

# Tomato bushy stunt virus at 5.5-Å resolution

F. K. Winkler, C. E. Schutt & S. C. Harrison\*

Gibbs Laboratory, Harvard University, 12 Oxford St., Cambridge, Massachusetts 02138

G. Bricogne

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

*The coat of tomato bushy stunt virus is built from protein subunits having rigid domains connected by a flexible hinge. Two states of the hinge are present in the  $T = 3$  icosahedral structure. Each subunit has a binding site for RNA on its inner surface.*

DESIGN principles of regular viruses were first clarified by Caspar and Klug<sup>1</sup> who showed that quasi-equivalence could reconcile the variability of subunit packing (required to construct a shell of sufficient size) with the specificity of subunit bonding, necessary for self-assembly. The small, spherical RNA viruses pose two important structural questions: how is quasi-equivalent packing of identical protein subunits achieved, and how is the nucleic-acid chain accommodated in the particle? We describe here a 5.5-Å resolution electron-density map of tomato bushy stunt virus (TBSV), and derive some initial answers.

The TBSV particle comprises a single molecule of RNA (MW  $1.5 \times 10^6$ ), 180 coat protein subunits (MW 41,000 each) and perhaps one copy of a larger polypeptide<sup>2-4</sup>. The coat protein forms a  $T = 3$  icosahedral surface lattice, with prominent dimer clustering at the outside of the particle<sup>5</sup>. This design implies that a single kind of polypeptide chain packs in relation to its neighbours in three slightly different ways (Fig. 1). X-ray crystallography<sup>6,7</sup> has revealed the orientation of local symmetry axes and suggested that the differences between quasi-equivalent contacts is related to RNA-induced perturbations.

The structure has been solved using multiple isomorphous replacement, combined with phase refinement by icosahedral averaging. Problems for data collection due to the large unit cell were solved using a focusing X-ray camera<sup>8</sup>, and a strategy for oscillation photography appropriate to the restriction that only one exposure can be obtained per crystal<sup>9</sup>. As in the 5-Å resolution study of the tobacco mosaic virus disk<sup>10</sup>, a problem of comparable magnitude, exploitation of the non-crystallographic symmetry<sup>11</sup> has been important for obtaining a good map.

## Data collection and phase determination

TBSV crystals, obtained as described before<sup>7</sup>, are cubic, space group 123,  $a = 383$  Å. Each lattice point of the body-centred cubic cell is occupied by one virus particle, such that the crystallographic point group is the tetrahedral subgroup of the viral icosahedral symmetry.

X-ray intensities were measured by oscillation photography from crystals of the native virus and of two derivatives,  $\text{PtCl}_6^{2-}$  and  $\text{Hg}^{2+}$  (Table 1). Coordinates for the  $\text{PtCl}_6^{2-}$  derivative were available from earlier work, refined in projection to 9 Å (ref. 7). A difference-Fourier  $hk0$ -projection for the  $\text{Hg}^{2+}$  derivative at 9 Å, (from a  $6^\circ$  precession photograph), using signs from  $\text{PtCl}_6^{2-}$ , revealed two sets of quasi-equivalent sites (six sites per icosahedral asymmetric unit). One of these corresponds to the  $\text{CH}_3\text{Hg}^+$  sites used in the 16-Å work<sup>7</sup>, and the other to positions near the subunit tips. After conventional refinement, the average figure-of-merit was 0.57 (Table 2). Fivefold molecular averaging was then carried out using recently developed methods<sup>11</sup>. After one cycle of phase refinement the mean figure-of-merit was 0.76 (Table 2). Previous

experience suggested that additional cycles were unlikely to improve the map further<sup>7,10</sup>. A second map was computed using the combined phases, icosahedrally averaged<sup>11</sup>.

## Electron density map

It is convenient to look at the electron density map in sections normal to a crystallographic twofold (arbitrarily chosen as the Z-axis). The 16-Å resolution map<sup>7</sup> shows that the quasi-threefold axis is approximately parallel to the nearest strict twofold (Fig. 1), so that quasi-threefold related electron density features appear in almost the same Z-sections. The triangle shown in Fig. 1 is thus a useful choice for the icosahedral asymmetric unit, and we chose to contour on acetate sheets a region somewhat larger than this triangle, indicated by the rectangular outline.

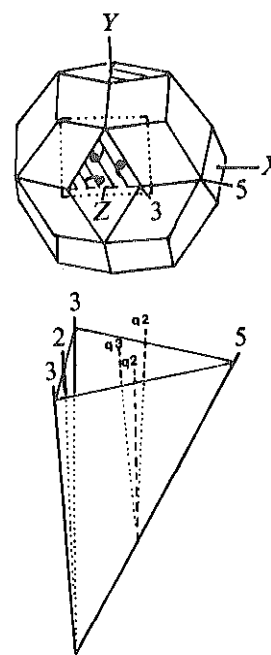


Fig. 1 Orientation of TBSV symmetry axes and choice of icosahedral asymmetric unit. *a*, Rhombic triacontahedron, showing some icosahedral symmetry axes. This figure has icosahedral symmetry, and it is useful for representing a  $T = 3$  icosahedral surface lattice. The triangular half-face (shaded) is an asymmetric unit of the surface, that is, this region, replicated by the various symmetry operations, will cover the entire surface. In a  $T = 3$  lattice, there is a local threefold axis near the centre of this triangle, relating the three quasi-equivalent subunits in the icosahedral asymmetric unit. Note that X, Y and Z are twofold axes. The rectangle indicates the area of the electron-density map, computed in sections normal to Z, chosen for contouring (compare Fig. 2). *b*, An asymmetric unit of volume in a region governed by icosahedral symmetry. This volume, replicated by the various symmetry operations, will fill the rhombic triacontahedron in (*a*); it can be thought of as the domain swept out by the upper shaded triangular half-face in (*a*) in a series of concentric rhombic triacontahedra of continuously increasing radius. Orientations of local (quasi) symmetry axes (denoted by small numbers) are also shown. The quasi-twofold and threefold axes around each fivefold axis intersect it, about 60 Å out from the particle centre.

\*To whom correspondence should be addressed.

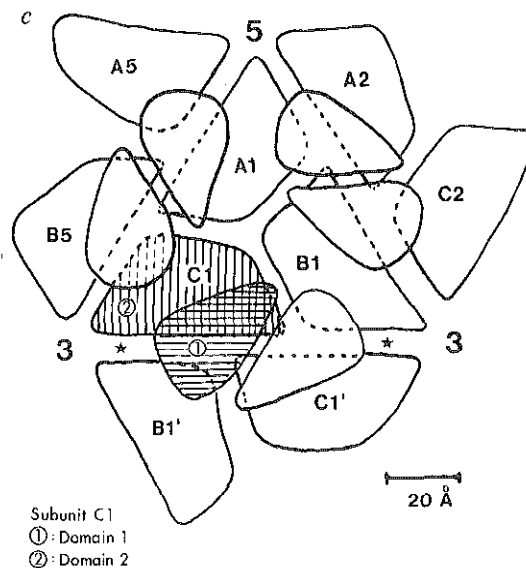
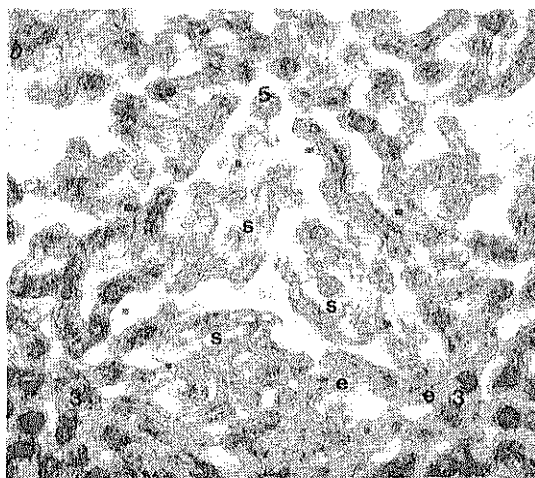
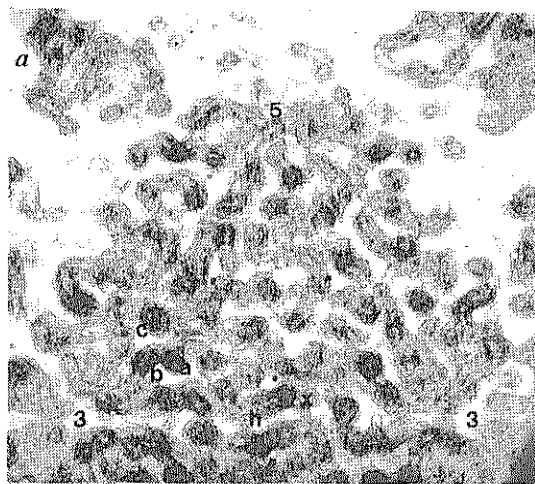


Fig. 2 Views along Z through sections of the TBSV 5.5-Å electron density map. *a*,  $Z = 125\text{--}140\text{ \AA}$ ; *b*,  $Z = 110\text{--}125\text{ \AA}$ . Figure 1 gives orientation. Some features of secondary structure are indicated by lower-case letters: *X* is the inter-domain connection; *h*, the tangential rod; *a*, *b* and *c*, some radial rods; *s*, the saddle-like feature; and *e*, the extra density. The subunit boundaries may be judged by reference to (*c*). *c*, Packing of subunits and domains. A1, B1 and C1 represent subunits packed in one icosahedral asymmetric unit, related to each other by quasi-threefold rotation; A2, . . . , A5, are related to A1 by fivefold rotations; similarly, B1, . . . , B5 and C1, . . . , C5; B1' is related to B1 by a strict dyad, A2 to B1 and B5 to A1, by quasi-dyads. The domains are indicated in rough outline: barrel-shaped domains 1 protrude outwards from the surface, making only twofold contacts; the more extensive domains 2 make contacts across all adjacent symmetry axes, forming a continuous polyhedral shell. The shading of C1 indicates which domain 1 is associated with the labelled domain 2. Domains 1 for C2, B1' and A5 are not shown. Stars indicate the interrupted contact ('cleft'), occupied by RNA. The scale is the same for (*a*), (*b*) and (*c*).

Table 1 Statistics of data collection to 5.5 Å

	Native	PtCl <sub>6</sub> <sup>2-</sup>	Hg <sup>2+</sup>
Total fully recorded reflections measured	69,000	75,000	65,000
No. of independent reflections measured	26,700	27,100	26,700
$R_{\text{sym}} = \frac{\sum  I_i - \bar{I} }{\sum I_i}$	0.100	0.133	0.145
Mean isomorphous difference $= \frac{ I_{\text{nat}} - I_{\text{der}} }{I_{\text{der}}}$	—	0.29	0.56

Data were collected by oscillation about [101], the shortest axis in reciprocal space; rotation from 0° to 20° (where 0° corresponds to [101] coincident with the X-ray beam) includes more than 90% of the 29,000 independent reflections to 5.5 Å. The crystals thus always present a rather plate-like aspect to the X-ray beam, which is much smaller than the crystal diameter, so absorption effects are minimised. The oscillation range of a single 20–24-h exposure was 55–65° of arc. Exposure usually exhausted the crystal lifetime, so that we could not add partially recorded reflections from abutting photographs<sup>12</sup>. Partial reflections, about 1/3 of those recorded, were not included in the data set but the ratios of the measured intensities of some of them to the intensities of corresponding symmetry-related whole spots on the same film, were used to obtain an accurate estimate of crystal orientation and hence an accurate discrimination between fully and partially recorded reflections<sup>9</sup>. Because of the small cross-fire in the focused beam and the small mosaic spread of the crystals, a 10–15' overlap of adjacent exposures ensured that all reflections were fully recorded on film. Films were densitometered using a PDP-11/20-linked Optronics Photoscan and a modified version of the program SCAN 12 (ref. 13). The output included an integrated intensity and a standard deviation for each reflection, the latter determined from the actual standard deviation of the measured background. The three films in a pack were scaled to each other by the method of Fox and Holmes<sup>14</sup>. The approximately 3,000 fully-recorded reflections were sufficient to generate a Wilson-type plot of  $\langle I \rangle$  against  $\sin^2 \theta / \lambda^2$ , and these plots were used to determine relative scale and temperature factors for each film pack.

The electron-density features within this icosahedral asymmetric unit begin at about  $Z = 112\text{ \AA}$  and extend out to  $Z = 170\text{ \AA}$ . At  $Z \sim 110\text{ \AA}$  there are faint and rather wispy but relatively continuous features. From  $Z = 112\text{--}140\text{ \AA}$ , the volume is entirely filled with dense features, and at larger values of  $Z$ , density is concentrated more and more around the strict and quasi dyads, leaving large gaps in between. The icosahedrally-ordered part of the virus structure can thus be described as a roughly 30-Å thick, compact polyhedral shell, from which 90 protrusions, the dimer-clustered morphological units so clearly seen in electron micrographs, emerge.

### Boundaries and domains

Figure 2*a* shows the composite electron density of sections from the central part of the polyhedral shell. Local threefold symmetry is visible in the central part of the asymmetric unit, and there is striking similarity of corresponding parts of quasi-equivalent subunits. Assignment of subunit boundaries was aided by the requirement that symmetry-related density features near an axis must belong to different subunits, even though they may seem weakly connected. A good illustration of the clarity of most boundaries is the gap running more or less along the lower edge of the icosahedral asymmetric unit (connecting the strict threefolds) in Fig. 2*a*. The quasi-equivalent boundaries, not as pronounced in this view, show up equally well if viewed along the local dyad.

The most outstanding characteristic of the subunit is that it falls into two very distinct, globular domains (Fig. 3). The smaller of these (domain 1) forms half of one of the dimer-clustered protrusions; the larger (domain 2) forms part of the polyhedral shell. Domain 2 has a rather distinct, saddle-like feature just against its innermost surface, which we describe in the next section. The two domains are connected at only one place by a narrow

**Table 2** Phase determination

	Site	Heavy-atom parameters				
		Z	B	x	y	z
PtCl <sub>6</sub> <sup>2-</sup>	A	55	53	-0.0157	0.1493	0.3060
	B	49	46	0.0756	0.0481	0.3014
	C	53	48	-0.0619	0.0203	0.3001
	D	19	35	-0.1132	0.0059	0.2933
<i>R<sub>c</sub></i> = 0.58, <i>R</i> = 0.56, r.m.s. <i>f<sub>c</sub></i> /r.m.s. <i>E</i> = 1.86, mean $\Delta F$ /mean <i>F</i> = 0.12						
Hg <sup>2+</sup>	A1	80	25	-0.0372	0.0905	0.3339
	B1	71	27	0.0358	0.0923	0.3344
	C1	81	35	0.0010	0.0271	0.3360
	A2	71	25	-0.0680	0.1205	0.4234
	B2	66	28	0.0822	0.0976	0.4277
	C2	76	25	-0.0146	0.0145	0.4301
<i>R<sub>c</sub></i> = 0.55, <i>R</i> = 0.53, r.m.s. <i>f<sub>c</sub></i> /r.m.s. <i>E</i> = 1.9, mean $\Delta F$ /mean <i>F</i> = 0.26						

Z, occupancy in electrons, on experimental absolute scale<sup>7</sup>; B, heavy-atom 'temperature factor'; R,  $\Sigma||\Delta F_{\text{obs}}| - |\Delta F_{\text{calc}}|| / \Sigma|\Delta F_{\text{obs}}|$ , all reflections; *R<sub>c</sub>* (Cullis *R* factor), *R* for centric reflections; *E*, lack of closure at best phase; *f<sub>c</sub>*, heavy-atom structure factor; r.m.s., root-mean square.

Statistics of molecular averaging up phase refinement.

Figures of merit: initial (DIR), 0.57; final (combined), 0.76.

Comparison of calculated and observed structure factors:

$$R = \langle ||F_{\text{obs}}| - |F_{\text{calc}}|| \rangle / \langle |F_{\text{obs}}| \rangle = 0.37.$$

$$\text{Mean phase difference} = \langle |\alpha_{\text{calc}} - \alpha_{\text{DIR}}| \rangle = 47^\circ.$$

$$\text{Mean phase change after combination} = \langle |\alpha_{\text{final}} - \alpha_{\text{DIR}}| \rangle = 25^\circ.$$

Alternate rounds of phasing and icosahedrally-constrained least-squares refinement<sup>7,15</sup> used starting coordinates obtained as described in the text. Coordinates and occupancies for both derivatives were refined in projection, first at 9 Å then at 5.5 Å, and finally in three dimensions against a set of approximately 3,500 reflections representing a random sample of resolution and intensity ranges, selected with the criterion that the precision of heavy-atom intensity differences meet a pre-determined minimum. After refinement in projection, self and cross-difference maps gave no evidence for additional sites, a conclusion confirmed in three dimensions. Site names correspond to the subunit to which they were subsequently seen to bind (compare Fig. 2c), and only one icosahedral asymmetric unit of sites is given. The quoted values of *R<sub>c</sub>* and so on, refer to the 3,500 well-measured reflections used for the refinement. For icosahedral phase refinement, an initial map, computed on a 1.1-Å grid for accurate interpolation<sup>11</sup>, was averaged about the icosahedral fivefold axes to determine the molecular envelope; density not part of the particle to which these axes refer is smeared out by this procedure, enhancing the boundary contrast. The final envelope for an entire virus particle is a rhombic triacontahedron (Fig. 1) with roughly cylindrical 'pillars' at each twofold and quasi-twofold position, where subunit tips protrude<sup>7</sup>. Averaging and setting all density outside this envelope to zero yielded a symmetrised map (2.2-Å grid), from which structure factors were computed. The agreement between observed and calculated amplitudes is tabulated: most of this residual comes from the scatter of the figure-of-merit around its mean value in resolution shells<sup>11</sup>. Calculated and DIR phases were combined using the Sim formula<sup>16</sup>.

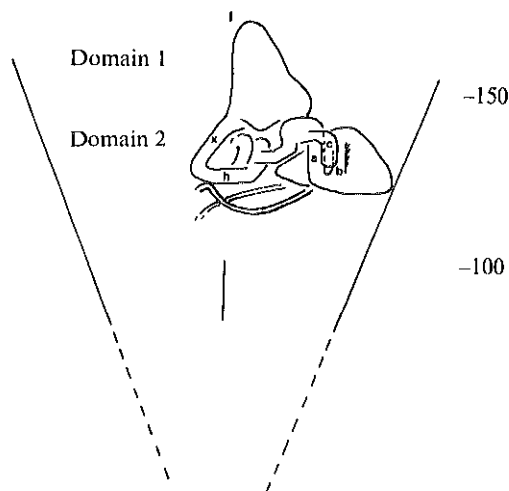
region of strong electron density: each domain thus appears to contain one terminus of the polypeptide chain.

Domain 1 is about 30 × 30 × 25 Å in overall dimensions. It consists of two approximately parallel layers, oriented such that one layer forms an extended and tight dimer contact with domain 1 of the twofold (or quasi-twofold) related subunit. The general distribution of electron density in this domain resembles very strongly the β-barrel structures, also found dimer-clustered, in immunoglobulin fragments<sup>17,18</sup> and in superoxide dismutase<sup>19</sup>. The outer layer of the barrel is markedly curved; the inner appears quite flat, especially in the contact region. The strand directions in each layer cannot be established with certainty in a 5.5-Å map, but there is definite indication of a radial rather than tangential course. The outer layer of the barrel is markedly curved; the inner seems quite flat, especially in the contact region. The strand directions in described below.

The connection of domains 1 and 2 is well defined and occurs at the point marked with a cross in Figs 2a and 3. Domain 2 itself is considerably more complex than domain 1, both in its internal structure and in its interactions with adjacent subunits. Nevertheless, its boundaries appear rather well determined, with one major ambiguity (see below). In tracing a boundary, we used the symmetry criteria described earlier and assumed that the domain is a reasonably compact globular unit. Figure 2c shows schematically the relation of domains 1 and 2 and the nearest-neighbour packing in the protein shell.

Domain 2 has a number of rodlike features indicative of α-helices (these are absent in domain 1). These rods are oriented more or less radially with one important exception, which we refer to as the tangential rod (h, in Fig. 2a). As seen in Figs 2a and 3, one end of this rod connects with the feature that forms the junction between the two domains. The tangential rod is 15 Å long (2-3 turns of α-helix) and located close to the dyad, such that the distance between twofold-related helical axes is 10 Å. Each rod is tilted slightly out of the plane normal to the dyad, and we have chosen the enantiomorph such that the tilted rods have a left-handed relationship. This packing is appropriate for right-handed α-helices<sup>10,21</sup>. Closely-packed, radially-oriented rods in domain 2 show the same relative tilt, and the twist of the outer wall in the

domain 1 β-barrel is also consistent with this choice. The three most prominent radially-directed rods (a,b,c in Fig. 2a) seem to be in direct contact. Their lengths correspond to 2-3 turns. Rods a and b are tightly packed, whereas rod c is somewhat farther away. The principal ambiguity in boundary tracing, mentioned above, concerns the assignment of c and some connected density to subunit C1 (Fig. 2c). The alternative assignment of c to A1 leads to



**Fig. 3** Schematic representation of the electron density in one subunit, viewed normal to the adjacent dyad and looking toward that dyad from the quasi-threefold. If the drawing is taken to represent C1 in Fig. 2c, then the direction of view is downwards in the plane of that figure. The lobe at the outside of the subunit (*Z* = 140-170 Å) is the β-barrel of domain 1; the cross indicates its connection to the more extensively α-helical domain 2; the 'saddle' lies beneath domain 2, and a tentative interpretation in terms of two stretches of RNA backbone is shown in the diagram.

a more extended domain 2, but the choice is arbitrary. In the pentamer (or hexamer) contact regions near the bottom of domain 2, the density has a layer-like character, suggestive of  $\beta$ -sheet. Rods a and b are packed partly against this sheet-like feature. There are a number of shorter pieces of rodlike density, less clearly  $\alpha$ -helices. One of these occurs close to the local threefold axis: the three such symmetry-related rods appear to form a sort of rudimentary, left-handed, three-stranded coiled-coil<sup>21</sup>.

We estimated the volume occupied by all the strong features of one subunit as  $51,000 \pm 6,000 \text{ \AA}^3$ , where the upper and lower limits correspond to generous and restricted estimates of surrounding side chain, Fourier cutoff and so on. If all this volume were protein, the corresponding MW would be  $39,000 \pm 5,000$  (taking  $\bar{v} = 0.74 \text{ cm}^3 \text{ g}^{-1}$ ). Twelve nucleotides, our upper estimate for RNA in the saddle feature, occupy about  $3,400 \text{ \AA}^3$  (taking  $\bar{v} = 0.51 \text{ cm}^3 \text{ g}^{-1}$ ). We can thus account for most of the protein subunit, although some 10% of the chain could be outside strong density in the present map.

### Nucleic acid

A significant problem in interpreting the TBSV electron-density map has been to establish criteria for recognising RNA. Strength of the density alone proved insufficient: the features described below are among the strongest in the map but no stronger than the  $\alpha$ -helices. This is not surprising because partial occupancy of polynucleotide binding sites and other sorts of static disorder would diminish the corresponding density. A general picture of the nucleic acid might be one of ordered regions of backbone, bound to protein, interspersed with more disordered regions at smaller radii. The overall appearance of the map at about  $110 \text{ \AA}$  from the particle centre is consistent with this view: strong features trail off towards the centre in little streamers of density. The distribution of this streaky density suggests that the polynucleotide chain may enter or leave a binding region in more than one way, but the strength of the beaded features to which they join indicates a definite conformation in the binding site.

A striking aspect of the map is the complicated and cross-connected texture of features between  $110$  and  $120 \text{ \AA}$ , in contrast to the rather clear sheet-like or rod-like density of domains 1 and 2. In this region, each icosahedral asymmetric unit shows three very similar, though not precisely identical, saddle-shaped features, about  $30 \times 20 \times 10 \text{ \AA}$  in extent (*s* in Fig. 2*b*). Each of these saddles forms the innermost density associated with one protein subunit. In addition, there is a set of strong features with no quasi-equivalent counterpart extending from near the strict twofold toward the strict threefold axes; we refer to this as extra density (*e* in Fig. 2*b*). Since the protein subunits are identical and since RNA is the only other significant component, this extra density is likely to represent part of the nucleic acid chain. Inspection of a  $5.5\text{-\AA}$  resolution map of yeast-phenylalanyl tRNA (computed as an  $F_{\text{calc}}$  map by A. Jack, MRC Laboratory of Molecular Biology, Cambridge) shows that the sugar phosphate backbone has a quite characteristic appearance: a continuous stretch of density with barely resolved maxima spaced at distances of  $6\text{--}7 \text{ \AA}$ , at or near the positions of phosphates, giving it a beaded aspect. Both the extra density and certain parts of the edges of the saddles have just this  $6\text{--}7\text{-\AA}$  beading. Moreover, the saddle lying beneath subunit C is connected to the extra density at several points, with the  $6\text{--}7\text{-\AA}$  periodicity continuing from one to the other. We therefore conclude that a significant part of the density between  $110$  and  $120 \text{ \AA}$  represents polynucleotide chain.

The saddle features form rather distinct lobes of density just against the bottom of domain 2. We followed possible RNA backbone configurations in this region, although, in view of the ambiguities involved, such a tracing is rather speculative at this resolution. We believe that we can recognise two pieces of RNA backbone (Fig. 3). One of these emerges from the particle interior and passes near the strict or quasi-twofold; the second forms the outer rim of the saddle. We detected 10–12 possible phosphates. Density between these chains could in part be nucleotide bases, and the backbones are so oriented that they could be joined by two or three base pairs in the A-RNA geometry.

No simple description would seem to summarise protein-RNA contacts. The structure of domain 2 shows that several portions of the polypeptide chain, separated in the amino acid sequence, must come together to form the shallow pocket in which the saddle feature lies. Stretches of  $\alpha$ -helix, as well as a region that may be  $\beta$ -sheet, are involved.

### States of the protein subunit

The clarity of the electron density map has facilitated quantitative analysis of subunit similarities. The coat protein assumes two clearly different conformational states, depending on whether it is near a local or a true dyad in the  $T = 3$  surface lattice. The individual domains 1 and 2 are essentially invariant, and a hinge between the rigid domains relates the two subunit conformations.

Corresponding well defined electron density maxima could easily be recognised in all three quasi-equivalent subunits and their superposition determined after appropriate rotation and translation. Referring to Fig. 2*c*, subunits A and B superpose very well, whereas A (or B) and C do not. Taken separately, however, domains 1 and 2 of A (or B) can be superposed with their analogues in C as well as A and B can be matched to each other. This implies that the relative orientation of domains 1 and 2 in C is different from their relative orientation in A or in B, but that individual domains are nearly invariant. The hinge between domains is

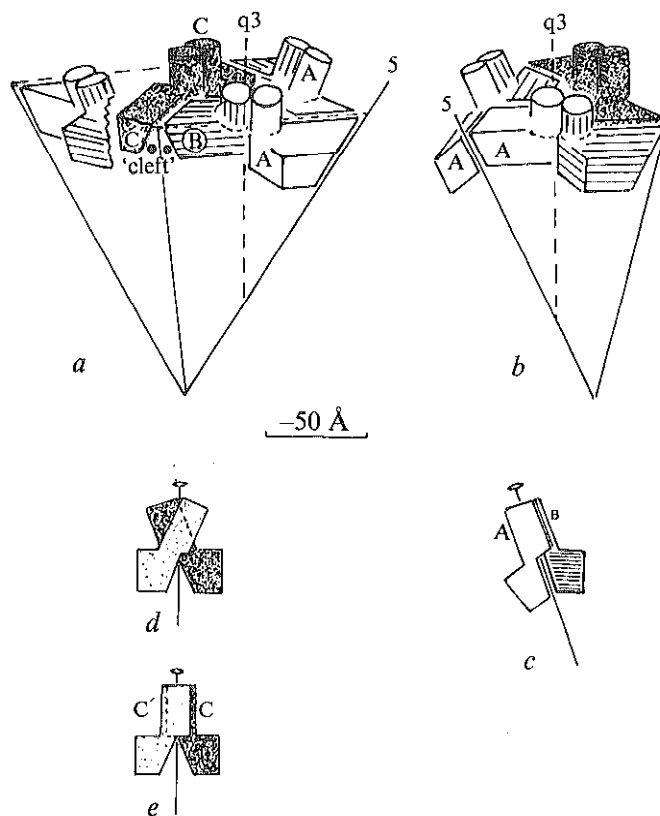


Fig. 4. Packing diagrams, showing subunits and relative orientation of the two domains. *a*, View similar to Fig. 1*b*, with a second icosahedral asymmetric unit replicated by the strict twofold axis. Each subunit is represented by a cylinder (domain 1) on a block-like base (domain 2). The three quasi-equivalent subunits are shown: white (A), hatched (B) and stippled (C). *b*, View of an icosahedral asymmetric unit, rotated  $120^\circ$  with respect to the view in (*a*). Considering just the domain 2 contacts, note that the A-A contact about the fivefold axis and the B-A contact across the quasi-dyad are quite tight, whereas the B-C' and C-C' contacts about the quasi-sixfold and strict twofold axes are opened up to produce a cleft. The difference in the angle of contact across quasi (A-B) and strict (C-C') dyads requires a change in the angle between domains 1 and 2 to preserve the dimer pairing of domains 1. This is shown in (*c-e*): the A-B contact (quasi-dyad) is depicted in (*c*). If the contact between the domains 2 were altered, without a relative reorientation of domains 1 and 2 in each subunit, the C-C' contact (strict dyad) would appear as in (*d*). In fact, there is rotation by approximately  $20^\circ$  about a hinge between domains 1 and 2, producing the sort of contact shown in (*e*).

located just above the end of the horizontal  $\alpha$ -helix (Fig. 3). The hinge rotation axis is approximately parallel to the axis of the tangential rod, and the angle of rotation is about  $20^\circ$  (for C, relative to A or B). There are small, though significant, differences between A and B, but the hinge position is clearly the same for both.

The inter-subunit contacts of the hinged protein are represented in Figs 2c and 4, which show how the hinge leads to quasi-equivalent subunit bonding. Domains 1 have large and essentially identical contacts across both local and strict twofold axes. Domains 2 are likewise in extensive contact around the local threefold axis, which, for these domains, approximates well to an exact triad. Schematically, the polyhedral shell formed by all the domains 2 can be represented by 60 wedge-like trimers in continuous contact across local dyads but with only their outer edges in contact across strict dyads (Fig. 4). This produces a cleft connecting strict threefold (quasi-sixfold) axes. The quasi-sixfold relationship is thus very inexact: domains 2 of C1 and B5 in Fig. 2c are related essentially as A1 and A2, whereas the analogous contact between B1' and C1 (star in Fig. 2c) is largely broken by the cleft. This is in fact the only place where contacts are not basically the same for all three quasi-equivalent subunits. It is the hinge between domains 2 and 1 of subunit C that permits the cleft to open (Fig. 4e); without it, TBSV protein could form only a simple  $T = 1$  icosahedral shell. Since 'extra' electron density features appear in the cleft, but not in the tight contacts between C1 and B5 or A1 and A2, RNA may play a significant role in regulating assembly of correct-sized,  $T = 3$  particles. RNA-free,  $T = 1$  structures have indeed been observed in studies of reassembly of the related turnip crinkle virus<sup>2,3</sup>.

## Domains and hinges

The generalisation that large polypeptide chains will fold into smaller, compact domains has gained considerable strength from the several proteins solved thus far with subunit molecular weights greater than 30,000 (refs 24–26). The two domains of TBSV protein have approximate molecular weights of 15,000 and 25,000; in the scheme of Levitt and Chothia<sup>27</sup> they seem to be members of class II (all  $\beta$ ) and III ( $\alpha + \beta$ ), respectively.

The TBSV subunit is not the first example of a protein with flexibly-linked domains. The immunoglobulins are the classic case of such a structure<sup>17</sup>, and the relation between light chains in the Mcg Bence-Jones dimer<sup>18</sup> is precisely analogous to the situation in TBSV: one of the chains in this dimer (the chain with a

'heavy-chain-like' configuration) bends at a hinge point, in order to conserve strong inter-molecular contacts. This type of structure could lead to controlled polymorphism in assembly.

The presence of RNA in the cleft at the C1/B1' interface seems to dictate the nature of inter-molecular contact and hence the choice of a  $T = 1$  or  $T = 3$  design. Since the angle of the contact is determined by the hinge angle, binding of ligands to residues in the hinge region, or at the domain 1/domain 2 interface, could likewise influence the mode of association. In general, a subunit with flexibly-connected rigid domains can be a useful element for controlled switching between distinct states of a macro-molecular assembly. Work on a 2.8-Å resolution map of TBSV is underway.

We thank James Crawford for the SCAN 12 oscillation film scanning program, A. Jack for programs and a tRNA map, D. L. D. Caspar for advice, W. N. Lipscomb for use of an Optronics scanner and P.D.P. 11/20, Howard Bailey for technical assistance, Athena Andriadis for aid in data processing and the Harvard Chemistry Department for computing facilities including a remote terminal to the Columbia University IBM 360/91. The work was supported by a grant from the NCI. S.C.H. acknowledges a USPHS Research Career Development Award and F.K.W. a fellowship from the Swiss National Fund.

Received October 5; accepted December 10, 1976.

- <sup>1</sup> Caspar, D. L. D., and Klug, A., *Cold Spring Harb. Symp. quant. Biol.*, **27**, 1–24 (1961).
- <sup>2</sup> Weber, R., Rosenbusch, J., and Harrison, S. C., *Virology*, **41**, 763–765 (1970).
- <sup>3</sup> Dorne, B., and Pinck, L., *FEBS Lett.*, **12**, 241–243 (1971).
- <sup>4</sup> Ziegler, A., Harrison, S. C. and Leberman, R., *Virology*, **59**, 509–515 (1974).
- <sup>5</sup> Finch, J. T., Klug, A., and Leberman, R., *J. molec. Biol.*, **50**, 215–222 (1970).
- <sup>6</sup> Jack, A., Harrison, S. C., and Crowther, R. A., *J. molec. Biol.*, **97**, 163–172 (1975).
- <sup>7</sup> Harrison, S. C., and Jack, A., *J. molec. Biol.*, **97**, 173–191 (1975).
- <sup>8</sup> Harrison, S. C., *J. appl. Crystallogr.*, **1**, 84–90 (1968).
- <sup>9</sup> Schutt, C. E., and Winkler, F. K., in *The Rotation Method in Crystallography* (edit. by Arndt, U. W., and Wonacott, A. J.) (North-Holland, Amsterdam, 1976).
- <sup>10</sup> Champness, J. N., Bloomer, A. C., Bricogne, G., Butler, P. J. G., and Klug, A., *Nature*, **259**, 20 (1976).
- <sup>11</sup> Bricogne, G., *Acta crystallogr.*, **A 32**, 832 (1976).
- <sup>12</sup> Arndt, U. W., Champness, J. N., Phizackerley, R. P., and Wonacott, A. J., *J. appl. Crystallogr.*, **6**, 427–463 (1973).
- <sup>13</sup> Crawford, J., thesis, Harvard Univ. (1977).
- <sup>14</sup> Fox, G. C., and Holmes, K. C., *Acta crystallogr.*, **20**, 886–891 (1966).
- <sup>15</sup> Dickerson, R. E., Weinzierl, J. E., and Palmer, R. A., *Acta crystallogr.*, **B24**, 997–1003 (1967).
- <sup>16</sup> Sim, G. A., *Acta crystallogr.*, **12**, 813–815 (1959).
- <sup>17</sup> Colman, P. M., Deisenhofer, J., Huber, R., and Palm, W., *J. molec. Biol.*, **100**, 257–282 (1976).
- <sup>18</sup> Schiffer, M., Girling, R. L., Ely, K. R., and Edmundson, A. B., *Biochemistry*, **12**, 4620–4631 (1973).
- <sup>19</sup> Richardson, J. S., Thomas, K. A., Rubin, B. H., and Richardson, D. C., *Proc. nat. Acad. Sci. U.S.A.*, **72**, 1349–1353 (1975).
- <sup>20</sup> Chothia, C., *J. molec. Biol.*, **75**, 295–302 (1973).
- <sup>21</sup> Crick, F. H. G., *Acta crystallogr.*, **6**, 689–691 (1953).
- <sup>22</sup> Rao, S. T., and Rossmann, M. G., *J. molec. Biol.*, **76**, 241–256 (1973).
- <sup>23</sup> Leberman, R., and Finch, J. T., *J. molec. Biol.*, **50**, 209–213 (1970).
- <sup>24</sup> Poljak, R. J., *et al.*, *Proc. nat. Acad. Sci. U.S.A.*, **70**, 3305–3310 (1973).
- <sup>25</sup> Stoltz, T. A., Fletterick, R. J., Anderson, W. F., and Anderson, C. M., *J. molec. Biol.*, **104**, 197–222 (1976).
- <sup>26</sup> Blake, C. C. F., and Evans, P. R., *J. molec. Biol.*, **84**, 585–601 (1974).
- <sup>27</sup> Levitt, M., and Chothia, C., *Nature*, **261**, 552–558 (1976).

5  
1  
4

5  
3  
4  
2