CHAPTER 7

CONCLUSIONS

The result of two cycles of 12Å phase refinement and four cycles of 8Å refinement, starting with the best full parameter model, was a map, calculated with F^{obs} and derived phases. This was averaged 30-fold and skewed into a rectangular box, somewhat larger than the icosahedral asymmetric unit. It was contoured at six equally spaced electron density levels between zero and the most positive peak. A map of the compact structure was prepared similarly, using the high resolution refined phases and structure factors between 35 and 8 Ångstroms resolution; a temperature factor was applied to these data to simulate the loss of detail at high resolution due to thermal disorder, the value of $B = 340Å^2$ being derived from the Wilson plot slopes in figure 3.6. On both maps the pseudo atom positions (one atom per amino acid residue) of the respective models were superimposed using the DRWMOL program (J. E. Ladner). These were colour coded by domain to assist the examination of the domain contacts.

7.1 Interpretation of the Maps.

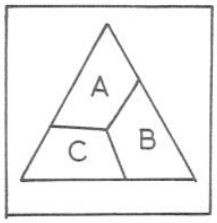
Comparison of the shape of the electron density for corresponding parts of the model showed strikingly good

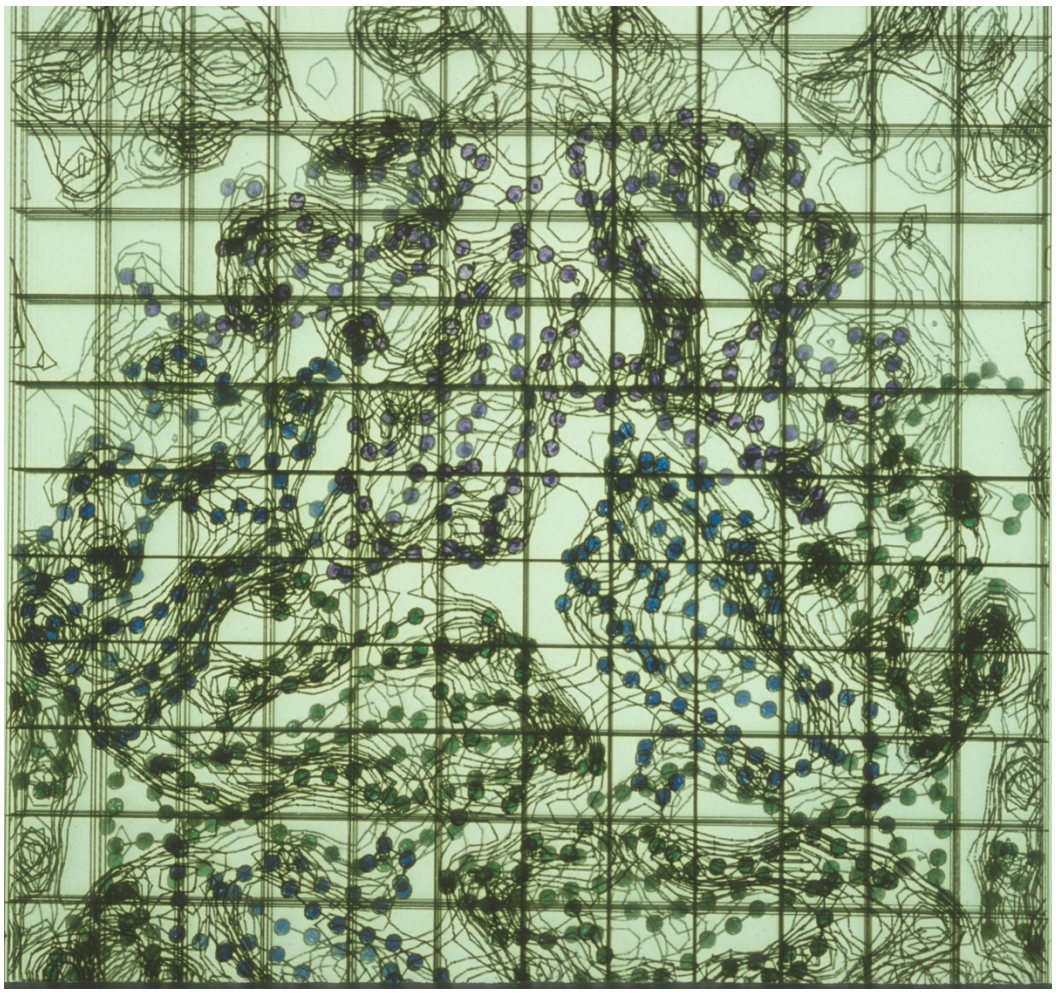
agreement between the two maps. Features with weaker or stronger density than average corresponded well, and the solvent regions outside the subunit boundaries, with certain exceptions mentioned below, had very low levels of noise. This last statement includes the large cleft which forms in the region of the trimer contacts in the compact structure. This cleft is over 80Å long and is 20Å wide in the middle. It is clearly visible in plate 7.1. A detailed interpretation of all the differences of structural significance follows.

The outside of the particle is referred to as 'up' and the inside as 'down', corresponding to the z sectioning of the maps (see figure 1.1). The sequence information, which is incomplete, but only considered in regions where it is accurate, comes from the laboratory of R. Sauer (personal communication). The amino acid numbering scheme is that of Harrison et. al. (1978), offset by 66 residues, which is the number that was disordered at the N-terminus in the high resolution compact structure. All distances that are quoted are for the separation of the pseudo atoms that represent the amino acid centroids.

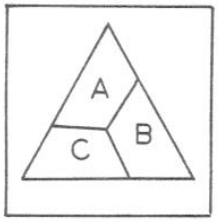
7.1.1 Beta Annulus.

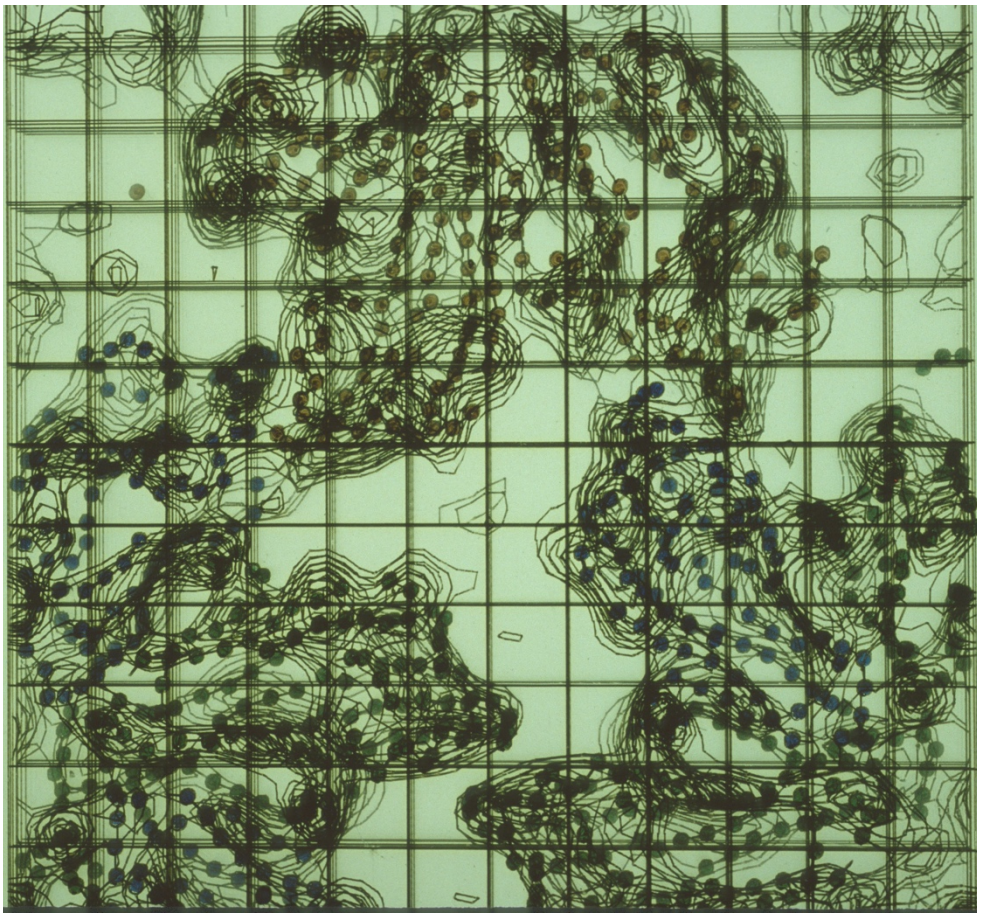
This feature is one of the strongest of the expanded structure. At 8 Ångstroms resolution, the compact structure shows positive electron density only in the





(a) 8 Ångstrom map of the compact structure with the pseudo atom model superimposed (one circle per amino acid residue). The map was calculated with high resolution refined phases, structure factors between 35 and 8 Ångstroms, a temperature factor of 340Å^2, and 5-fold symmetry averaging. Sections from z = 116 to 128Å are included.





(b) 8 Ångstrom map of the expanded structure with the pseudo atom model superimposed. Sections from z = 134 to 146Å are included.

Plate 7.1. General views of the electron density maps for compact and expanded TBSV. 7 adjacent sections spaced 2Å apart are superimposed, so that most of the S-domain regions are shown. The grid is 10Å square. The x axis runs from left to right across the page, the y axis up the page, and the z axis out of the page. These conventions are used throughout the plates and figures of this chapter.

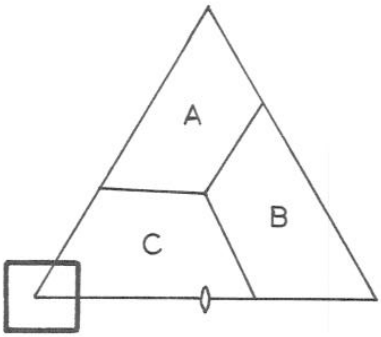
regions of the three 3-stranded β sheets. The expanded one has these as well as strong density at the top of the domain where the strands pass very close to the 3-fold axis (in the region of residue 82). At the bottom there is a significant continuation of the density which extends about 8Å into the interior of the particle (plate 7.2), below residue Ile 67 which is the first ordered residue of the high resolution structure. The estimated extent of this extra density is enough to accommodate about 4 residues in an alpha helix or 3-4 rather large residues in an extended conformation. The sequence immediately preceding Ile 67 is:

Arg Arg Lys Gln Gly Asn Gln Gln Ile
60 67

The conformation of the three consecutive basic residues is unlikely to be the same in the 30 symmetry equivalent positions because of its strong interaction with the phosphates of the RNA, which does not have the icosahedral symmetry (Harrison, et. al., 1978; section 7.1.8, below). It is therefore reasonable that these should be the first disordered residues below the beta annulus of the expanded structure.

7.1.2 The Extended Arm.

The 'arm' is the feature of the compact structure for which the quasi equivalence of the three positions breaks down; it appears that the same is true for the expanded structure. The model of the C position arm based



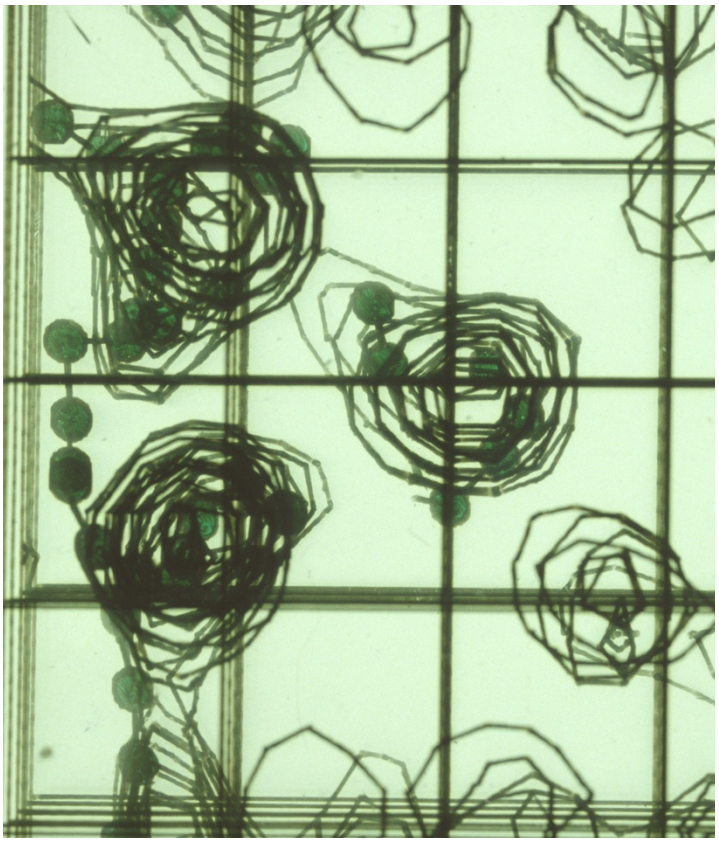
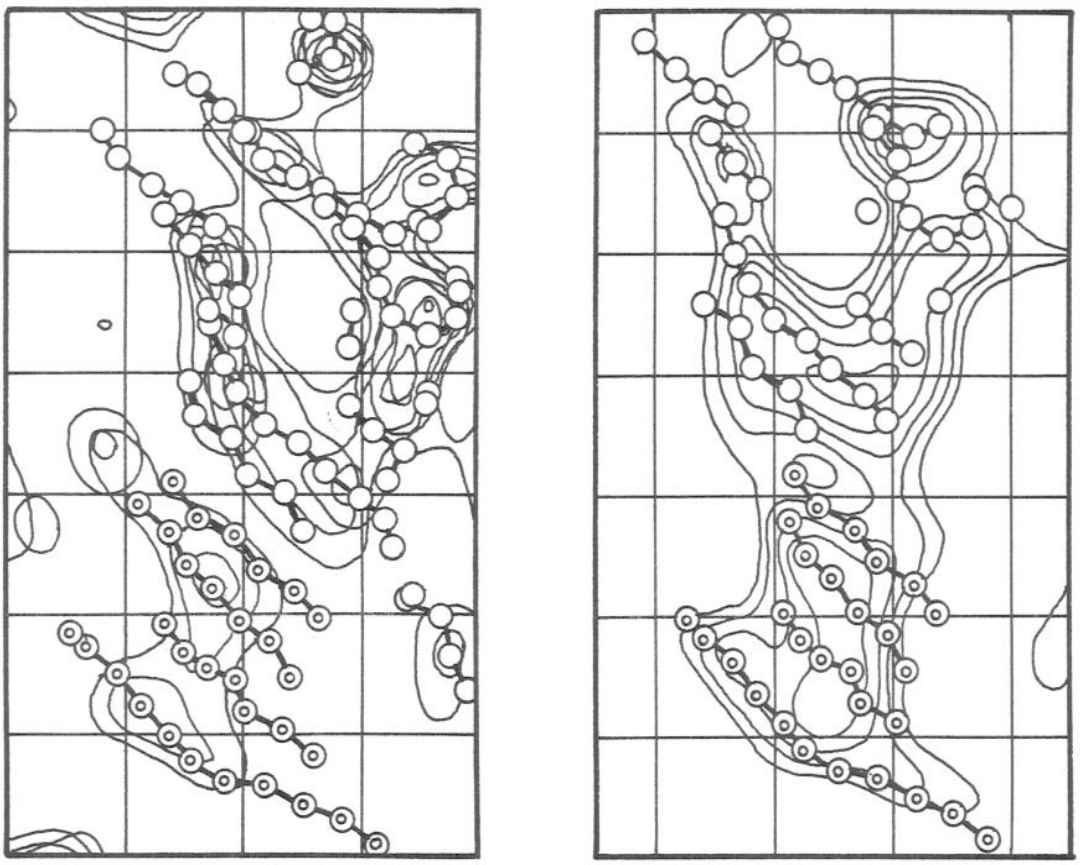


Plate 7.2. The underside of the beta annulus viewed from the interior of the expanded particle. Sections from z = 118 to 130Å are included.

on the position of its S-domain appears to be correct: there is strong density over its whole length, very similar to that of the compact structure. One interesting difference is that the side chain of Arg 98 is clearly visible in the compact 8 Ångstrom map and not ordered in the expanded one, whereas Lys 92 and Lys 94 are ordered in the latter (as a single projection hanging below the arm) but not the former.

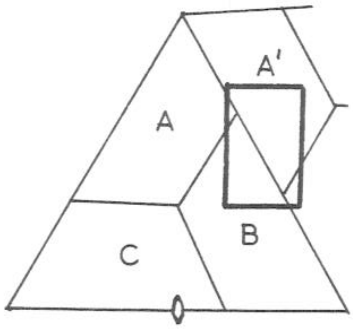
In the A and B position subunits, the first residue visible in the high resolution compact structure is Thr 99, which is 1 or 2 residues before the first strand of the horizontal β sheet across the base of the S-domain (see figure 1.3). Major breaking of the quasi symmetry arises here because this β strand is hydrogen bonded to the arm in the C position subunits ('divided' contact), but is hydrogen bonded to its quasi equivalent mate in the A and B positions ('direct' contact). The bonding extends over at least 6 residues, so that the sheet is seen to continue over both domains. In the expanded structure, this sheet is torn because the A and B position S-domains slide about 14Å past each other; this is just enough to break all continuity of the sheet, with no overlap of these strands (figure 7.1). However, examination of the electron density shows there to be a distinct bulge beyond the first strand of both the A and B position S-domains (plate 7.3), just where the next strand in the sheet would lie, which can be interpreted in three ways:



○ A position

◎ B position

Figure 7.1. Sections of the electron density at the base of the S-domain AB contacts in the compact (left, z = 114Å and 116Å superimposed) and expanded (right, z = 132Å) structures. The relevant regions of the model (one circle per amino acid residue) are included. The shearing of the β sheet contact is clearly apparent. A photograph of a greater number of sections in this region is shown in plate 7.3.



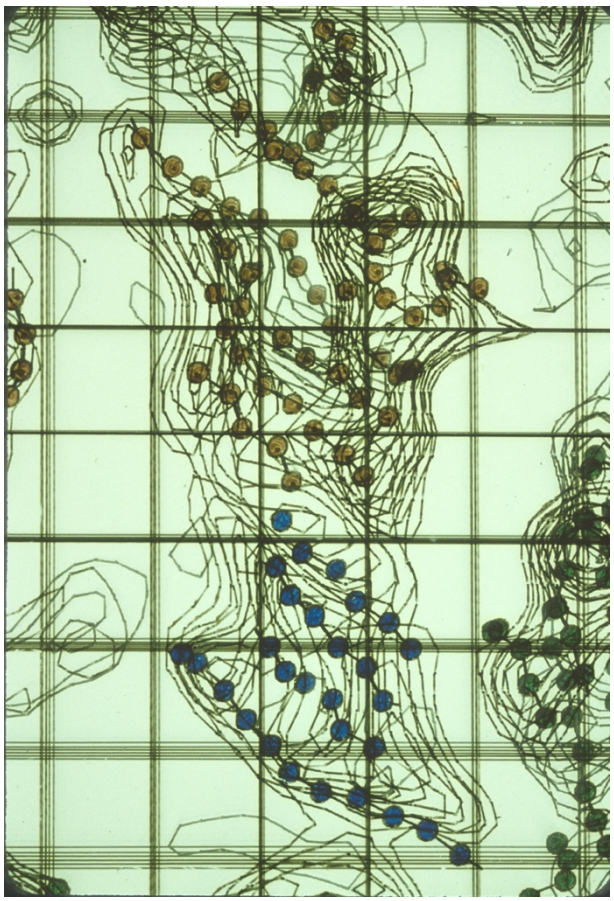


Plate 7.3. The base of the S-domain AB contact region of the expanded particle. Sections from z = 126 to 136Å are included. For interpretation, see figures 7.1 and 7.2.

i) as the side chains of that first strand. This is unlikely, since the strand is known to be in the β pleated configuration in which the side chains point up and down, not sideways.

ii) as the extended arm folded back into the same configuration as in the C position subunit. This is shown in figure 7.2(a).

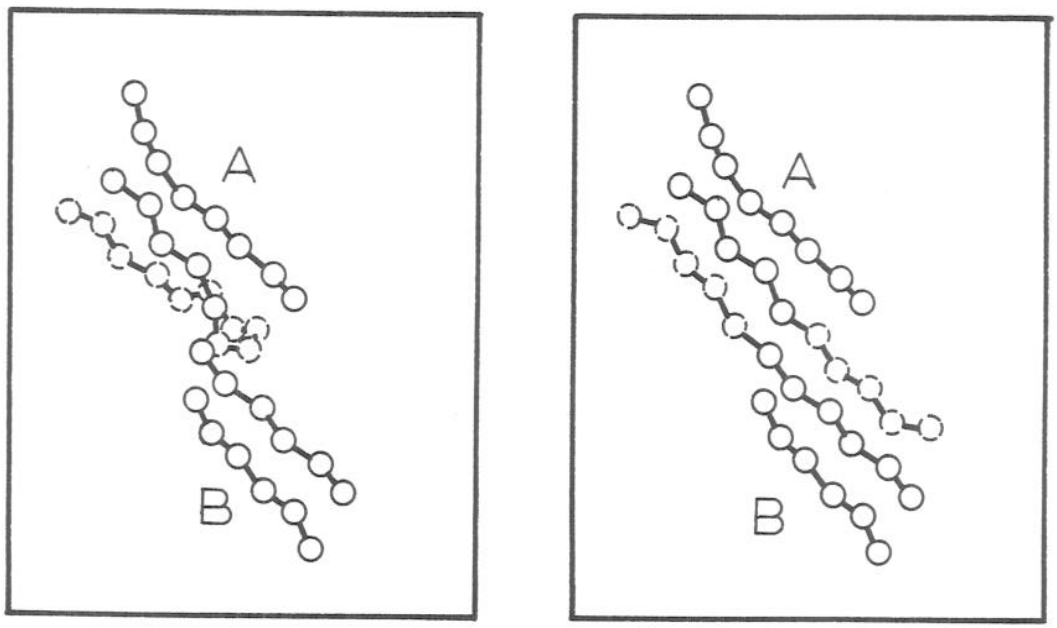
iii) as the extended arm continuing straight across the AB interface to form the same contacts as in the C position subunit, but with the opposing domain. This is shown in figure 7.2(b).

On closer examination, the second possibility is disallowed too, because the beta turn between the strand in question and the arm itself takes up so much space that the A and B position versions would collide. Besides, if it were true, a large protrusion of density underneath the domain boundary would be expected, analogous to that seen for the C position, but is not observed.

Instead, the last possibility would fit the electron density extremely well (compare figure 7.2(b) with plate 7.3). The sequence of the new β strand is:

Gly	Arg	Thr	Ser	Gly	Ser	Val	Thr	Val	Thr	Ser	His
97			100					105		?	108

This includes the old strand (103 to 108) and its hypothetical continuation (97 to 102). It is interesting to observe that the alternating pattern of hydrophobic and hydrophilic residues on the right hand side is conserved on the left



(a) Arm folded back (A position arm only).

(b) Arm extended.

Figure 7.2. Interpretation of the interruption of the S-domain AB contact, shown in figure 7.1. See text for explanation.

too.

The functional significance of such a structure is that it clamps the particle into the expanded state, by fixing the relative location and orientation of all the S-domains. Without some specific mechanism such as this, there is no reason that the expanded particle should not have a variable size, and so a completely indefinite structure.

In the case of the B position arm, and to a much lesser extent the A, the density continues up around the side of the opposing S-domain as far as the radial (vertical) α helix. The contact with this helix appears to be analogous to that of the compact structure, which is with the piece of chain immediately preceding the hinge. From the point of termination of this extra density, there is a clear path to the exterior of the virus. Thus there is a suggestion that the arm does indeed loop out of the virus in its expanded state, as is suggested by its sensitivity to proteolytic cleavage there (see chapter 1).

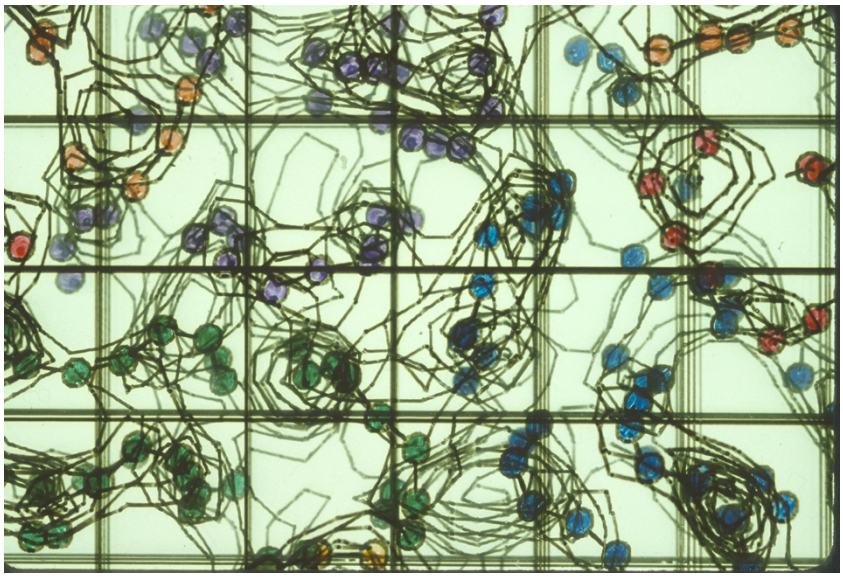
7.1.3 S-Domains.

The fit of the S-domains into the electron density is excellent except for one distinct feature. The trimer contact loops of the compact structure (residues 224 to 230), which run close to the quasi 3-fold axis that approximately relates the A, B and C subunits to each other, are

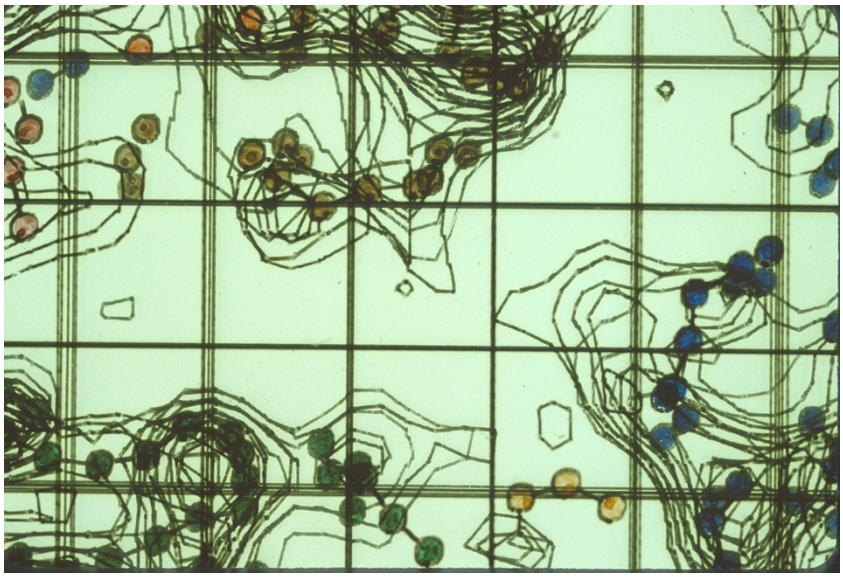
shown in plate 7.4(a). The corresponding region of the expanded structure is shown in plate 7.4(b). The situation for the B position subunit is the most pronounced: a large loop of density, pointing outward into the cleft is visible, unaccounted for by the model, while the loop of the model points upwards and is completely without density. The indication is that the loop adopts a new configuration, displaced downwards by about 10Å at the tip. This is entirely reasonable since the loop makes very little contact with the rest of its own S-domain, instead forming a contact of hydrogen bonds and salt bridges (via one of the Ca^{++} sites) with the other S-domains in the compact structure.

That this feature is not some 'ripple' artefact of the reciprocal space truncation to low resolution is clear from examination of its shape, cycle by cycle, during the course of the 8 Ångstrom phase refinement. After the first cycle, there is density around the model of the loop, which diminishes as the refinement proceeds, and is replaced by the new density.

The situation is more or less the same in the A and C position subunits. In A, the loop folds down only about half as far as B, but seems to have well-defined density. In C, there is very little new density at all, so the loop must be assumed to have become disordered (see also section 7.1.5).



(a) The trimer contact region of the compact virus showing the interaction of the three quasi equivalent trimer contact loops. Sections from z = 130 to 140Å are included.



(b) The corresponding region of expanded TBSV showing the new density that indicates folding down of the trimer contact loops into the cleft. Sections from z = 146 to 154Å are included.

Plate 7.4. Evidence for the conformational change of residues 225 to 230 (the trimer contact loop) during expansion.

7.1.4 The Hinge.

Because the P-domains are in completely new positions, it is necessary for their hinges to adopt new configurations. The connectivity between the S- and P-domains is uncertain: it is possible that this can be determined by using the constraint that the length of the hinge must be less than that of a fully extended polypeptide chain, which may exclude one of the two alternatives. In the case of the AB dimer, it seems likely that the expansion involves a rotation of $+30^{o}$ rather than -150^{o}, but whether the CC dimer rotates $+103^{o}$ or -77^{o} is a completely open question.

In all the three hinges, it was found that the last three residues of the model of the S-domain were out of density. For purposes of dividing the molecule, the S-domain had been (somewhat arbitrarily) terminated at residue 270; it is now clear that residue 267 would have been a better choice as this is the last residue to appear in density.

Residue 273 was the first to be included in the P-domain model. Consideration of the possible paths of the chain between 267 and 273 indicates that the $+30^{o}$ rotation possibility is the only realistic one for the AB dimer, even though there is no density to suggest either one. The situation for the CC dimer is not so clear. The hinge is probably long enough to span the gap in either rotational position, but the $+103^{o}$ position would require it to be less

extended and to make fewer close contacts with the top of the S-domain. It is therefore tentatively assigned that way around.

In both cases, the compact to expanded transformation could probably take place by flexion at residues 268 to 271, which we will call the 'expansion hinge', and which lies a few residues nearer the N-terminus than the 'hinge' (residues 271 to 274) whose flexion differentiates the A and B position subunits from the C ones in the compact structure (Harrison, et. al., 1978).

7.1.5 P-Domains.

The most noticeable feature about the P-domains is that they are generally weaker in electron density than the S-domains. This is not an unanticipated phenomenon, as it was observed both at low and high resolution with compact TBSV, and also with TCV. What is different, is that the C position domain is quite a bit weaker than the A and B, particularly at its base, where the peaks are rarely more than two contour levels high. This cannot be explained as a truncation artefact, as this would affect the A and B positions as well. Instead, it must be concluded that the C position domains have a larger amount of spatial or thermal disorder than does the rest of the structure.

To test this hypothesis, models were built with a temperature factor introduced for the various domains, and

these were compared with the observed dataset. This was done by enlarging the radius of gyration, r_g, of each atom with the formula,

$$r_g'^2 = r_g^2 + \frac{3B}{8\pi^2}$$

An R-factor search for B_P (temperature factor for all P-domains) gave a minimum of R = .4538, r = .5426 at $B_P = 200\text{Å}^2$, compared with R = .4566, r = .5340 for the model with $B_P = 0$. Inclusion of a positive temperature factor for the S-domains only led to an increase in the R-factor. Varying the temperature factor for the A and B positions (B_{PAB}) and C position (B_{PC}) separately gave an even better agreement of R = .4481, r = .5516 at $B_{PAB} = 80\text{Å}^2$, $B_{PC} = 600\text{Å}^2$, with a clear minimum in both variables. This supports the above hypothesis of the CC dimers being considerably more disordered than the AB ones.

$B = 600\text{Å}^2$ corresponds to an r.m.s. fluctuation in each of the positional coordinates of about 3Å (Debye, 1914). This is sufficient to remove most of the high resolution detail of this P-domain, and explains the weakening of its density in the map.

7.1.6 S-P Contacts.

The base of the P-domain dimer is the shape of a rather shallow wedge. In the case of the C position dimer in the compact structure, the ridge of the wedge runs

roughly parallel to the extended arm, along the line joining neighbouring 3-fold axes (see figure 1.2). After rotation by 103°, this ridge lies perpendicular to that direction. The P-domains are displaced inwards by 5Å relative to the S-domains and this is accommodated by driving the wedge into the groove that opens between them, when they part during expansion. Some close contacts do arise from this transformation: the closest (3Å) is between the trimer contact loop and the base of the P-domain; this is known, however, to be relieved by the reconfiguration of the loop (section 7.1.3). The next closest contact (4Å) is to the S-domain half of the hinge, which is also known to move out of the way. Given these rearrangements, the closest contacts are between C143 and C289 and between B186 and C307, both of which are a healthy 8Å. These two contacts and their 2-fold related images form a four-legged support that keeps the P-domain dimer located in its observed position and prevents it from rotating in either direction.

The AB dimer similarly makes close (4Å) contacts to the loop and hinge in the existing model that will not exist when the respective modifications are made. The operational contacts appear to be between A149 and A309 (5Å) and between B149 and B356 (6Å), with all others greater than 8Å.

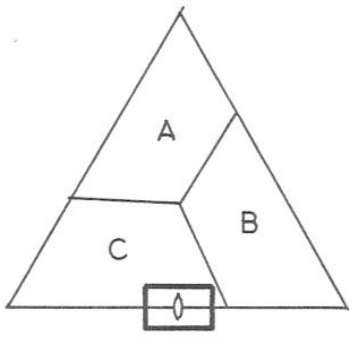
In both cases, the contact is not at all extensive, there being no large areas in close proximity, merely

turns that happen to collide. It is suspected that the final positions of the P-domains are therefore a consequence of some more basic (and functionally significant) driving mechanism located somewhere else in the structure.

7.1.7 S-S Contacts.

The majority of these are conserved in the expansion mechanism, but the trimer and dimer contacts are modified (see figure 1.2 and chapter 4). A fairly good estimate of the closeness of a contact at 8 Ångstroms resolution is whether the bottom contour level 'bridges' the density from one side to the other. The β sheets of the S-domain, for instance, have several contour levels bridging from strand to strand. As can be seen in plate 7.4(a), there is already the beginning of a cleft in the trimer contacts of the compact structure, and very few crossbridges in the density. Thus this contact probably contains a large amount of solvent, and does not lose many stabilising contacts during the expansion, except, of course, the salt bridges associated with the Ca^{++} binding.

The new contact between the A and B position S-domains has been described above (section 7.1.2). The corresponding region of the CC S-domain dimer contact is shown in plate 7.5. The absence of density here and the separation of the pseudoatoms of the model by over 8Å suggests that the contact is completely broken. The ine-



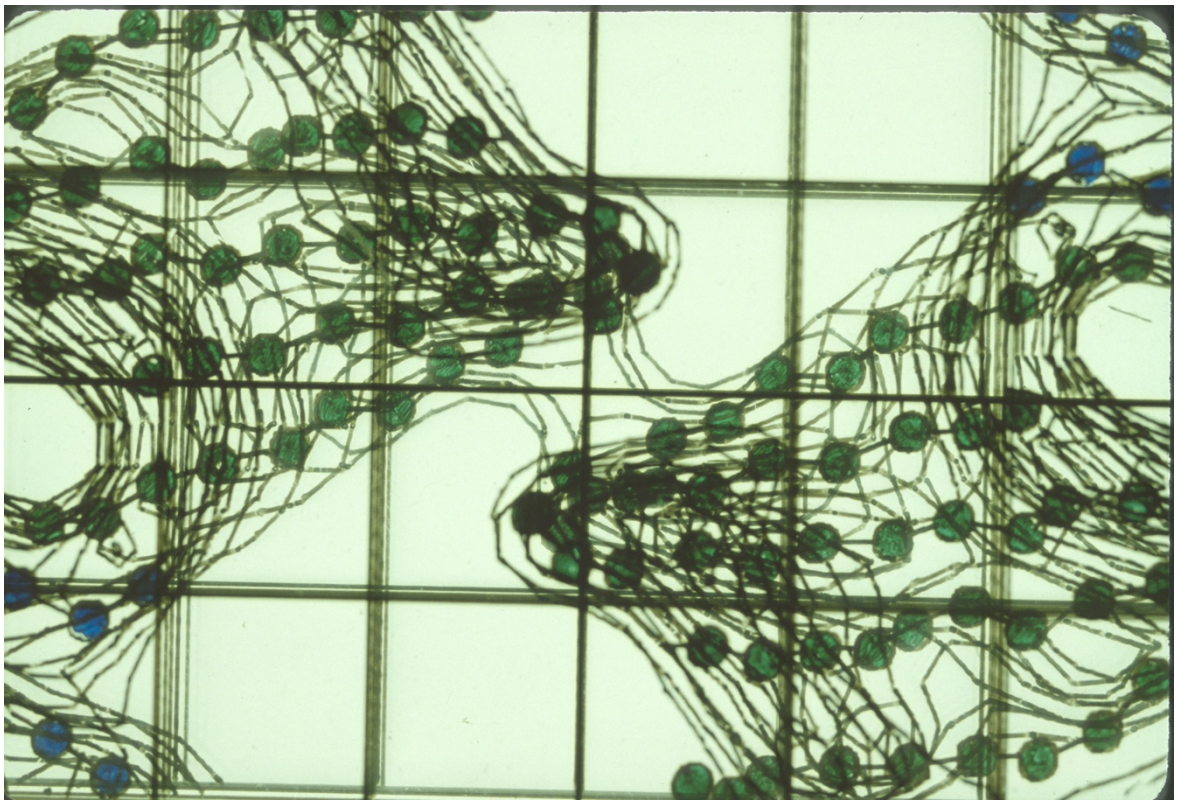


Plate 7.5. View of the interaction of neighbouring C position S-domains down the virus 2-fold axis. The contact here appears to be very limited in extent, and may involve no direct interaction of the amino acid side chains. Sections from z = 128 to 140Å are included.

quivalence of the AB and CC dimer contacts in the expanded state can also be seen from the S-domain rotational parameters in table 6.8: $\alpha_{SA} = 4.5^{o}$ is a significantly greater rotation than $\alpha_{SBC} = 2.1^{o}$ and so causes the A and B position S-domains to remain in contact but not the C position ones.

The tangential (horizontal) α helix makes a side-by-side antiparallel contact with its symmetry related mate in both the AB and CC positions of the compact structure, consisting largely of hydrophobic residues. It appears that these residues become exposed to solvent in all three subunit positions in the expanded state.

7.1.8 The Viral Interior.

The RNA-containing interior between 50Å (the inside of the envelope) and 128Å (the base of the beta annulus) shows no interpretable features. There is one highly prolate island of density, located just below the inner protein surface, that might be the average of a few ordered RNA positions, but might equally well be protein. As all the other features of the map are accounted for, there is no evidence for R-domains, either protruding through the cleft or not.

The 8 Ångstrom map of the compact structure has a 15Å diameter spherical feature just below the C position S-domain. This is interpreted as being an R-domain that is less well ordered than the rest of the structure (since it

vanishes at high resolution), but even this is missing in the map of the expanded state.

7.2 Discussion.

Detailed examination of the electron density maps has shown that the structure is consistent with the model of the expanded state, with a limited number of modifications, that are easily rationalised. The expanded state is characterised by a relatively rigid network of S-domains, stabilised by a specific new contact between the bases of the A and B S-domains. The new contact involves an entirely new conformation of the extended arm in those subunits. The arm and beta annulus of the C position subunits are conserved. The P-domains are rearranged on the surface of the virus by fairly large rotations about their local diad axes. The AB P-domain dimer is fairly well ordered, but the CC dimer is not, as indicated by its large temperature factor (600$\AA^2$) in the model calculations, by its poor refinement of heavy atom positions and by its weak density in the final map. There are two possible explanations for this disorder:

i) Spatial disorder (imperfect icosahedral symmetry, leading to a smeared-out average structure), that may be caused by the particle-particle contacts in the crystal. It seems unlikely, though, that this would affect the C positions only, as all the contacts involve both kinds of dimers (section 2.3.2). Moreo-

ver, if this were true, symmetry arguements suggest it would be the AB dimers that would be disordered, not the CC ones.

ii) Thermal disorder. There is a line of reasoning that explains why the CC dimers are more disordered than the AB ones. The S-domain constellation is stabilised by the network of P_{AA}, D_{AB}, H_{BC} and H_{CB} contacts, with an almost non-existent D_{CC} contact (plate 7.5). The internal motional modes of such a structure would involve at least small fluctuations of the separation of the C position S-domains (not enough for these to appear disordered themselves at 8 Ångstroms, apparently) but these fluctuations are likely to be amplified greatly in the P-domain dimer that is sitting across the interface. It is perhaps this thermal disorder, acting through the crystal contacts, that leads to the relatively low diffraction limit of the crystals.

7.2.1 Expansion Mechanism.

Once the Ca^{++} ions are removed from the virus at low pH, the aspartate residues in the Ca^{++} binding sites become protonated (Durham, et. al., 1977). When the pH is raised, these titrate at the same point as that at which the expansion takes place. Presumably, the mutual repulsion of the charged aspartates is the 'trigger' for the expansion. It is expected, therefore, that in the absence of any

structural change, the repulsive force between the S-domains in the trimer contact would be a continuous monotonic function of pH, roughly following the shape of the anomalous component of the titration curve. Counterbalancing this force is the attractive force between the RNA and the numerous basic amino acid side chains across the base of the S-domains, and the forces arising from the free energy of the various other contacts, particularly the β sheet across the AB interface. Beyond a critical pH, these restraining forces are overcome and the particle switches to its expanded state. The 'switch' is located in the base of the S-domain AB dimer contact, which undergoes the specific conformation change to reach the expanded state.

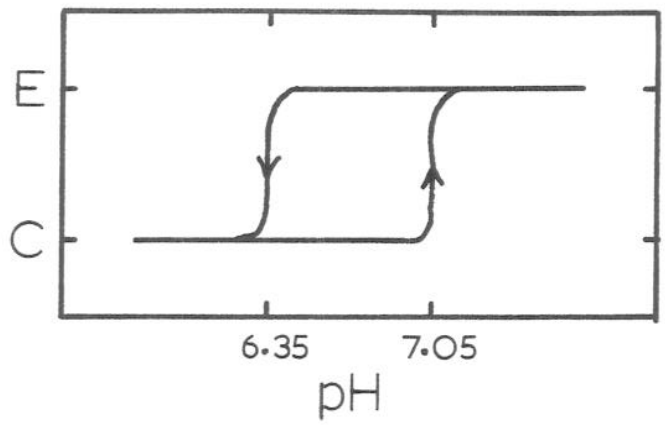
Lastly, the P-domains adopt new positions, which are most likely minimum-energy arrangements. They drop down slightly relative to the S-domains, probably to utilise the exposed surface area of the cleft that forms between these. The fact that the CC P-domains are partially disordered implies that it is very unlikely that they have a switch function, even though they do have the most noticeable positional change of the whole structure.

7.2.2 Hysteresis.

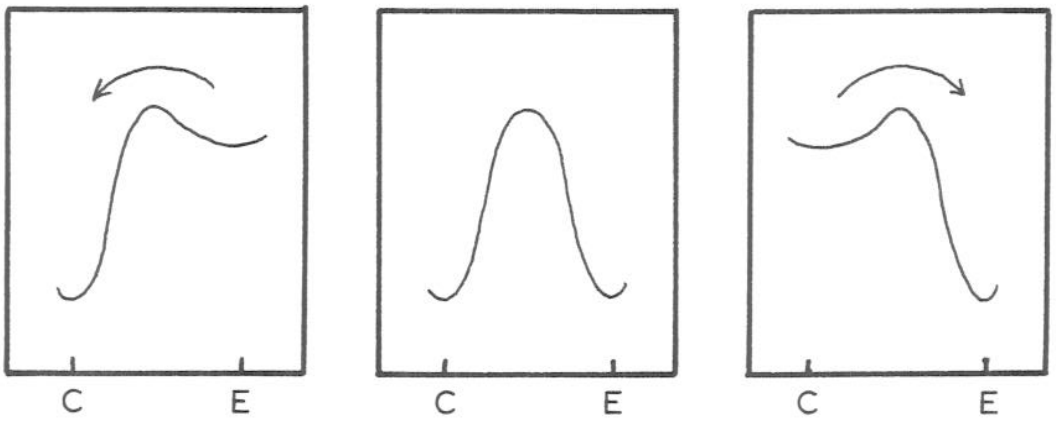
The pH hysteresis observed in the titration experiments and the corresponding hysteresis seen in the particle radius, measured by small angle scattering (see chapter 1),

is an apparent breakdown of the principle of microscopic reversibility. For every pH value, the system has one global minimum of free energy, which should correspond to the equilibrium state; here it is possible for either state to exist over a range of pH values, depending on its immediate history. The only reasonable explanation for this is that the transition between states is kinetically prohibited within this pH range, i.e. the phase transition takes place so slowly that it is never observed in practice. That slow reaction mechanisms are involved is apparent from a remark by Jacrot (1975) that the titration 'equilibrium' pH was only reached after about 30 minutes in the vicinity of the phase transition.

The proposed model for the origin of hysteresis in TBSV is shown in figure 7.3. The free energy barrier separating the compact and expanded states is so large that in the range $6.4 < pH < 7.0$ the velocity in either direction is essentially zero. For $pH < 6.4$, the expanded state is sufficiently unstable that the transition to the compact state proceeds at a reasonable rate, while the reverse is true for $pH > 7.0$. The rate constants must vary very sharply with pH for this to be true. At pH 6.9, TBSV is stable in the compact state for at least a day, with less than 10% being converted to the expanded form, as seen by solution x-ray scattering; however, when the compact virus is titrated to pH 7.1, the majority is converted to expanded within a few minutes, as seen by its proteolytic



(a) Schematic plot of the pH hysteresis in the TBSV particle radius.



pH < 6.3 6.4 < pH < 7.0 pH > 7.1

(b) Free energy diagrams of the compact ('C') to expanded ('E') transition.

Figure 7.3. Kinetic model of the pH hysteresis of TBSV. Note: the energy barrier is so large that equilibrium is never reached in the critical pH range.

sensitivity. The free energy of the compact state must therefore be increased by at least 12 kJ $mole^{-1}$ for a pH change of 0.2 units, and possibly much more since these figures were chosen very conservatively. This would clearly be impossible for the titration of a single group, but since there are 60 (and possibly 180) titrating groups acting cooperatively in the virus particle, this is not at all unreasonable. The figures reflecting the stability of the expanded state near to the reverse transition at pH 6.4 are comparable, so we must expect a cooperative mechanism in this direction too.

In the Ca^{++} free particle, the aspartates of the Ca^{++} binding site(s) titrate anomalously near neutrality. The local buildup of negative charge at pH 7.0 can adequately explain the destabilisation of the compact state there that gives rise to the 'up' transition.

The detailed mechanism for the 'down' (expanded to compact) transition at pH 6.4 is harder to identify. The 20Å solvent-filled cleft ensures that the aspartates of the Ca^{++} site will now behave normally (pK_A = 4). In the proposed model of figure 7.2(b), the stability of the expanded state is attributed to the additional strand of β sheet in the AB S-domain contact. We must look for a force that destabilises this at pH values just below neutral. Here are two models:

i) A conformational change inside the RNA or associated with the R-domains could be transmitted mechanically by the arms and so, by tearing this fifth strand away from its β sheet contacts could destabilise the expanded state. This mechanism is attractive because it implies a function for the extrusion of the R-domains in the expanded state (see below); however, it does not explain the quenching phenomenon. Another difficulty is that lysine/arginine interactions with phosphate groups are not likely to titrate significantly near neutrality; more likely, there would be a strong dependence of the expanded state stability upon ionic strength, which has not been observed.

ii) Histidine 108 is the last residue of the proposed new β contact (section 7.1.2). No details of the conformation of this amino acid can be determined with the present level of knowledge, but it is a candidate for a group titrating near neutral to trigger the 'down' transition.

7.2.3 Quenching.

In section 1.2.2, the proteolytic cleavage properties of rapidly recontracted ('quenched') TBSV were discussed. The observed accessibility to cleavage of the arm in a particle that was clearly of the compact diameter was interpreted as being due to the trapping the entire R-domains, that had migrated to the outside of the particle

while it was expanded.

In view of the structure of the expanded state, we can see that either of these is a reasonable presumption. The large cleft has a maximum width of 20Å. The 60 N-terminal amino acids would form a globular domain of normal density with a diameter of about 25Å. If it were not spherical, it would have a small dimension less than this. Thermal motion of the C position S-domains could enlarge the gap enough for even a sphere to pass through.

In the case of the B position subunit, ordered density corresponding to the arm is visible most of the way to the exterior surface of the virus. This is very direct evidence in support of the extruded R-domain interpretation of the proteolysis results.

7.2.4 Assembly.

The assembly mechanism proposed in the introduction is completely consistent with the expanded state, as it is now seen, in the role of an intermediate. The presence of an ordered C position arm and beta annulus is a prerequisite for the structure of any hypothetical intermediate, for instance. There are some features of the expansion which offer additional support for this assembly model:

i) The AB dimer is a natural choice of assembly unit as it appears to be fairly tightly bonded in the expanded structure. It is not known whether the conformation

observed in the expanded state exists in free dimers in solution, but it is worth investigation (for TCV).

ii) Because the CC dimer contacts are weak (or non-existent) in the S-domain and there is no well-ordered mode of binding the CC P-domain dimer to the S-domains, the hypothetical 'fenestrated cage' assembly intermediate of 60 subunits linked by beta annuli must be very floppy. This is probably advantageous from the point of view of assembling the remaining 60 dimers into the gaps. The structure would gradually become more rigid as the assembly proceeded, but some flexibility would remain even at the stage with only one pentagonal window unfilled.

There is a consequence of the first compaction scheme proposed above that is also relevant for assembly. If it is the pulling of the RNA from the interior upon the arms of the A and B position subunits that triggers compaction of the virus, then this is an automatic method of ensuring that capsids only become compact when they contain RNA. Furthermore, it could be that a sufficient number of arms must be pulled before this can happen, so that the capsid can only close when it contains close to its full length of 4,500 bases of RNA.

This is also consistent with the model that the R-domains lie on the exterior of the newly formed empty capsid, during assembly. As the RNA enters the shell, the

basic R-domains also pass to the inside, one at a time, to neutralise it. When an R-domain is on the outside, its arm is pulled up against the inner surface of its opposing S-domain, thus jamming the switch in the expanded state. Only when all the R-domains have passed inside is the switch free to change state.