

MECHANISM OF RNA VIRUS ASSEMBLY AND DISASSEMBLY¹

S. C. Harrison⁴, P. K. Sorger², P. G. Stockley³,
J. Hogle¹, R. Altman⁵, R. K. Strong

Department of Biochemistry and Molecular Biology
Harvard University, Cambridge, Massachusetts 02138

ABSTRACT This paper describes a series of experiments on assembly and structural rearrangements of isometric RNA virus particles in vitro, performed with a view to answering some questions about assembly mechanism, RNA specificity, requirements of the packaging process, and possible steps in disassembly. The experiments have been carried out primarily with turnip crinkle virus (TCV).

¹This work was supported by PHS Grant CA-13202, awarded by the National Cancer Institute, DHHS, and by the National Science Foundation Grant PCM 82-02821.

²Present address: MRC Laboratory for Molecular Biology, Hills Road, Cambridge CB2 2QH, England.

³Present address: The Biotechnology Unit, Department of Genetics, University of Leeds, Leeds LS2 9JT, England.

⁴Present address: Scripps Clinic and Research Foundation, 10666 North Torrey Pines Road, La Jolla, California 92037

⁵Present address: School of Medicine, Stanford University, Stanford, California 94305

TCV STRUCTURE

We first review the structure of TCV, using the description to introduce what we see as the important questions concerning assembly and disassembly.

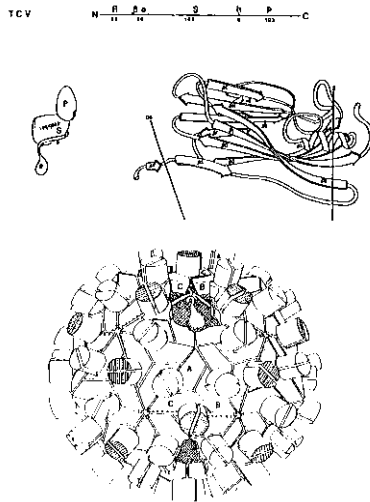


FIGURE 1. Organization of TCV. The 351-residue polypeptide chain is organized in four regions: an internal, positively charged domain (R), a connecting arm (a), a β -barrel domain that forms the shell (S), and a projecting domain (P). The fold of the chain in the S domain is shown in the middle right. The packing in the virion is shown at the bottom. There are three non-symmetry related environments: A, B and C, with 60 subunits in each type of position. The connecting arms of C subunits extend along the inner edge of the S domain (dashed lines), and they interdigitate around three-fold symmetry axes to form a β -annulus. The corresponding chain segments are marked e and β in the top part of the figure. The hinge (h) between S and P domains has one conformation in A and B subunits and a different one in C subunits.

The data in Table 1 and the overview in Fig. 1 summarize what we know about TCV from chemical and crystallographic studies. The three-dimensional structure was determined by using the atomic coordinates of tomato bushy stunt virus (TBSV), which provide a suitable first approximation for phase refinement (Hogle et al., 1986). As in TBSV, the polypeptide chain can be considered to comprise four segments: an N-terminal RNA binding region (R), an arm involved in conformational switching (a), a domain that forms the major part of the viral shell (S), and a projecting C-terminal domain (P). The S domain, shown in the secondary structural diagram in Fig. 1, is homologous to the S domains of other plant viruses such as southern bean mosaic virus (SBMV: Abad Zapatero et al., 1980), as well as to the principal domains of VP1, VP2 and VP3 in the picornaviruses (Rossmann et al., 1985; Hogle et al., 1985). Thus, it appears that the shells of many of the small, positive-stranded RNA viruses have a similar architecture, with 180 S domains (here, all identical; in picornaviruses, of three different types). The S domains of different viruses vary principally in their N-terminal extensions, in loops formed by insertions between elements of secondary structure (especially after β C, β E and β G), and in C-terminal extensions (most elaborately in these plant viruses, where the C-terminal extension forms a complete P domain).

TABLE I

Composition of TCV and TBSV⁺

	Protein:	Mr	Copies/virion	RNA
TCV	coat	39K	178	4.0 kb
	p80	78K	1	
TBSV	coat	42K	178	4.5 kb
	p80	84K	1	

p80 is a covalently linked dimer of the coat protein in both cases.

⁺ See Ziegler et al. (1974), Hopper et al. (1984), Morris and Carrington (1986), Stockley et al. (1986).

The structure can be considered as an assembly of subunit dimers - 120 in one conformation (A/B) and 60 in a second (C/C). The key conformational differences involve a hinge between S and P domains and the configuration of the arms. Only very subtle differences occur within S domains themselves, probably in response to packing differences in the capsid, which exert somewhat different forces on domains in A, B and C locations (see Harrison, 1984, for a description of these changes in TBSV). The C/C dimers have arms folded in an ordered way along the inner edge of the S domain, interdigitating around particle threefold axes so as to form a coherent inner scaffold. The arms of A/B dimers project in a less ordered way into the interior of the particle.

RNA packaging in these particles appears to have evolved to accommodate irregularities and changes in the nucleic acid. The R segments are not uniformly ordered, and since their principal characteristic is a high concentration of lysine and arginine residues, they serve to clamp onto the RNA. Moreover, there are no defined binding sites for RNA on the inner face of the S domain, although a number of basic residues project inward and provide further positive charge to neutralize the phosphates. Thus most of the interaction with RNA is with parts of the polypeptide chain that are flexibly linked to the S domain shell, and no highly ordered RNA is seen crystallographically. The nucleic acid is nonetheless very tightly packed - as tightly as tRNA in crystals - and there appears to be little motion (Munowitz et al., 1980). The internal disorder is therefore a reflection of the irregularity of the non-repeating RNA structure and of the way the R segments accommodate to it.

The R segment of TCV can be prepared by chymotryptic cleavage of dissociated virus (data not shown). A strong cleavage occurs at tyr 66, as in expanded virions. Studies of the purified peptide, which corresponds to residues 1-66 (as confirmed by amino acid analysis), show little evidence for a rigidly folded structure. In particular, urea gradient gels do not give evidence for a cooperative unfolding transition, and the peptide is equally sensitive to digestion by thermolysin at 0 and 25 °C (data not shown). The properties of the isolated R region are thus consistent with its role in the virion as an accommodating clamp.

As indicated in Table 1, the minor species p80, present in one copy per virion (Ziegler et al., 1974), is a covalent dimer of the coat protein (Stockley et al., 1986). The covalent linkage is at or near the N terminus, but its chemistry is not yet known. Since this part of the subunit is not rigidly folded, there are no new constraints on the contacts, and p80 can fit into the shell just as a 'normal', unlinked dimer. We believe that the structure is thus composed of 178 copies of p40 (89 noncovalent dimers) and 1 copy of p80. A number of other plant viruses with T=3 icosahedral shells appear to contain about one copy of species having twice the M_r of the coat protein (Rice, 1974), suggesting that these dimers have a conserved function.

EXPANSION OF TBSV AND TCV

A number of T=3 plant viruses have been shown to undergo an expansion at neutral to alkaline pH, when divalent cations are sequestered (Incardona and Kaesberg, 1974; Robinson and Harrison, 1982; Kruse et al., 1982). The expanded form of TBSV is a relatively well-ordered particle, and its crystal structure is known (Fig. 2; Robinson and Harrison, 1982). There are Ca^{++} binding sites at the subunit interfaces indicated in Fig. 2, and removal of the ions leaves apposed aspartic acid groups titrating at about pH 7. The instability of negatively charged groups in close vicinity causes the subunit interfaces to dissociate and the rearrangements shown in Fig. 2 to occur. The principle changes can be described as rigid-body motions of domains with respect to each other. TCV undergoes a similar transition. It has not been crystallized, but small-angle X-ray scattering shows a similar increase in radius in both viruses, and we can reasonably suppose that the expanded TCV is essentially similar in structure to expanded TBSV.

An additional rearrangement accompanying expansion of TBSV and TCV, not evident from the crystallographic results, is exposure of the subunit arms. Native TBSV is completely resistant to digestion by various proteases, but proteolysis of expanded virions leads to cleavage. With chymotrypsin, one third of the subunits are cleaved to 35 kD fragments, and the remaining two-thirds, to 30 kD fragments (Fig. 3). With V-8 protease, only two-thirds of the subunits are sensitive to degradation, to fragments of about 38 kD. Evidence that these cleavages occur in the

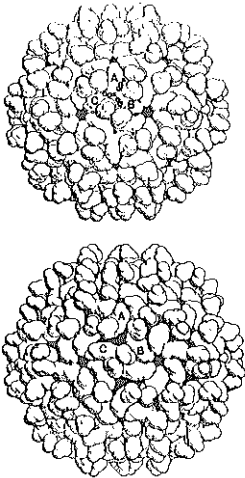


FIGURE 2. The compact and expanded structures of TBSV (from Robinson and Harrison, 1982). Some of the Ca sites that regulate the expansion are shown as circles in the upper figure. All positions symmetry-related to these also have cation sites. The corresponding interfaces move apart in the expanded structure.

amino-terminal part of the subunit is described by Golden and Harrison (1982). The position of V-8 cleavage can be uniquely assigned to glu 36, the only negatively charged residue in the first 100. The 30 kD chymotryptic fragment had probably been cleaved at phe 95, and the 35 kD chymotryptic fragment at leu 48 or leu 53. All these cleavages lie in the R segment or in the arm. The two-thirds/one-third ratio suggests differential behavior of A, B and C subunits. Expanded, cleaved and recontracted TBSV can be crystallized, and a 2.9 Å difference map shows no significant changes in ordered density (Altman, unpublished). In particular, the region around phe 95 in the C subunit arm is undisturbed. Thus, C subunits have been cleaved near residue 50 A and B subunits at residue 95.

Similar chymotryptic cleavage is observed in expanded TCV (Fig. 4), but here the extent of digestion is less extensive (one-third of the subunits) and ionic-strength dependent. The cleavage point is at tyr 66, as determined by direct sequence analysis (see Hogle et al., 1986). Sedimentation profiles show that the cleaved TCV is a homogeneous species - that is, cleavage has occurred on one-third of the subunits of all virions, not on all subunits of one-third of the virions.

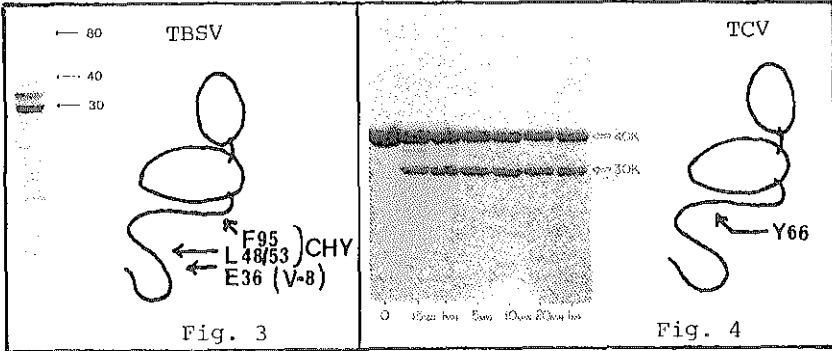


FIGURE 3. Proteolytic cleavage of expanded TBSV. (a) Gel lane shows endpoint of chymotryptic cleavage of expanded virus (3 hours, room temperature, 1:100 w/w enzyme/virus). One third of the subunits have been cleaved to an approximately 35 kD species and two thirds to an approximately 30 kD species. (b) Diagram showing sites of chymotrypsin and V-8 cleavage using schematic representation of folded subunit in C conformation (see Fig. 1). The cleavages at positions 36 and 48/53 probably occur only on A/B subunits.

FIGURE 4. (a) Cleavage of expanded TCV by chymotrypsin at low ionic strength (50 mM Tris, 5 mM EDTA, pH 8.5). Note that one third of subunits are cleaved to 30 kD species. (b) Position of chymotryptic cleavage (tyr 66).

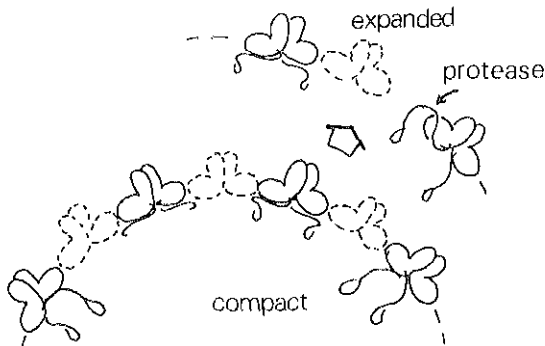


FIGURE 5. Model for looping out of arms when TCV and TBSV shells expand. In order to be exposed to chymotrypsin, the arms must move through a distance of at least 20 Å.

The proteolytic accessibility of arms and R segments is unexpected, since gaps in the expanded structure are not wide enough to admit molecules as large as chymotrypsin or V-8 protease. Our observations may be explained either by assuming that the expanded particle 'breathes' extensively, permitting entry of protease, or by postulating extension of the arm itself. Several arguments indicate that breathing alone cannot account for the cleavages. (1) Proteolytic cleavage of arms in rapidly recontracted virus suggests that accessible arms have been trapped in an exposed configuration (Sorger, unpublished). (2) RNase does not gain access to the particle interior, as judged by failure to observe RNA degradation when expanded TBSV is exposed to the enzyme (Kruse et al., 1982; Sorger and Harrison, unpublished). RNase A is substantially smaller than chymotrypsin. (3) TCV arms can be cleaved at low but not at moderate ionic strength. The radius of the expanded particle is identical in the two cases, and we therefore have no reason to believe that S domain arrangements are different. The most straightforward explanation of the observed cleavage of 1/3 of the subunits is a salt-dependent conformational change in one of the three classes of arms (probably A or B, see below). (4) V-8 protease cleaves 2/3, but not all, of the TBSV R domains. Such differential accessibility is better explained by specific conformational changes of the arms than by a general breathing of the particle. Each of these arguments is not complete by itself, but taken as a whole, the observations are difficult to reconcile with a picture other than emergence of a particular class of arms under particular expansion conditions.

A diagram with our interpretation of these results appears in Fig. 5. The tight packing of RNA and R segments in the viral interior is sufficiently relaxed on expansion that arms can loop out through the gaps between subunits shown in Fig. 2. There is good evidence from NMR studies for increased internal mobility in expanded TBSV (Munowitz et al., 1980). The similarity between plant and picornavirus structures suggests a parallel between the capacity of the subunit arm to extrude in this way and the loss of VP4 when picornaviruses bind to cells (Rueckert, 1985; Hogle et al., 1985; Rossmann et al., 1985). Some expansion of the latter structures must occur, in order to permit exit of VP4.

TCV ASSEMBLY

The mechanism of in vitro assembly can conveniently be studied, since disassembly of TCV can be carried out under relatively mild, non-denaturing conditions. Dissociation occurs at pH 8, 1mM EDTA (i.e., expansion conditions), if the salt concentration is raised to 0.5M (or higher); it proceeds far more rapidly at 0°C than at room temperature. Efficient reassembly can be achieved by two-step dialysis (first to pH 6-7, then to 0.1 M NaCl) or by dilution to pH 7, 0.1M NaCl (Sorger et al., 1986).

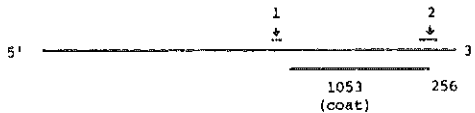
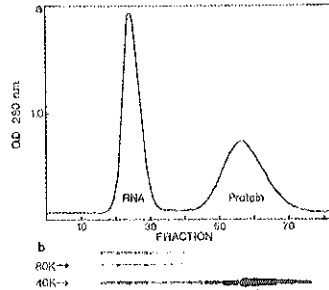
The rp-complex.

The products of TCV disassembly are dimers of the coat protein subunit p40 (Golden and Harrison, 1982), and a complex of TCV RNA with six subunits of p40 and the unique copy of p80 (Fig. 6: Sorger et al., 1986). The complex of RNA and protein, which we term 'rp-complex,' is stable and resistant to high salt concentrations (1.5M). Several lines of evidence show that the protein is bound to specific sites or structures on the RNA. (1) When lightly labelled by reductive methylation with ³H-formaldehyde, rp-complex is competent to reassemble. The label is retained on RNA as rp-complex if the reassembled particles, made with unlabelled coat protein dimers, are disassembled in the usual way. (2) Particles, reassembled with protein-free RNA (phenol extracted) and normal p40 dimers, can be disassembled to yield rp-complex (and free coat dimers). This experiment also shows that p80 is not required for assembly (see below). (3) Extensive digestion of rp-complex with T1 ribonuclease yields a set of seven protected fragments about 30-50 bases in length. The sequences of these fragments have been determined; three correspond to segments of the 3' one-third of the genome for which sequence is known (Carrington et al., submitted to J. Mol. Biol.). These sequences and their location are shown in Fig. 7.

Reassembly.

The properties of the in vitro reassembly reaction can be summarized as follows: (1) It proceeds from coat-protein dimers and rp-complex, or from coat-protein dimers and free RNA. In the latter case, the product lacks p80, which is

FIGURE 6. (a) Elution profile from Sephadex G-150 of TCV dissociated as described (Sorger et al, 1986). (b) SDS-polyacrylamide gel pattern of fractions from the G-150 column. Loadings from the 'protein' peak were only 1/20 as great as those from the 'RNA' peak. Protein in the 'RNA' peak shows that it is in fact an rp-complex (six subunits per RNA molecule).



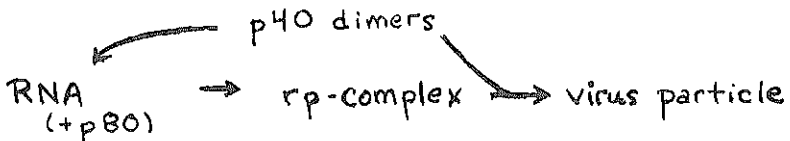
- (1) GAUUCACACAUCCUACAACACACACGACUCAUC
- (2) GAAAAUGGUCACAACUGAGGAGCAGCCAAAGG
GUAAAUGGCAAGCACUCAGAAUUUAGUACGGUAA

FIGURE 7. Sequences of three RNA fragments found in ribonuclease T1 limit digest of rp-complex. Fragments were isolated by filter binding of complex after digestion. Sequence 2 corresponds to two contiguous fragments. Four other T1 fragments were reproducibly isolated, but their sequences are not found in the 3' one-third of the genome for which nucleotide sequence is presently available (Carrington et al., submitted to J. Mol. Biol.).

therefore not required for morphologically correct assembly. (2) The conditions of optimal assembly are approximately physiological (pH 7, 0.1M NaCl). (3) Assembly requires RNA and intact protein subunits. No defined structures are obtained in the absence of RNA; T=1 RNA-free, 60-subunit particles are obtained from subunits that have been proteolytically cleaved (Fig. 8). The T=1 particles have all subunits in the A/B conformation - the arm is required for conformational switching, as the character of the internal framework clearly implies. (4) The viral shell grows by addition of subunit dimers to an initiating structure. There is no evidence for defined intermediates, other than rp-complex and coat-protein dimers. Moreover, the curvature of the shell is correctly determined as it forms. That is, the selection of A/B or C/C conformation occurs as a dimer adds - subsequent adjustment or rearrangement does not occur (Fig. 9). Multiple nucleation on a single RNA molecule can lead to non-coalescing partial shells, since independently nucleated arm frameworks cannot in general merge with each other, but even these partial shells have correct local morphology (Fig. 9). (5) Assembly is selective for viral RNA. Although heterologous RNA can be packaged - 18S or 28S chick ribosomal RNA is a particularly efficient substrate - competition experiments show a strong preference for TCV RNA (Sorger et al., 1986). Free RNA is as effective as rp-complex, since under the conditions we have studied, rp-complex formation is not rate limiting. That is, when free TCV RNA is used, rp-complex is rapidly reconstituted, and free RNA thus competes as effectively as preformed complex when challenged with rRNA (Sorger et al., 1986). It is clear from the formation of 'monsters' under conditions of multiple nucleation (Fig. 9) that unique initiation is physico-chemically as well as biologically significant, and the nucleating structure can account for both accuracy and specificity in assembly.

Mechanism.

The reassembly reaction can be represented as follows:



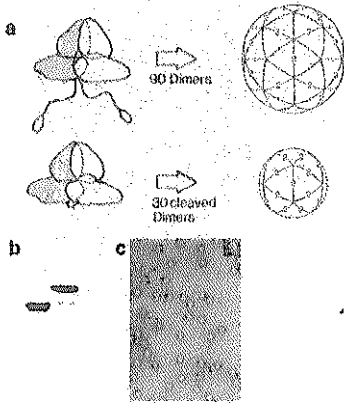


Fig. 8

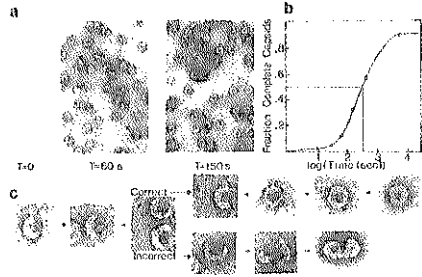


Fig. 9

FIGURE 8. (a) Intact TCV subunit dimers can reassemble to form normal T=3 capsids in the presence of RNA. Cleaved (30K) subunits assemble into RNA-free, 60-subunit, T=1 particles. (b) SDS-polyacrylamide gel of cleaved (left) and intact (right) subunits. (c) Reassembly from cleaved subunits yields T=1 particles shown in this micrograph. A T=3 particle is shown for comparison in the inset. (Sorger et al, 1986).

FIGURE 9. Dissociated TCV was reassembled by dilution to pH 7. Samples were withdrawn at various times, applied to carbon-coated EM grids, and negatively stained with uranyl acetate. (a) Typical micrographs at indicated times. (b) Fraction of particles assembled, expressed as a ratio of all visible fragments, plotted versus time. (c) A synthetic 'time course' of assembly for both native and aberrant products. The 'monster' particles appear to represent multiple nucleation on one RNA. (Sorger et al, 1986).

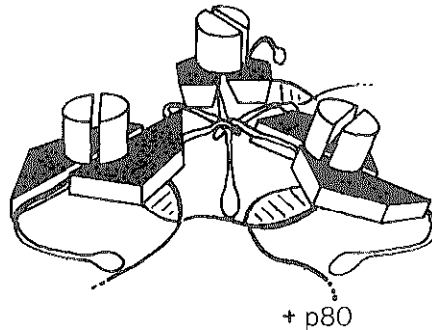


FIGURE 10. Proposed structure for the rp-complex: a trimer of dimers, connected by a β -annulus, bound to a defined structure on the viral RNA. The binding of p80, also a part of rp complex from normal virions, is probably at a different site on the RNA. For further discussion, see text.

It is clear that understanding the structure of the rp-complex is critical for visualizing the mechanism. The structure we have proposed for the complex is shown in Fig. 10. The evidence for this model comes largely from the experiments summarized above. Since rp-complex can be produced by assembly in protein-free RNA, since it survives a round of reassembly/disassembly, and since it appears as a product of gentle dissociation, it is likely to be a sub-structure of the virion itself. That is, the interactions among its subunits and between these subunits and RNA are likely to persist in the assembly process and in the completed particle. Moreover, since assembly can occur on heterologous RNA, we suggest that bonding properties of the subunits, rather than tertiary structure of an RNA scaffold, determine the organization of the initiating complex. There are two types of trimer interactions in the structure - the set of three C subunits linked by a β -annulus and the 'ABC' trimer linked by Ca^{++} -site interfaces (Hogle et al. 1983). The Ca^{++} -mediated, local threefold contacts are selectively destabilized in the conditions in which we prepare rp-complex, making the 'ABC' structure an unlikely model. The set of C/C dimers joined by interdigitating arms is, by contrast, an attractive model, since it is easy to see how binding of three arms or R domains to a particu-

lar RNA structure could stabilize a β -annulus. A set of three dimers forming a structure of this type is also a uniquely effective starting point in a model for accurate assembly, illustrated in Fig 11. Note that in the structure depicted in Fig 10, we show the arms in the tight conformation they adopt in a C/C dimer. In the free rp-complex, the linkage between β -annulus and S-domain may be relatively flexible, with addition of further dimers (Fig 11) required to lock it into the precise conformation shown.

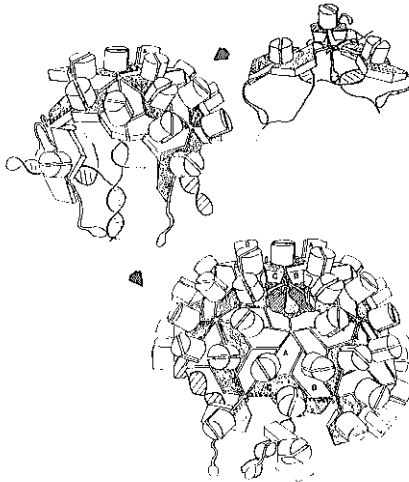


FIGURE 11. Proposed mechanism for TCV assembly. (Sorger et al, 1986).

The assembly model in Fig 11 is based on the postulated rp-complex structure, on the assumption that subsequent dimer incorporation will occur at sites of extensive interaction with the existing assembly, and on the cooperativity of the conformational switch in the coat protein dimer. The first additions of dimers to the rp-complex thus occur around the initial β -annulus, and these dimers adapt the A/B conformation because the existing structure prevents folding of their arms. The formation of B-C and C-B interfaces also locks in the arms of the rp-complex

around the β -annulus. As addition continues (Fig 11b), new positions for β -annular interaction are determined by the arms of the 'other subunit' in each of the initial three dimers. That is, the two polypeptide chains in C/C dimers interact closely in the region where arms fold back into S domains, and locking in of the folded arms in the first β -annulus also stabilizes the folded arm and the C conformation of the second member of each dimer. Propagation of the structure thus consists of β -annulus formation by dimers that consequently have the C/C conformation, in alternation with 'filling-out' steps by dimers that retain the A/B conformation. Recruitment of dimers to the assembly is probably accelerated by non-specific interaction of R domains with RNA (Fig 11b), and this interaction ensures that RNA is gathered into the shell as it forms.

Correct closure is guaranteed by the mechanism depicted in Fig 11, provided that multiple nucleation has not occurred on the same RNA molecule (see 'monster' in Fig 9). A possible position for p80, the unique, covalently-linked coat-protein dimer, is as the last unit to enter the shell (Fig 12). If it binds near one end of the RNA, this termination step can ensure complete packaging by preventing a protruding end or loop from inhibiting closure (Fig 12). A more important function might then be in the initiation of disassembly. Wilson and co-workers have observed that SBMV will direct *in vitro* synthesis of viral protein from intact virions and suggest a 'weak capsid element' that is displaced to allow ribosome binding (Brisco et al, 1985).



FIGURE 12. Diagram indicating one proposal for the function of p80, as a 'last in, first out' element of the viral shell. The shaded dimer represents the covalently linked p80.

SBMV and many other plant viruses appear to contain a species similar to p80 - that is, a component with M_r equal to twice that of the coat protein (Rice, 1974). Like TCV, these viruses assemble from dimers. Thus, there appears to be some importance to whatever function is served by the covalently linked species.

An important feature of the mechanism shown in Fig 11 is that it involves purely local interactions in the determination of long-range structural characteristics (curvature and closure). The 'decision' to adopt an A/B or C/C conformation is completely determined by interactions possible when a dimer adds. Moreover, these interactions do not need to be sensitive to signals transmitted across a domain or to small conformational distortions of a folded structure. When a dimer 'docks' into a shell there are two possible outcomes. If a β -annular interaction can form with an arm already folded and set to nucleate it, this interaction will occur, and a C/C dimer will result. Otherwise, the arms will remain unfolded, and the dimer will have the A/B conformation. A mechanism of this kind is robust to small changes in the conditions under which assembly occurs and to mutations that do not directly destabilize the critical interactions.

REFERENCES

- Abad-Zapatero, C., Abdel-Meguid, S.S., Johnson, J.E., Leslie, A.G.W., Rayment, I., Rossmann, M.G., Suok, D., and Tsukihara, T. (1980). *Nature* 286, 33-39.
- Brisco, M.J., Hull, R.C. and Wilson, T.M.A. Submitted to *Virology*.
- Golden, J.S. and Harrison, S.C. (1982). *Biochemistry* 21, 3862-3866.
- Hogle, J., Kirchhausen, T., and Harrison, S.C. (1983). *J. Mol. Biol.* 171, 95-100.
- Hogle, J.M., Chow, M. and Filman, D.J. (1985). *Science* 229, 1359.

- Hogle, J., Maeda, A. and Harrison, S.C. (1986). J. Mol. Biol., in press.
- Hopper, P., Harrison, S.C., and Sauer, R. J. Mol. Biol. 177, 701-713 (1984).
- Incardona, N.L. and Kaesberg, P. (1974). Biophys. J. 4, 11.
- Kruse, J., Kruse, K.M., Witz, J., Chauvin, C., Jacrot, B., and Tardieu, A. (1982). J. Mol. Biol. 162, 393-417.
- Morris, T.J. and Carrington, J.C. (1986). In Isometric Plant Viruses with Monopartate RNA Genomes, ed. R. Koenig (Plenum, New York).
- Munowitz, M.G., Dobson, C.M., Griffin, R.G., and Harrison, S.C. (1980). J. Mol. Biol. 141, 327-333.
- Rice, R.H. (1974). Virology 61, 249-255.
- Robinson, I.K. and Harrison, S.C. (1982). Nature 279, 563-568.
- Rossmann, M.G., et al. (1985). Nature 317, 145.
- Rueckert, R.R. (1985). In Virology, ed. B. Fields (Pergamon, New York).
- Sorger, P.K., Stockley, P.G. and Harrison, S.C. (1986). J. Mol. Biol., in press.
- Ziegler, A., Harrison, S.C., and Leberman, R. (1974). Virology 59, 509-515.

