallographic asymmetric unit contains one-sixth of the virus particle corresponding to ten identical icosahedral units.

The packing of particles in the type II crystals was compared with that in other known crystal forms of SBMV by Akimoto *et al.* (1975). They also distinguished between the two possible orientations of the icosahedral symmetry axes in the cell. Subsequently, Johnson *et al.* (1976) reported a 22·5 Å-resolution electron density map obtained by *ab initio* molecular-replacement phase determination. The initial set of 40 Å-resolution signs was derived from the transform of a sphere with a radius equal to the average radius of the particle. The sign determination was then extended to 22·5 Å resolution by repeated averaging of the electron density and computation of structure factors to progressively higher resolution. The resultant electron density map suggested that the protein subunit had a bilobal structure with one lobe associated with hexamer or pentamer clusters and the other lobe close to the quasi-threefold axes.

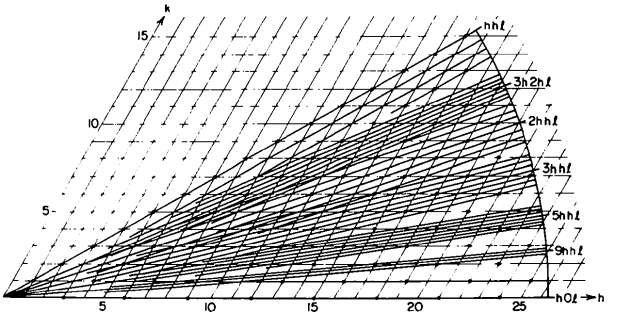


Fig. 2. Scheme used for collecting 11 Å-resolution film data for the native and K_2HgI_4 derivative SBMV crystals. Pairs of non-overlapping adjacent layers were taken on a single film pack for the closely spaced upper layers in all cases except for the (*h0l*) and (*hhl*) reciprocal-lattice planes. Such double-layers films were processed twice, once for each independent layer.

This paper presents the determination of the heavy-atom positions for a K_2HgI_4 derivative. Their refined parameters were used to compute a single isomorphous replacement map at 11 Å resolution. The phases were improved by two cycles of a molecular-replacement averaging procedure.

In discussing the structural results, it is necessary to differentiate between three different types of symmetry elements (Fig. 1*b*): crystallographic, icosahedral and local quasi-symmetry. The virus particle has exact icosahedral 532 symmetry with one of its threefold and three of its twofold axes coinciding with crystallographic symmetry axes (Fig. 1*a*). In a $T = 3$ surface lattice (Caspar & Klug, 1962), there are additional quasi- or local symmetry axes giving rise to approximate symmetry relations between neighboring protein subunits.

Data collection

SBMV of the cowpea strain (Ghabrial, Shepherd & Grogan, 1967) was isolated and crystallized in the type II form with the methods previously described by Akimoto *et al.* (1975). The original SBMV sample was kindly supplied by Dr R. J. Shepherd. The heavy-atom derivative was prepared by crystallizing SBMV in the presence of 0·116 m*M* K_2HgI_4. Although these crystals were generally smaller and were less readily obtained than the native crystals, they showed no discernible changes in cell dimension. When the concentration of K_2HgI_4 was halved, the diffraction pattern showed only very small differences with respect to the native data. Differences in intensities between native and K_2HgI_4 derivative crystals reach a plateau at about 0·07 m*M* K_2HgI_4, indicating that a limited number of specific sites had been substituted. Derivative crystals prepared by soaking K_2HgI_4 into native crystals were normally cracked and showed poor diffraction patterns.

Conventional screened precession photographs were used to collect 11 Å-resolution data for both native and heavy-atom derivative crystals according to the scheme shown in Fig. 2. The width of the annulus on the layer-line screen was 1 mm. 34 and 24 reciprocal-lattice planes were photographed for the native and derivative data sets, respectively. A new crystal was used for each film with 40 h exposure of Cu $K\alpha$-radiation from an Elliott rotating-anode X-ray generator. A 0·2 mm focal cup and a 0·2 mm diameter collimator were used for the native-data collection. A 0·1 mm focal cup and two perpendicular focusing mirrors (Harrison, 1968) were used in the K_2HgI_4 data collection. The crystal-to-film distance was 10 cm for all films. The precession angle was set at $\mu = 4°0'$ for zero-layer lines, but was reduced for upper levels so as to maintain the same exposure time without loss of resolution. Intensity measurements were made with an Optronics 'photoscan' film scanner. The film-scanning procedure (Matthews, Klopfenstein & Colman, 1972; Ford, 1974) determined the intensity of a reflection by a profile-fitting method. Such a technique is less affected by random error in optical density measurements than a summation integration process, especially in the case of the numerous weak reflections. Different films within one film pack, symmetry-related reflections within one film plane, and different film planes were successively scaled together for each compound by the method of Hamilton, Rollett & Sparks (1965). The results of scaling the heavy-atom data to the native data with a least-squares procedure (Adams *et al.*, 1969) are summarized in Table 1.

Table 1. *Data-processing results*

Data	Number of film packs	r^a (%)	Number of significant independent reflections	Fitting statistics[b] $\lvert\bar{\Delta}\rvert$	$(\bar{\Delta}^2)^{1/2}$	$\bar{F}$	R(%)
Native	32	10·7	5722	275	349	973	28·2
K_2HgI_4	24	13·5	4939				

Notes: (*a*) $r = \sum_h \frac{\sum_l (I_{hl} - I_h)}{n} \Big/ \sum_h I_h$

where I_h is the mean of the n reflections I_{hl} on films 1, 2, . . . , n. (*b*) $\lvert\bar{\Delta}\rvert$ and $(\bar{\Delta}^2)^{1/2}$ are the mean and r.m.s. differences with respect to the native data. $\bar{F}$ is the mean native structure amplitude for the terms used to compute $\lvert\bar{\Delta}\rvert$ and $(\bar{\Delta}^2)^{1/2}$. $R = \lvert\bar{\Delta}\rvert/\bar{F}$.

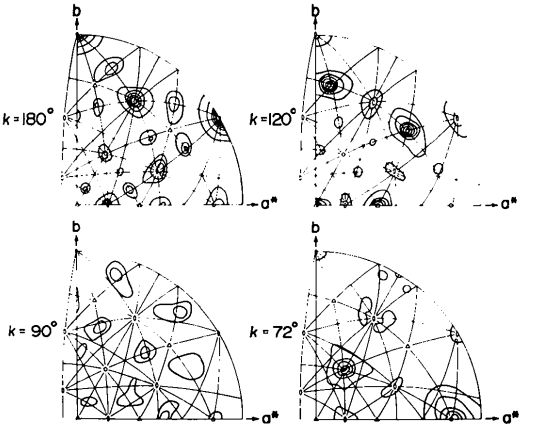


Fig. 3. Stereographic projection of the rotation function, computed with native SBMV data limited between 13 and 11 Å resolution, showing sections $\kappa = 180$, 120, 90 and 72°. The radius of integration was 50 Å. There were 1975 independent reflections of which the 118 largest terms were selected to represent one of the Patterson syntheses. The assumed icosahedral symmetry elements are superimposed.

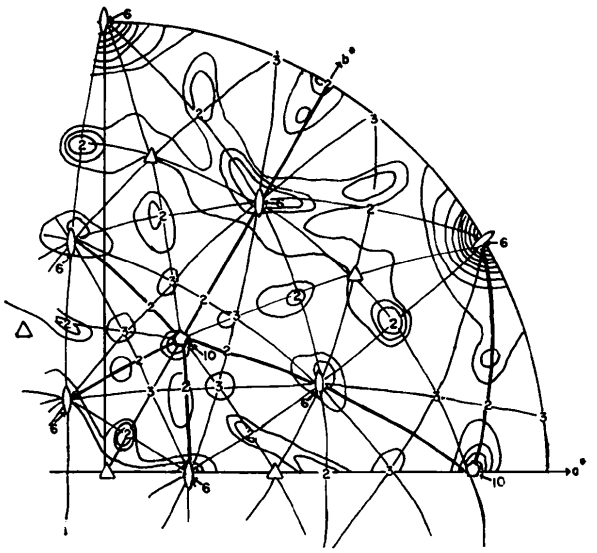


Fig. 4. Stereographic projection of the $\kappa = 180°$ rotation function section for the native SBMV data with a $T = 4$ surface lattice superimposed. Resolution of data was limited to 20–11 Å. Radius of integration was 100 Å.

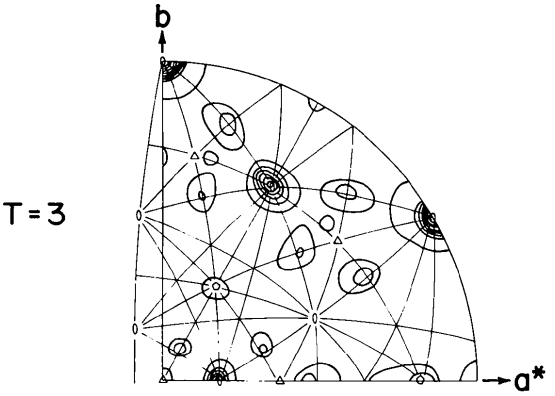


Fig. 5. Stereographic projection of the $\kappa = 180°$ rotation function section for a centric SBMV model with $T = 3$ symmetry. A $T = 4$ surface lattice is superimposed.

Rotation-function results

A rotation function (Rossmann & Blow, 1962) was computed (Fig. 3) as a crude check of the native data. Although a variety of conditions were tested (radius of interpolation, maximum and minimum resolution limits and the number of large terms representing one of the Patterson syntheses), only minor alterations in the appearance of the function were observed. An icosahedral distribution of large peaks, corresponding to the previously determined orientation of the virus (Akimoto *et al.*, 1975; Johnson *et al.*, 1976), was immediately obvious. A number of other systematic background features were, however, also present. These had the distribution of radial, quasi $T = 4$, local symmetry elements (Fig. 4). Not only would it be surprising to observe quasi-symmetry but a $T = 4$ arrangement would require 240 protein subunits, which would be inconsistent with previous determinations.

The molecular weight of the subunit had been established at around 28 500 by a number of investigators. Sehgal & Sinha (1974) found a molecular weight of 28 000 using SDS gels. Tremaine (1966) and Ghabrial *et al.* (1967) found a molecular weight of 29 000 using amino acid composition and minimal sulfhydryl content. Hill & Shepherd (1971) used SDS gels on the severe strain of SBMV and found a molecular weight of 26 700 ± 300. In addition, Lonberg-Holm & Korant (Appendix I) redetermined the molecular weight of the protein subunit from the cowpea strain of SBMV and found it to be 28 250 ± 250. The molecular weight of the complete virus had been determined as $6{\cdot}6 \times 10^6$ by Price (1946) and Yphantis (1964). A sedimentation coefficient of 112*S* (Appendix I) for the cowpea strain of SBMV was consistent with the earlier work. Thus, even if the 21% RNA content (Ghabrial *et al.*, 1967) is somewhat high, it was difficult to avoid the conclusion that there were only 180 protein subunits per virus particle.

Structure factors were, therefore, computed for a variety of models based on $T = 3$ and $T = 4$ symmetry,

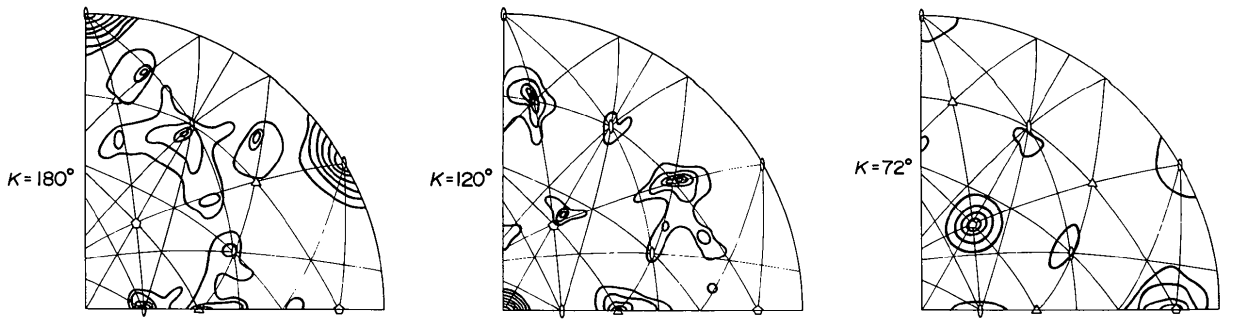


Fig. 6. Stereographic projection of the $\kappa = 180$, 120 and 72° sections for the $(F_{K_2HgI_4} - F_{native})^2$ difference rotation function, which demonstrates the icosahedral distribution of the heavy atoms.

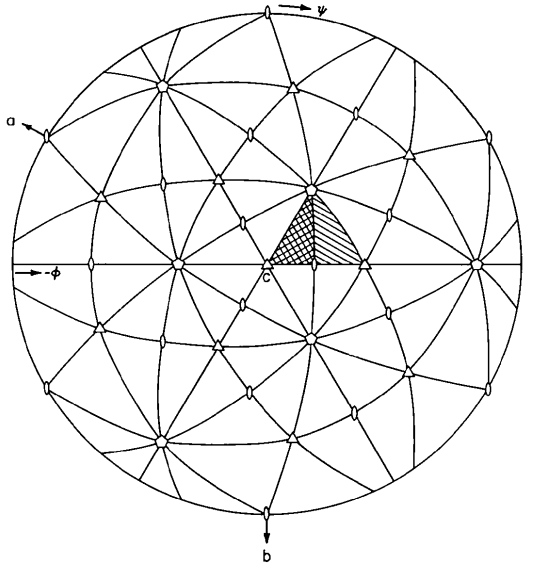


Fig. 7. Stereographic projection of the orientation of the icosahedral virus in the $R32$ cell, showing a typical icosahedral asymmetric unit (shaded). Only half of the hatched area need be searched for a possible heavy-atom site as trial positions in each half of the shaded area are related by a mirror inversion. The chosen definitions of the polar coordinates ψ and φ are also shown.

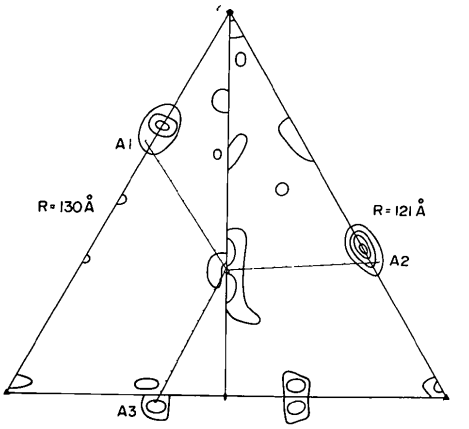


Fig. 8. Patterson search function results for radii of 130 Å (right) and 121 Å (left). The three major peaks are related by an approximate threefold axis positioned near the center of the icosahedral asymmetric unit.

and corresponding rotation functions were calculated. Models based on $T=3$ symmetry showed patterns similar to those observed for the SBMV data whenever the electron density tended to have a centric distribution (Fig. 5). It was concluded that the systematic background features in the observed SBMV rotation function were due to a pseudo-centric distribution of mass and not due to $T=4$ quasi-symmetry. The origin of the $T=4$ distribution for the smaller features on the function is not entirely clear.

A self-rotation function was also computed for the difference Patterson synthesis with $(F_{K_2HgI_4}-F_{native})^2$ coefficients to determine whether the heavy-atom positions maintained icosahedral symmetry. The conditions chosen for this function (Fig. 6) were similar to those selected for the native rotation function. The icosahedral symmetry was apparent, thus permitting a systematic search of the Patterson synthesis for an icosahedral distribution of heavy atoms.

The Patterson search

Systematic searches for heavy atoms in chemically identical environments related by non-crystallographic symmetry can be used when a straightforward interpretation of the Patterson synthesis would be difficult or impossible (Argos & Rossmann, 1974, 1976). These techniques have been used to determine the heavy-atom positions for the icosahedral $T=1$ satellite tobacco necrosis virus (Lentz *et al.*, 1976). The problem is somewhat more difficult for SBMV in view of the approximate nature of the $T=3$ quasi-symmetry. Here, any one icosahedral asymmetric unit will contain multiples of three atoms related by the quasi-threefold symmetry axis whose orientation and position is not precisely known. Only vectors between trial atomic positions related by 532 symmetry can thus be treated, neglecting all vectors between and among the other two $T=3$ related atoms. Possible solutions of the heavy-atom sites should, nevertheless, obey the quasi-symmetry.

The 11 Å-resolution difference Patterson map was computed on a grid with intervals near 3·7 Å in each direction. Trial heavy-atom positions were tested within the icosahedral asymmetric unit (Fig. 7) defined by the crystallographic threefold axis, the icosahedral twofold axis and the icosahedral fivefold axis. The test positions were separated by an angle of 1° in ψ and φ corresponding to a distance of less than 2·7 Å [see Rossmann & Blow (1962) for the definition of spherical coordinates, as well as Fig. 7]. Search sections at a constant radial distance, R, were generated between 118 Å $< R <$ 157 Å in steps of 3 Å. The fit of each vector distribution to the Patterson map was tested at 6642 positions in all. The criterion of fit was a modified sum of the Patterson densities at the end of the vectors (Argos & Rossmann, 1976).

The search function was inspected for peaks which were related by an approximate threefold axis near $\psi=79°$, $\varphi=-110{\cdot}9°$. This was readily accomplished by constructing a trigonal ruler of the type depicted in Fig. 8. While quasi-equivalent peaks might be expected to lie on the same radial section, deviation from this assumption would occur if the quasi-threefold axes

Table 2. *Set of peaks approximately related by a quasi-threefold axis*

	Peak 1				Peak 2				Peak 3			
Set	Height	ψ (°)	φ (°)	R (Å)	Height	ψ (°)	φ (°)	R (Å)	Height	ψ (°)	φ (°)	R (Å)
A	31	67·5	−117·0	129	41	78·0	−98·0	122	21	91·0	−117·5	118
B	15	89·0	−97·7	127	14	64·5	−110·5	130	15	89·0	−124·2	127
C	15	89·0	−97·7	127	14	61·5	−105·0	130	12	84·0	−128·0	130
D	11	84·0	−106·0	143	13	74·0	−111·0	147	11	84·0	−116·0	143
E	11	84·0	−116·0	142	13	74·0	−111·0	147	13	81·3	−105·0	142
F	13	85·0	−94·0	140	16	63·7	−111·5	147	12	87·0	−126·0	136

Largest peaks not among above atom sets

	Height	ψ (°)	φ (°)	R (Å)
In special position	24	73·0	−101·0	151
In general position	18	79·0	−105·0	148

Table 3. *Refined* K_2HgI_4 *heavy-atom parameters*

		Fractional coordinates			Polar coordinates			Shape parameters[b]
Site	Occupancy[a]	x	y	z	ψ	φ	R (Å)	(Å)
A_1	58	−0·1839	0·0594	0·1396	66·7	−117·0	128·9	10
A_2	99	−0·0531	0·0504	0·1554	77·7	−97·5	121·3	10
A_3	66	−0·1864	−0·1039	0·1358	91·8	−117·9	116·2	10

Notes: (*a*) The scale of the structure amplitudes was chosen to give less than 100 e for the largest site. (*b*) Kept constant throughout the refinement.

Table 4. *Refinement error analysis*

Exact definitions of the quantities given will be found in Buehner, Ford, Moras, Olsen & Rossmann (1974).

Resolution	42 Å	28 Å	21 Å	17 Å	14 Å	12 Å	11 Å	Overall
R.m.s. closure E	2251	1181	827	675	673	793	1056	
R.m.s. small f	950	929	1197	994	944	1067	1001	
R.m.s. differences	2862	1602	1544	1035	981	1181	1441	
R.m.s. F errors	440	168	206	255	257	229	301	
R modulus	197	110	67	62	68	67	86	72
R weighted	67	111	33	36	51	41	101	52
Number per annulus	62	240	482	746	1059	1406	654	4656
Mean figure of merit	0·19	0·26	0·39	0·38	0·38	0·36	0·29	0·35

were not radial. For instance, the quasi-threefold axes are parallel to the adjacent icosahedral twofold axes in tomato bushy stunt virus (Harrison & Jack, 1975; Winkler, Schutt, Harrison & Bricogne, 1977). In the latter situation the expected radial distance of quasi-threefold related points would increase slightly as ψ decreases for the icosahedral unit selected in Fig. 7. In Table 2 are shown combinations of the largest search function peaks which might be related by a quasi-threefold axis. By far the largest peaks in the search function (set *A* in Table 2) were related by a quasi-threefold axis (Fig. 8) approximately parallel to its neighboring twofold axis. The set of three largest peaks was, therefore, accepted as trial heavy-atom positions.

Refinement of heavy-atom sites

The quasi-threefold related peaks, suggested by the search function, were refined in a single isomorphous least-squares procedure adapted to include the icosahedral symmetry (Rossmann, 1976). Thus only three atoms, A_1, A_2 and A_3 (Table 2), were refined independently, although 30 atoms were positioned by icosahedral symmetry within the crystallographic asymmetric unit.

The final parameters and an analysis of the refinement are given in Tables 3 and 4, respectively. A difference Fourier map, based on the refined phases, gave no indication of any minor sites consistent with $T = 3$ symmetry.

The R factor and figures of merit are reasonable for a single isomorphous replacement refinement. They do not, however, constitute proof of the correct interpretation of the Patterson synthesis. A test was, therefore, conducted in which only two of the sites were included in the refinement to observe if they gave the position of the third. After refinement of sites A_1 and A_2 alone, phases were generated and used to compute a difference electron density map. The known sites had heights of 83 and 151 on the averaged map, respectively, whereas site A_3 had a height of 27 arbitrary electron density units. Site A_3 was the third highest peak on the map. There were peaks of 25 and 21 units on the icosahedral fivefold and threefold axes, respectively, while all other peaks were less than 20 units.

The refined positions were examined to determine the orientation and position of the quasi-threefold axes. Angular deviations from the threefold relation are less than 7° (Fig. 9) and the normal of the plane defined by the heavy atoms makes an angle of 2·5° with the neighboring icosahedral twofold axis.

The heavy-atom distribution is close to centric; thus any phase determination based on the K_2HgI_4 derivative will yield both the real structure and a weaker enantiomorphous image. The degree of centricity imposed on the structure is dependent on the displacement of the heavy atoms (about 4 Å for site A_1) from the pseudo-mirror planes in relation to the resolution of the data. However, this effect is probably not too serious in the present study because the gross distribution of density is itself roughly centric to the limits of 11 Å resolution, as demonstrated by the rotation-function results.

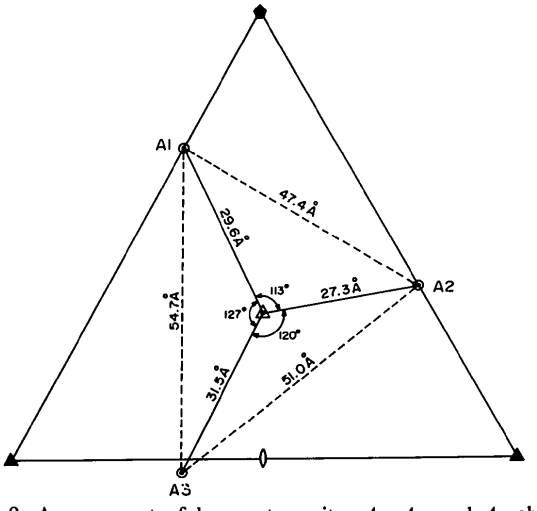


Fig. 9. Arrangement of heavy-atom sites A_1, A_2, and A_3 about the quasi-threefold axis. Angles and distances between quasi-equivalent points are shown.

Table 5. *Criteria of convergence of phase refinement during real-space molecular replacement*

	Mean phase change	Standard deviation of electron-density points related by icosahedral symmetry[a]	Figure of merit	R (%)
SIR	52°	96	0·35 (SIR)	
1st cycle		48	0·68 (Sim weighted)	61
2nd cycle	12°	44	0·75 (Sim weighted)	50

Note: (*a*) $\sigma = (\Sigma_{j=1}^{n} \sigma_j / \Sigma_{j=1}^{n} \rho_{av}^2)/n$ where n is the total number of unique grid points in the molecular envelope; $\sigma_j = [\Sigma_{j=1}^{m} (\rho_j - \rho_{av})^2]/(m-1)$ where m is the number of icosahedrally related points in the crystallographic asymmetric unit; and $\rho_{av} = (\Sigma_{j=1}^{m} \rho_j)/m$.

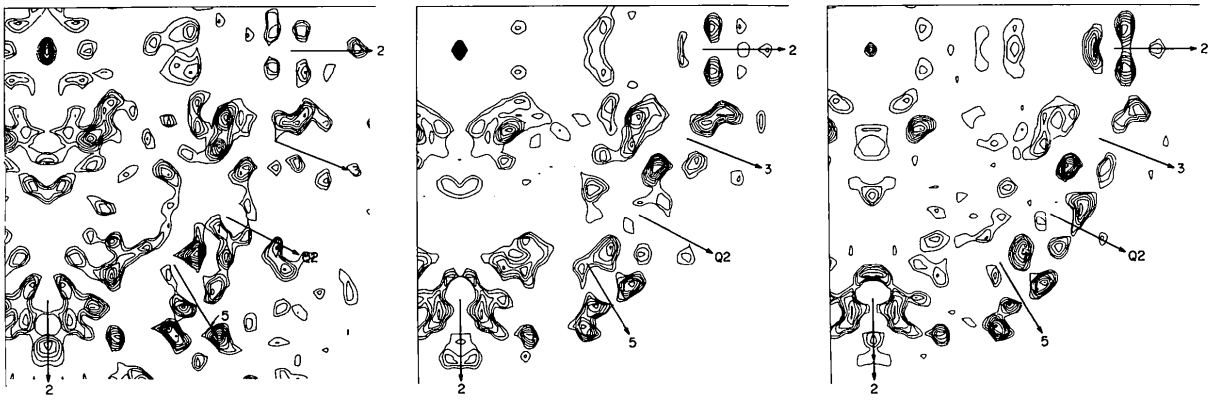


Fig. 10. Equatorial sections viewed down an icosahedral twofold axis through the virus showing improved quasi-symmetry. Maps were computed (*a*) with single isomorphous replacement phases, (*b*) by averaging the single isomorphous replacement electron density and (*c*) from phases after two cycles of molecular replacement.

Phase refinement by real-space averaging

Phase refinement and phase improvement by utilization of the redundant information present in electron density maps containing multiple copies of chemically identical subunits has been employed in a number of recent crystallographic studies of biological macromolecules (Champness, Bloomer, Bricogne, Butler & Klug, 1976; Winkler *et al.*, 1977; Argos, Ford & Rossmann, 1975). The 22·5 Å structure of SBMV (Johnson *et al.*, 1976) was also obtained with the real-space averaging technique (Buehner, Ford, Moras, Olsen & Rossmann, 1974; Bricogne, 1974, 1976).

In the 11 Å structure determination of SBMV, a single isomorphous electron density map (Rossmann & Blow, 1961) was calculated on a grid suitable for 11 Å resolution (3·4 Å per grid step in x and y, 3·7 Å per grid step in z). The map was then averaged over the ten non-crystallographic icosahedral asymmetric units. The averaged map, with background set to the average

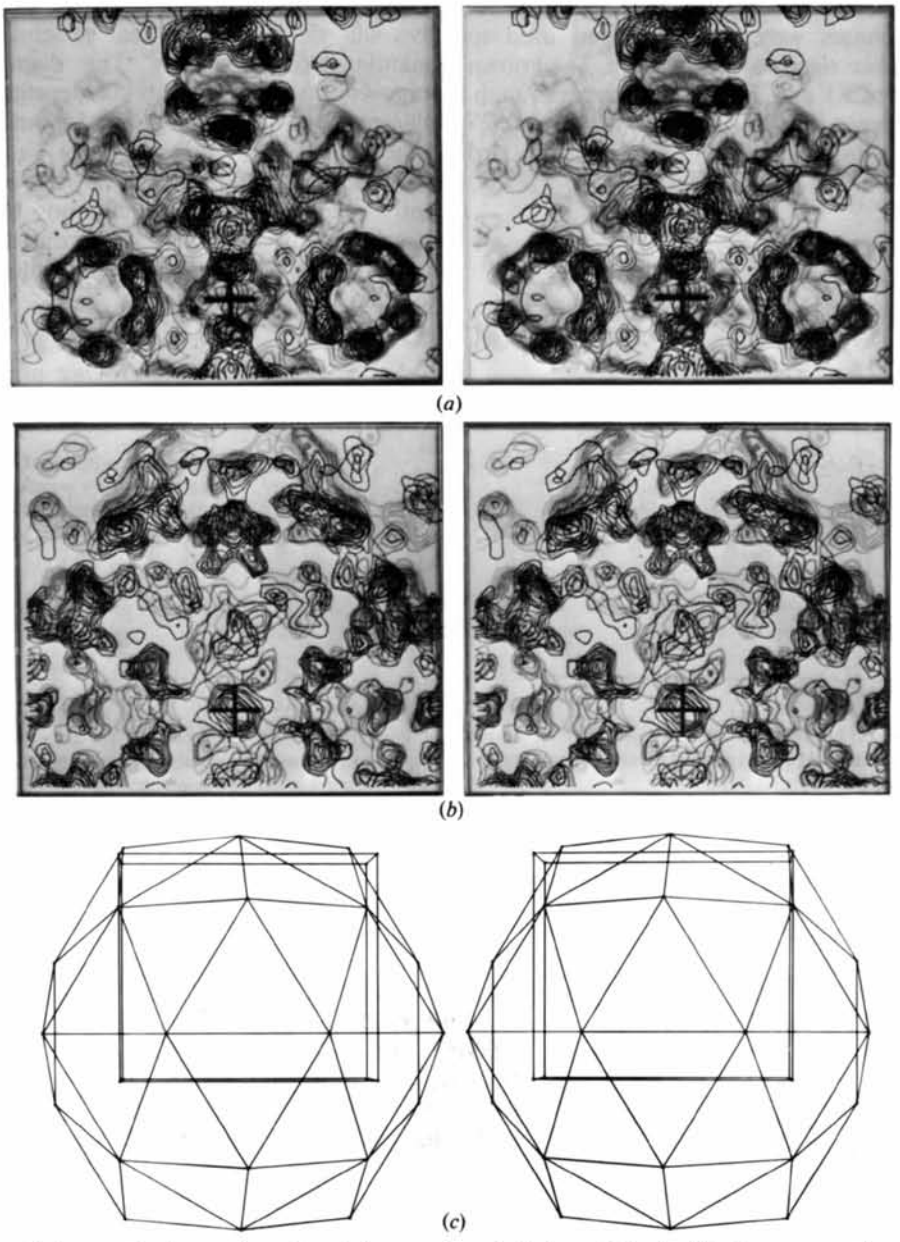


Fig. 11. Stereoview of electron density sections viewed down an icosahedral twofold axis. The heavy-atom sites are shown for one icosahedral unit. (*a*) Sections $R = 118$–145 Å, near the surface of the virus. Density at the very top of the figure is associated with an adjacent virus particle. Particle contacts are directly along the fivefold icosahedral symmetry axes. (*b*) Sections $R = 100$–115 Å, further toward the center of virus. (*c*) Stick model of $T = 3$ surface lattice oriented as are the electron density maps in (*a*) and (*b*).

electron density outside the molecular envelope, was Fourier transformed to obtain calculated amplitudes and phases. A new electron density map was computed using calculated phases and observed amplitudes modified by Sim weights (Sim, 1959, 1960). Two iterations of this procedure were necessary to obtain reasonable convergence (Table 5). No use was made of the isomorphous replacement phases after the initial calculation of the single isomorphous map.

A comparison is shown in Fig. 10 of the single isomorphous replacement electron density, of the density after averaging and of the density distribution after phase refinement. An equatorial section, perpendicular to an icosahedral twofold axis, was chosen for this purpose. Evidence for the validity of the phase-refinement technique is apparent by the improvement of the quasi-twofold axis in the plane of this section.

Description of the structure

The electron density described below is the result of two cycles of molecular replacement with a final averaging. Parts of the virus electron density are shown in Fig. 11(*a*) (near the surface) and in Fig. 11(*b*) (below the surface). The virus is viewed down an icosahedral twofold axis. The region of the electron density displayed is indicated in Fig. 11(*c*).

The distribution of density in the present map is similar to that previously observed at 22·5 Å resolution, but it shows more detail. In the earlier low-resolution map, the principle features were rings of density about the quasi-sixfold axis and fivefold axis and a blob of density at the quasi-threefold axis. The density distribution suggests a bilobal protein subunit with one lobe at a quasi-threefold axis and the other near a quasi-sixfold axis or a fivefold axis. The most striking feature of the 11 Å map is the clustering of six distinct lobes about the quasi-sixfold axis and five lobes about the fivefold axis. Furthermore, three distinct lobes are visible around, but not on, the quasi-threefold axis.

The hexamer and pentamer clusters are different in appearance from the clusters about the quasi-threefold axis. The former show six features which gradually weaken with decreasing radius and eventually approach each other forming a 'cup'. The densities around the quasi-threefold axis are flattened and splay away toward the quasi- or icosahedral twofold axis as the radius decreases, forming a connection between the two lobes.

The heavy-atom positions have been marked on Fig. 11 for three subunits related by a threefold axis. The sites are surprisingly far from the virus surface, at an average radius of 125 Å, yet they are easily accessible to solvent. They occur in the neck of density between the lobes of the protein subunit not far from the quasi- or icosahedral twofold axes.

In general, the $T = 3$ symmetry is well maintained near the surface of the virus. However, at about a radius of 118 Å, there is very little density, while at even smaller radii the density is weak and does not obey the quasi-symmetry. It must, therefore, be presumed that most of the features at radii less than 118 Å represent ordered RNA structure. As the radius decreases further the amount of ordered RNA decreases, as judged by the progressive weakening of distinct electron density features.

Conclusions

The reliability of the result of the SBMV structure investigation at 11 Å resolution is supported by: (1) the electron density features as derived by single isomorphous replacement are consistent with the features of the 22·5 Å-resolution *ab initio* molecular replacement result; (2) the electron density map is consistent with $T = 3$ symmetry near the surface of the virus corresponding to the protein coat; (3) the internal features lack symmetry, as would be expected if this space were occupied by RNA; (4) feedback experiments to verify the heavy-atom positions were successful. Thus, with the solution of the heavy-atom positions in the first isomorphous derivative, the major technical problems for a high-resolution determination have been solved. Although the distribution of heavy atoms is pseudo-centric, the present phase determination has been employed to deconvolute the heavy-atom sites of other derivatives, permitting the computation of a 5 Å-resolution electron density map (Suck, Rayment, Johnson & Rossmann, 1977).

We are grateful to Dr Patrick Argos for helpful discussions and to Sharon Wilder for assistance in the preparation of the manuscript. We would like to thank the Purdue University Computer Center staff for assistance.

This work was supported by the National Institutes of Health (grant No. AI 11219), the National Science Foundation (grant No. BMS74-23537) and a small supply grant from the Eli Lilly Co. JEJ was the recipient of a National Institutes of Health postdoctoral fellowship (No. 6 F22 CA00557) and DS of a Deutsche Forschungsgemeinschaft postdoctoral fellowship.

APPENDIX I

Molecular-weight determination of the southern bean mosaic virus coat-protein subunit*

The sedimentation rate (*S* value) of native SBMV virus particles was determined by zonal gradient sedimen-

* By Karl Lonberg-Holm and Bruce D. Korant.

tation. A sample containing 900 μg of purified SBMV was mixed with purified ^{14}C-labeled polio virus type 2 (Lonberg-Holm, Gosser & Kauer, 1975) and diluted to a volume of 100 μl with saline phosphate (0·15 *M* NaCl containing 0·05 *M* sodium phosphate, pH 7·3). The sample was layered on a linear sucrose gradient (4·9 ml volume, 10–25% sucrose in saline phosphate) and was centrifuged 50 min at 45 000 r.p.m. and at 2° in an SW50.1 rotor (Spinco Division, Beckman Instruments Inc.). The tube containing the gradient was punctured at the bottom, and fractions containing 2 drops each were collected. Aliquots of each fraction (20 μl) were counted for ^{14}C, and then 3·0 ml water was added to the remaining portion of each fraction and the adsorption at 260 nm was measured. It was assumed that the polio-virus marker sedimented at 156*S* (Korant, Lonberg-Holm, Noble & Stasny, 1972). The sedimentation coefficient of SBMV was calculated as 112*S*.

The molecular weight of the coat-protein subunit of SBMV was determined by gel electrophoresis in sodium dodecyl sulfate (SDS) of the dissociated virus particles. Purified SBMV virus particles (1 mg ml^{-1}) were heated to 100°C in 0·1 *M* tris(hydroxymethyl)aminomethane(tris) buffer, pH 6·8, containing 1% SDS and 0·1% β-mercaptoethanol (Korant *et al.*, 1972). The dissociated, reduced proteins were then alkylated with 0·1 *M* iodoacetamide to block sulfhydryl groups, and dialyzed against 500 volumes of 0·1 *M* tris buffer, pH 6·8, containing 1% SDS and 7·5% glycerol. The SBMV protein subunit was analyzed by electrophoresis in a thin-layer polyacrylamide gel with SDS present (Knight, 1975). Included in parallel tracks of the gel were well-characterized proteins, including polio-virus coat proteins, lysozyme, chymotrypsinogen, ovalbumin and serum albumin. The proteins were fixed and stained upon completion of electrophoresis by placing the gel in 0·25% Coomassie blue in 10% acetic acid, 50% methanol and destaining in 7% acetic acid. The molecular weight of the SBMV subunit was estimated by measuring its mobility compared with standards.

From the known relation between electrophoretic mobility and molecular weight of proteins in these gels (Weber & Osborn, 1969), it was calculated that the SBMV coat-protein subunit, cowpea strain, has a molecular weight of 28 250 ± 250.

APPENDIX II

Averaging of electron density maps*

The method of averaging electron density, developed for the determination of lobster glyceraldehyde 3-phosphate dehydrogenase, used skew planes to represent the density on a grid related to the non-crystallographic point group. Non-crystallographically equivalent grid points could then be readily averaged. Bricogne (1976) has described a 'double sorting' approach used in studies of *B. stearothermophilus* glyceraldehyde 3-phosphate dehydrogenase and tobacco mosaic virus disk protein. The procedure described below, developed independently by the author, is similar to that of Bricogne.

It will be assumed that the orientation and point of intersection of the non-crystallographic axes are known, and an envelope can be defined to contain all the molecule or particle about that center. In stage one of the program (Fig. 12*a*), each grid point in the crystallographic asymmetric unit is transformed into one chosen

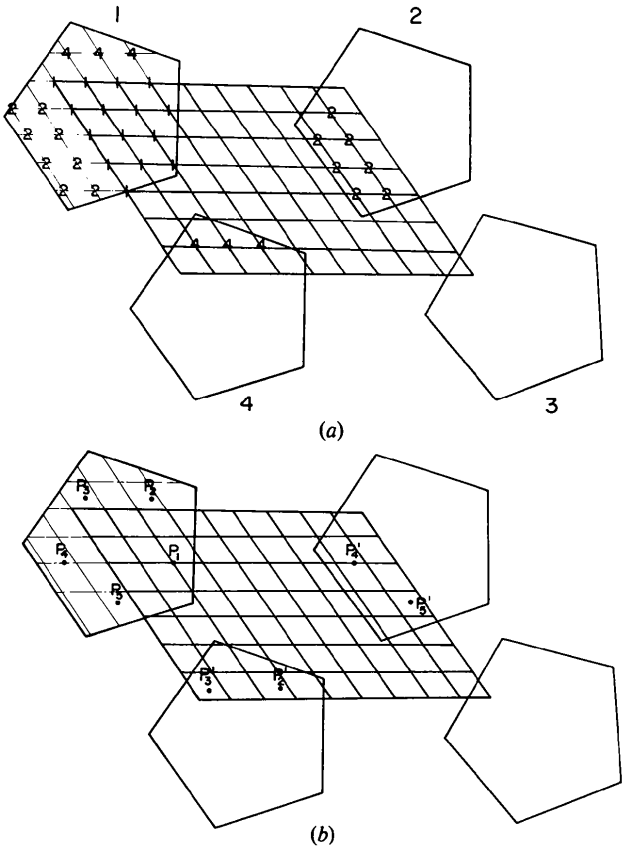


Fig. 12. A graphical representation of the averaging procedure described in the text. (*a*) The grid representing the asymmetric unit of a two-dimensional crystal with a single pentagon in the unit cell. Four pentagons contribute to the asymmetric unit. Step 1 of the averaging procedure is to fill the molecular envelope of one molecule by operating with crystallographic symmetry operators on all grid points in the asymmetric unit. (*b*) Averaging over the non-crystallographic symmetry. Each grid point within the asymmetric unit is operated on by the non-crystallographic symmetry operator to generate four non-crystallographically equivalent points (the operation has been carried out on only one grid point for clarity). These non-integral points are then transformed back to the equivalent position within the asymmetric unit to be evaluated.

* By John E. Johnson.

virus particle by testing to determine whether it is within or outside the chosen particle after a suitable crystallographic symmetry operation has been applied. The grid-point coordinates in the crystallographic asymmetric unit, a flag indicating the result of the test, and a number identifying the required symmetry operator are packed into one computer word and saved on tape I. It is useful to sequence the testing of grid points such that all points corresponding to sections in which the electron density is calculated are grouped together.

In stage 2 (Fig. 12*b*), each grid point generated in stage 1 is multiplied by the appropriate non-crystallographic symmetry operators. Since the points generated by non-crystallographic symmetry may lie in any region of the molecular envelope, it is necessary to transform those points beyond the limits of the crystallographic asymmetric unit to the related point within the asymmetric unit. Only the points originally on tape I will now necessarily lie on integral grid positions. Every set of non-crystallographically related points is given an index identifier. Each point together with the appropriate index identifier is packed into one word and stored on tape II. Thus for type II SBMV crystals, there will be ten times the number of points on tape II as on tape I. The contents of tape II are then sorted on the section number nearest to the non-integral points in the grid to facilitate efficient electron density evaluation.

It should be noted that the parameters generated on tape II depend only on the space group of the crystal, the grid size, the selected molecular center and the orientation of the non-crystallographic symmetry elements. As no reference has yet been made to the electron density, the previously described tape generation is carried out only once for a given grid on which the electron density is calculated.

To average the electron density, it is first necessary to evaluate the density at all the non-integral grid points. The electron density map is read sequentially, with two sections stored in the computer at one time. Each point stored on tape II lying on or between these sections is evaluated by interpolation according to the formula

$$\rho = \rho_k + \Delta Z(\rho_{k+1} - \rho_k)$$

and

$$\rho_k = \rho_{11k} + \Delta Y(\rho_{12k} - \rho_{11k}) + \Delta X(\rho_{21k} - \rho_{11k}) + \Delta X \Delta Y(\rho_{11k} + \rho_{22k} - \rho_{21k} - \rho_{12k})$$

where ρ_{ijk} are sampled electron density values at the point (x,y,z); $\Delta X = (x - x_1)/(x_2 - x_1)$; $\Delta Y = (y - y_1)/(y_2 - y_1)$; and $\Delta Z = (z - z_1)/(z_2 - z_1)$ where $x_1 \le x < x_2, y_1 \le y < y_2$, and $z_k \le z < z_{k+1}$.

These densities and the associated index identifiers are then stored on tape III. This tape is sorted on the index identifiers, thus placing each set of non-crystallographically related densities together and, furthermore, they will be in the same order as the grid points on tape I. The mean density and standard deviation for each of these sets is then calculated. A map containing the averaged electron density at all grid points in the asymmetric unit is generated by reading the grid locations from tape I and assigning to them the averaged density calculated from tape III. If the point does not lie within the molecular envelope, a density value can be inserted corresponding to the average solvent density, a density value specified by the user, or by the original solvent density. In order for the averaged electron density map to be in the same orientation as was the unaveraged map, it is essential that the grid points on tape I are generated in a sequence corresponding to sections of the unaveraged electron density.

The averaged density can then be Fourier transformed. The resultant calculated phases, after combining with isomorphous-replacement phases, can be used to recompute an improved electron density map. The new map can again be averaged by using the information already on tapes I and II.

References

ADAMS, M. J., HAAS, D. J., JEFFERY, B. A., MCPHERSON, A. JR, MERMALL, H. L., ROSSMANN, M. G., SCHEVITZ, R. W. & WONACOTT, A. J. (1969). *J. Mol. Biol.* **41**, 159–188.

AKIMOTO, T., WAGNER, M. A., JOHNSON, J. E. & ROSSMANN, M. G. (1975). *J. Ultrastruct. Res.* **53**, 306–318.

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

ARGOS, P. & ROSSMANN, M. G. (1974). *Acta Cryst.* A**30**, 672–677.

ARGOS, P. & ROSSMANN, M. G. (1976). *Acta Cryst.* B**32**, 2975–2979.

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

BRICOGNE, G. (1976). *Acta Cryst.* A**32**, 832–847.

BUEHNER, M., FORD, G. C., MORAS, D., OLSEN, K. W. & ROSSMANN, M. G. (1974). *J. Mol. Biol.* **82**, 563–585.

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

CHAMPNESS, J. N., BLOOMER, A. C., BRICOGNE, G., BUTLER, P. J. G. & KLUG, A. (1976). *Nature (London)*, **259**, 20–24.

FORD, G. C. (1974). *J. Appl. Cryst.* **7**, 555–564.

GHABRIAL, S. A., SHEPHERD, R. J. & GROGAN, R. G. (1967). *Virology*, **33**, 17–25.

HAMILTON, W. C., ROLLETT, J. S. & SPARKS, R. A. (1965). *Acta Cryst.* **18**, 129–130.

HARRISON, S. C. (1968). *J. Appl. Cryst.* **1**, 84–90.

HARRISON, S. C. & JACK, A. (1975). *J. Mol. Biol.* **97**, 173–191.

HILL, J. H. & SHEPHERD, R. J. (1971). *Phytopathology*, **61**, 1408–1409.

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

KNIGHT, E. (1975). *J. Biol. Chem.* **250**, 4139–4144.

KORANT, B. D., LONBERG-HOLM, K., NOBLE, J. & STASNY, J. T. (1972). *Virology*, **48**, 71–86.
LENTZ, P. J. JR, STRANDBERG, B., UNGE, T., VAARA, I., BORELL, A., FRIDBORG, K. & PETEF, G. (1976). *Acta Cryst.* A**32**, 2979–2983.
LONBERG-HOLM, K., GOSSER, L. B. & KAUER, J. C. (1975). *J. Gen. Virol.* **27**, 329–342.
MATTHEWS, B. W., KLOPFENSTEIN, C. E. & COLMAN, P. M. (1972). *J. Phys. E*, **5**, 353–358.
MILLER, G. L. & PRICE, W. C. (1946). *Arch. Biochem.* **10**, 467–477.
PRICE, W. C. (1946). *Am. J. Bot.* **33**, 45–54.
ROSSMANN, M. G. (1976). *Acta Cryst.* A**32**, 774–777.
ROSSMANN, M. G. & BLOW, D. M. (1961). *Acta Cryst.* **14**, 1195–1202.
ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24–31.
SEHGAL, O. P. & SINHA, R. C. (1974). *Virology*, **59**, 599–608.
SIM, G. A. (1959). *Acta Cryst.* **12**, 813–815.
SIM, G. A. (1960). *Acta Cryst.* **13**, 511–512.
SUCK, D., RAYMENT, I., JOHNSON, J. E. & ROSSMANN, M. G. (1977). *Virology*. In the press.
TREMAINE, J. H. (1966). *Virology*, **30**, 348–354.
WEBER, K. & OSBORN, M. (1969). *J. Biol. Chem.* **244**, 4406–4412.
WINKLER, F. K., SCHUTT, C. E., HARRISON, S. C. & BRICOGNE, G. (1977). *Nature (London)*, **265**, 509–513.
YPHANTIS, D. A. (1964). *Biochemistry*, **3**, 297–317.

Acta Cryst. (1978). B**34**, 578–584

The Crystal Structure of Colchicine. A New Application of Magic Integers to Multiple-Solution Direct Methods

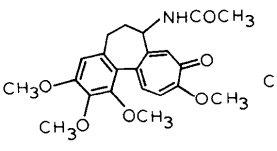
BY L. LESSINGER* AND T. N. MARGULIS

Department of Chemistry, University of Massachusetts–Boston, Boston, Massachusetts 02125, *USA*

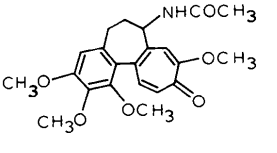
(*Received* 15 *June* 1977; *accepted* 9 *September* 1977)

The mitotic spindle inhibitor colchicine, $C_{22}H_{25}NO_6$, crystallizes as a dihydrate in space group $P2_1$, a = 17·08, b = 10·70, c = 13·88 Å, β = 117·9°, $Z = 4(C_{22}H_{25}NO_6.2H_2O)$/unit cell. The crystal structure was solved by a multiple-solution direct method in which unknown starting phases φ_i are represented using 'magic integers': $\varphi_i = m_i x$. An appropriate choice of integers m_i and sampling of the variable x allows a drastic reduction in computing time and a great increase in structure-solving capability compared to the widely used method in which all unknown phases are permuted among the four values 45, 135, 225, 315°. The crystal structure of colchicine dihydrate was refined by least squares to R = 0·052 for 2322 observed X-ray reflections. The two independent colchicine molecules have very similar conformations in the crystal. The troponoid rings have alternating bond lengths, and are not precisely planar. These rings make dihedral angles with the planar benzene rings of 53° in one molecule of colchicine, 51° in the other. The four independent water molecules are all found in a distinct region in the crystal, which is held together by a complex hydrogen-bond network. Two methoxy-group O atoms of colchicine act as hydrogen-bond acceptors, the one on the troponoid ring participating in a bifurcated hydrogen bond.

Colchicine (I), the principal active substance of the autumn crocus, has a remarkable range of biological effects. The most studied of these is its arrest of mitosis, believed to result from the specific strong binding of colchicine to the protein tubulin, preventing the assembly of tubulin into microtubules which form the mitotic spindle (Soifer, 1975). Colchicine is also capable of relieving the pain of gout (for which it has been used since ancient times), of inducing polyploidy in plants, and of causing particular malformations in embryos (Eigsti & Dustin, 1955).



(I)



(II)

To contribute toward understanding the specific interaction with tubulin, and other biological activity,

* Present address: Department of Chemistry, Barnard College, Columbia University, New York 10027, USA.