V.1. Structure and Models

THE STRUCTURE OF SOUTHERN BEAN MOSAIC VIRUS AT 5 Å RESOLUTION

M. G. Rossmann, J. E. Johnson, D. Suck, I. Rayment, and A. G. W. Leslie

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

Introduction

Southern bean mosaic virus is a small spherical virus of molecular weight $6.6 \times 10^6$ [1] containing 21 % RNA by mass [2]. The single protein component in the coat [3] has a molecular weight of 29,000 implying the presence of 180 protein subunits in the virus particle [4], a number consistent with $T = 3$ quasi-symmetry [5]. The virus has been crystallized in a number of different space groups [6–8] (Table 1) and the diffraction data were shown to be consistent with icosahedral symmetry [9].

The nearest neighbor contacts of 284 Å in the rhombohedral type I crystals are approximately along the direction of the threefold axes. However, in the rhombohedral type II crystals the particle contacts are directly along fivefold axes with a center-to-center distance of 318 Å. The radial distribution of intensities (vide infra) shows that the mean diameter of the virus is 284 Å, independent of the crystallization conditions. The difference in apparent radii thus reflects an extension of the virus particle along the fivefold axes.

The best diffracting crystals (Fig. 1) are type II in space group R32 (Table 1). These crystals contain one virus particle per unit cell located at the intersection of one threefold and three twofold axes (Fig. 2). Thus the crystallographic asymmetric unit contains one-sixth of the particle. Furthermore, averaging for molecular replacement [10–12] can be done between the ten different noncrystallographically related icosahedral units within the crystallographic asymmetric unit.

The 22.5 Å Resolution Map

In order to establish the initial phasing model, structure amplitudes were averaged within thin shells of reciprocal space and plotted as a function of resolution [34]. Variation of these amplitudes corresponded to the transform of a uniform sphere of diameter 284 Å,
<table>
<thead>
<tr>
<th>Type</th>
<th>Space group</th>
<th>Cell lengths (Å)</th>
<th>Angle</th>
<th>Setting</th>
<th>$Z$</th>
<th>$V_M$ (Å²/dalton)</th>
<th>Fraction of virus in asymmetric unit</th>
<th>Morphology</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magdoff's</td>
<td>C222₁</td>
<td>295 508 474</td>
<td>—</td>
<td>Orthorhombic</td>
<td>4</td>
<td>2.7</td>
<td>1/2</td>
<td>Rhombs. with threefold axis through obtuse corner</td>
<td>[6]</td>
</tr>
<tr>
<td>Type I</td>
<td>R32</td>
<td>542 542 542</td>
<td>$\alpha = 116^\circ 42'$</td>
<td>Rhombohedral</td>
<td>3</td>
<td>3.6</td>
<td>1/2</td>
<td>Rhombs. with threefold axis through acute corner</td>
<td>[7]</td>
</tr>
<tr>
<td>Type II</td>
<td>R32</td>
<td>318 318 318</td>
<td>$\alpha = 64^\circ 0'$</td>
<td>Rhombohedral</td>
<td>1</td>
<td>3.6</td>
<td>1/6</td>
<td>Diamond-shaped plates with the (001) face parallel to the plate</td>
<td>[8]</td>
</tr>
<tr>
<td>Type III</td>
<td>C222₁</td>
<td>522 330 530</td>
<td>—</td>
<td>Orthorhombic</td>
<td>4</td>
<td>3.5</td>
<td>1/2</td>
<td></td>
<td>[8]</td>
</tr>
</tbody>
</table>
Fig. 1. Diffraction photographs of type II R32 crystals. (a) Precession photograph of $h0l$ layer with $\mu = 4^\circ$. Observe the radial streaks of strong intensities corresponding to the icosahedral five-, three- and twofold directions. (b) Oscillation photograph with $1^\circ$ range taken on a rotating anode X-ray generator with two perpendicular focusing mirrors.
Fig. 2. Stereoscopic view of the SBMV particle in the type II crystal form. Both crystallographic (axes penetrating the icosahedron) and noncrystallographic icosahedral symmetry elements are shown. The a, b, and nonlabeled diagonal axes are twofold symmetry elements and the c-axis is a threefold symmetry axis.

Fig. 3. Structure amplitudes, averaged within shells of reciprocal space, shown in relationship to the Fourier transform of a sphere of diameter 284 Å. Inset is shown the complete spherical transform from infinity to 30 Å resolution.

to at least 35 Å resolution (Fig. 3). An electron density map was then calculated to 35 Å resolution utilizing the spherical phases and then averaged over the 10 noncrystallographic asymmetric units of the icosahedron. To avoid the effects of spurious features common at small radii, all density at less than 100 Å radius was set to a uniform value equal to the mean virus density. Density outside the virus envelope was set to zero. This averaged and modified map was Fourier transformed to give calculated structure factors to a resolution slightly higher than the starting 35 Å resolution. A revised electron density map was computed at this higher resolution based on the calculated phases and observed amplitudes modified with Sim’s weighting scheme [13, 14]. The resolution in subsequent cycles was increased by about one reciprocal lattice interval at a time. The final map at 22.5 Å resolution was
obtained after six cycles. Further extension was considered inappropriate in view of the increasing size of the imaginary portion of the structure factors with resolution.

The electron density of the virus at 22.5 Å resolution [34] is shown in Fig. 4a and 4b, while Fig. 4c (p. 554) shows a $T = 3$ model that is similarly oriented. Figure 4a shows the top segment of the particle viewed down the icosahedral threefold axis, which is also a quasi-sixfold axis, a feature distinctly apparent in the map. It is surrounded by three fivefold axes which protrude by about 15 Å from the remaining surface of the particle. Both the icosahedral fivefold axes and quasi-sixfold axes are surrounded by a ring of density. This is consistent with the demands of $T = 3$ quasi-symmetry which require the similar packing of five and six subunits with respect to these axes. In Fig. 4 it is apparent that SBMV [15] shows strong hexamer and pentamer clustering, but in addition there is a cluster around the quasi-threefold axes.

Mellema and Amos [16] showed that the protein associated around the fivefold axis of TYMV protrudes above the surface of the 300 Å diameter sphere by about 10 Å, a situation almost identical to SBMV. This similarity is probably due to equivalent solutions for accommodating the quasi-symmetry. The TYMV protein subunit has a MW of 20,000, about 9000 less than the SBMV subunit. In view of the similar size of the two viruses it would seem plausible that the extra 9000 MW protein of SBMV is clustered around the quasi-threefold axis, a volume essentially vacant in TYMV. This difference also expresses itself in the appearance of the virus in electron micrographs. While TYMV shows good differentiation between its morphological features, SBMV with its apparently higher surface protein density shows no distinct morphology [8, 17, 18].

**Determination of Heavy Atom Positions**

Systematic searches for heavy atoms in chemically identical environments related by non-crystallographic symmetry can be used when a straightforward interpretation of the Patterson would be difficult or impossible [19, 20]. These techniques have been used to determine the heavy atom positions for the icosahedral $T = 1$ STNV [21]. 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 be considered, 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 first derivative obtained for SBMV was prepared by soaking $K_2HgI_4$ into native crystals [22]. An 11 Å resolution difference Patterson was computed between the $K_2HgI_4$ derivative and native data. Trial heavy atom positions were used in a search within the non-crystallographic asymmetric unit limited to radial distance $R$ from the particle center to between 118 Å and 157 Å. The search function was inspected for peaks which were related by an approximate (quasi) threefold axis. This was readily accomplished by constructing a trigonal ruler of the type depicted in Fig. 5. 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 were not radial. For instance, the quasi-threefold axes are roughly parallel
Fig. 4a. Stereoscopic view of the 22.5 Å electron density map. In (a) are shown sections 10-24 at the top of the virus, and in (b) are sections 0-14 near the equator. The unit cell was cut in 96 sections perpendicular to $c$, corresponding to an interval of 8 Å.

In (c) is shown a stick model, oriented as in (a) and (b), showing $T = 3$ quasi-symmetry.
to the adjacent icosahedral twofold axes in TBSV [23, 24]. By far the largest peaks in the search function (Fig. 5) were related by a quasi-threefold axis approximately parallel to its neighboring twofold axis. These (A sites) were then refined by a least squares procedure.

Fig. 5. Difference Patterson search function results for radii of 130 Å (right) and 121 Å (left) of the $K_2\text{HgI}_4$ derivative. The three major peaks are related by an approximate threefold axis positioned near the center of the icosahedral asymmetric unit.
using single isomorphous replacement data, adapted to include icosahedral symmetry [25]. Thus only three atoms were refined independently, although 30 atoms were positioned by icosahedral symmetry within the crystallographic asymmetric unit.

The $K_2HgI_4$ single isomorphous replacement phases were used to compute an 11 Å resolution difference Fourier for the PHMBS derivative. This map showed a set of three quasi-threefold related B sites as well as peaks at the previous A sites. The B sites were now least-squares refined at 5 Å resolution and were used to calculate a new PHMBS difference Fourier. This map again showed substitution at all three A sites. Thus while $K_2HgI_4$ was substituted only at the A sites, PHMBS was substituted at A and B sites (Fig. 6). At a later stage, it was discovered that PHMBS had a minor substitution at three quasi-threefold related H sites.

With double isomorphous phases to 11 Å and single isomorphous phases from 11 Å to 5 Å resolution, it was now relatively easy to solve a uranyl acetate and a $PtCl_4$ derivative. The uranyl acetate derivative had rather special substitution on the quasi-threefold axes, site C, and on the icosahedral fivefold axes, site D (Fig. 6). The $PtCl_4$ derivative had three sets of quasi-threefold related sites (E near B, F near A, and G near the five- and quasi-sixfold axes). Presumably, the A, B, and H sites are cationic substitutions; the E, F, and G sites are probably methionine substitutions; and the C and D sites are cation sites possibly replacing the Ca$^{++}$ ions which hold the virus together.

**The 5 Å Resolution Map**

A multiple isomorphous replacement electron density map, based on the 11 Å $K_2HgI_4$ data and the 5 Å PHMBS and UAc$_2$ data, was computed on a grid of roughly 1.4 Å spacing. It was averaged over the 10 icosahedral units within the crystallographic asymmetric unit and displayed looking down an icosahedral twofold axis [26].
The overall features of the map were closely consistent with the earlier 22.5 Å map based only on molecular replacement. There were strong features around the quasi-threefold axis and icosahedral five- and threefold axes at a radius of about 140 Å (Fig. 7).

![Stereo view of radial sections at 5 Å resolution showing the protein coat between 133 Å < R < 148 Å. The T = 3 symmetry is clearly visible close to the quasi-threefold axis. Related features distant from this axis do not appear equivalent since the quasi-threefold axis is not radial.](image)

Two distinct characters in the electron density were immediately apparent. Between 118 Å and 155 Å radius, the electron density shows empty troughs between high and long ridges of density (Fig. 7). At about 118 Å radius the electron density, however, changes its appearance dramatically. At less than 118 Å radius the electron density is lower, shows smaller gradients, and contains discrete round peaks in ribbons of density separated by about 7 Å (Fig. 8). The two types of features are separated by a thin shell containing little density (Fig. 9). Furthermore, the $T = 3$ symmetry is clearly visible in the outer coat and quite absent in the interior. As the radius decreases below 118 Å, the features grow steadily.

![Stereo view of RNA portion of virus at 5 Å resolution between 103 Å < R < 118 Å. Notice the lack of $T = 3$ symmetry and the diffuse phosphate lumps separated by about 7 Å. The character of this part of the map differs remarkably from that seen in Fig. 7.](image)
less distinct and more diffuse. It is reasonable to assume that the crisp quasi-threefold related features near the outside of the virus are protein, whereas the inner features are RNA. The nucleic acid shows its greatest order where it is in close contact with the protein. These results are in reasonable agreement with those of Jacrot et al. [27] who used neutron diffraction and with those of Weintraub and Ragetti [18] who used electron microscopy.

Extensive lengths of RNA can be seen in Fig. 8. It is not possible to trace a single chain and, in some places, a number of chains appear to meet at a single point. Presumably this represents alternate RNA structures in different parts of the virus which are here superimposed by the crystallographic and noncrystallographic averaging. Although the RNA is mostly separated from the protein shell, regions of contact exist. At least 26 distinct phosphate sites can be counted in one icosahedral asymmetric unit. Thus at least 36% of the RNA has some form of order.

The 3.5 Å Resolution Map

An icosahedrally averaged map, based on 3.5 Å resolution data of the native SBMV, the PHMBS, UAc₂, and PtCl₄ derivatives as well as the 11 Å resolution K₂Hgl₄ data, was computed immediately prior to the Madras meeting. It shows all the same features as the earlier 5 Å map but, as expected, in greater detail. It has not been analyzed in detail. However in Fig. 10 is shown a model, as seen in a Richards optical comparator, of the three central quasi-threefold related α-helices (see central part in Fig. 7). It was possible to discern the hand of the helix (i.e. to select the correct hand of the map), the polarity of the helix direction (which end is N and which is C), and to recognize and possibly identify a few side chains. These side chains were fully equivalent in all three helices. The full polypeptide chain has, however, not been traced as yet at the time of writing.
Fig. 10. The three central α-helices, related by a quasi-threefold axis of side chains and the superimposed secondary structure, are seen in a Richards optical comparator. Note the clear indication of helix direction.
Swollen SBMV

Sehgal's group has shown [28] that, by removing Ca\(^{++}\) ions with EDTA at pH 7.5, the virus swells about 20% in diameter. It then becomes susceptible to RNA digestion and SDS degradation. Durham et al. [29, 30] have suggested that the lower Ca\(^{++}\) ion concentration within the cell cytoplasm is the trigger to the disassembly of many viruses.

Fig. 11. Crystals of swollen, metal-free SBMV. Length of crystals is about 0.5 mm.

It has been possible to grow swollen SBMV particles into crystals which diffract to at least 4 Å resolution (Fig. 11). It is planned to pursue a complete structural investigation of these crystals. This would thus permit a study of the conformational changes involved in disassembly both in terms of protein–protein and protein–RNA interactions.

Conclusions

The prospect of following the complete SBMV polypeptide chain at 3.5 Å resolution for each of the three quasi-equivalent protein subunits in one icosahedral subunit is good. Similarly, the prospect of observing about half of the RNA secondary structure is reasonable. Furthermore, it may be possible to investigate a change of conformation on removal of the Ca\(^{++}\) ions from the virus coat. Similar structural investigations, at an even more advanced stage, are already available for native TBSV [31] and TMV disc protein [32]. Other investigations are in process on TMV [33], STNV [21] and a variety of crystallized virus proteins and glycoproteins. Thus the next decade is likely to see a significant advance in the structural knowledge of viruses and, hopefully, its implication to viral functions.
Acknowledgements. This work was supported by National Institutes of Health grant no. AI 11219 and National Science Foundation grant no. BMS74-23537. We are most grateful for the very able assistance of Sharon Wilder; the Purdue University Computer Center and its staff; and Eric Henss, our instrument designer.

References

15. Abbreviations used: PHMBS, p-hydroxymercuribenzenesulfonate; SBMV, southern bean mosaic virus; STNV, satellite tobacco necrosis virus; TBSV, tomato bushy stunt virus; TMV, tobacco mosaic virus; TYMV, turnip yellow mosaic virus and UAc, uranyl acetate.