Title
Refined structure of desmodium yellow mottle tymovirus at 2.7 Å resolution11Edited by T. Richmond

Journal
Journal of Molecular Biology, 301(3)

ISSN
0022-2836

Authors
Larson, Steven B
Day, John
Canady, Mary A
et al.

Publication Date
2000-08-01

DOI
10.1006/jmbi.2000.3983

doi:10.1006/jmbi.2000.3983 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2000) 301, 625–642

JMB

Refined Structure of Desmodium Yellow Mottle Tymovirus at 2.7 Å Resolution

Steven B. Larson[1], John Day[1], Mary A. Canady[2], Aaron Greenwood[1] and Alexander McPherson[1]*

[1]*Department of Biochemistry University of California, Irvine CA, 92697-3900, USA*

[2]*Department of Molecular Biology, The Scripps Research Institute, La Jolla CA 92037, USA*

**Corresponding author*

Desmodium yellow mottle virus is a 28 nm diameter, $T = 3$ icosahedral plant virus of the tymovirus group. Its structure has been solved to a resolution of 2.7 Å using X-ray diffraction analysis based on molecular replacement and phase extension methods. The final R value was 0.151 ($R_{free} = 0.159$) for 134,454 independent reflections. The folding of the polypeptide backbone is nearly identical with that of turnip yellow mosaic virus, as is the arrangement of subunits in the virus capsid. However, a major difference in the disposition of the amino-terminal ends of the subunits was observed. In turnip yellow mosaic virus, those from the B and C subunits comprising the hexameric capsomeres formed an annulus about the interior of the capsomere, while the corresponding N termini of the pentameric capsomere A subunits were not visible at all in electron density maps. In Desmodium yellow mottle tymovirus, amino termini from the A and B subunits combine to form the annuli, thereby resulting in a much strengthened association between the two types of capsomeres and an, apparently, more stable capsid. The first 13 residues of the C subunit were invisible in electron density maps. Two ordered fragments of single-stranded RNA, seven and two nucleotides in length, were observed. The ordered water structure of the virus particle was delineated and required 95 solvent molecules per protein subunit.

Keywords: X-ray diffraction; water structure; Desmodium yellow mottle virus; macromolecular interactions; tymoviruses

Introduction

Desmodium yellow mottle virus (DYMV), first reported by Walters & Scott (1968, 1972), is a member of the tymovirus group of small spherical plant viruses. The $T = 3$ virion is composed of 180 chemically identical subunits arranged according to icosahedral symmetry. The virus infects a rather narrow range of plants that include Phaseolus and some other leguminosae, in which it produces necrotic lesions. The isometric particle, from electron microscopy, has a diameter of about 28 nm. The virus is composed of 65% protein and 35% nucleic acid. The nucleic acid is a single-stranded RNA genome containing 6.3 kb composed of 16.5% guanosine, 22.5% adenosine, 23.8% uridine and, characteristic of tymoviruses, an unusually high cytosine content of 37.2% (Scott & Moore, 1972). The protein subunits contain 188 amino acid residues with $M_r = 20{,}020$.

Purified preparations of DYMV contain two components when analyzed by ultracentrifugation, one sedimenting at 114 S and the other at 54 S (Scott & Moore, 1972). The latter component corresponds to RNA-free capsids. Formation of empty capsids suggests that virion assembly is primarily driven by protein-protein interactions that do not likely involve cooperative, co-condensation of protein and nucleic acid. The RNA retains infectivity even when deproteinized with phenol or detergent (Scott & Moore, 1972).

The type member of the tymovirus group is turnip yellow mosaic virus (TYMV), among the first discovered of the spherical viruses, and perhaps the most thoroughly studied by physical-

Abbreviations used: DYMV, Desmodium yellow mottle virus; TYMV, turnip yellow mosaic virus; CCMV, cowpea chlorotic mottle virus; PhMV, Physalis mottle virus; BIDG sheet, CHEF sheet, β-sheets of the β-barrel; STMV, satellite tobacco mosaic virus; ASA, accessible surface area; CGM, conjugate-gradient minimization; SA, simulated annealing.

E-mail address of the corresponding author: amcphers@uci.edu

chemical methods (Hirth & Givord, 1988; Matthews, 1991). The structure of TYMV has been determined to 3.2 Å resolution by X-ray crystallography (Canady *et al.*, 1996). The protein subunits were found to exhibit the canonical "jellyroll", anti-parallel β-barrel fold characteristic of virtually all spherical plant viruses. Similarities to cowpea chlorotic mottle virus (CCMV; Speir *et al.*, 1995) were noted, most important of which is the appearance of polypeptide annuli interior to the quasi 6-fold axes, constructed from the N termini contributed by the B and C subunits comprising the hexameric capsomere (referred to as a β-hexamer in CCMV). The recently reported structure of another tymovirus, Physalis mottle virus (PhMV), also exhibits these annuli around the quasi 6-fold axes (Krishna *et al.*, 1999).

DYMV and TYMV have 44% identity in their amino acid sequences and, though serologically distinct (Walters & Scott, 1968, 1972), were indicated by several studies (Walters & Scott, 1972) to be highly similar in structure. In addition to the major genomic RNA strand, TYMV contains a small subgenomic RNA. It is still unknown whether DYMV also contains a subgenomic RNA.

Results and Discussion

The final model consists of three crystallographically independent protein subunits designated A (residues 1 to 188), B (residues 1 to 188) and C (residues 14 to 188), 285 water molecules, and two generic RNA fragments of seven and two uridine nucleotides. The final R value is 0.151, with a corresponding R_{free} of 0.159 for 134,454 reflections having $F > 3\sigma$. The estimated error in the model coordinates is 0.25 Å and 0.30 Å, based on Luzzati plots (Luzzati, 1952; Kleywegt *et al.*, 1994; Kleywegt & Brünger, 1996) and the program SIGMAA (Read, 1986; Kleywegt & Brünger, 1996), respectively. Figure 1 illustrates through a Ramachandran plot (Ramachandran & Sasisekharan, 1968) that the structure is of high quality with only 1.5% of the residues outside of the most favored and allowed ϕ-ψ regions.

Although there is only 44% sequence identity between DYMV and TYMV, there is a high degree of amino acid homology (67%) elsewhere. As a result, the structures of the DYMV subunits and indeed the entire virion structure are remarkably similar in most respects to the corresponding subunits and overall structure of TYMV (Figure 2). Therefore, it seems most expeditious to describe and analyze the structure of DYMV by comparison with TYMV, cursorily treating the similarities and emphasizing the differences between the two.

Virion architecture

Capsids of $T = 3$ viruses are composed of three icosahedrally distinct subunits, designated A, B and C after the nomenclature used by Harrison *et al.* (1978). A subunits assemble, as shown in

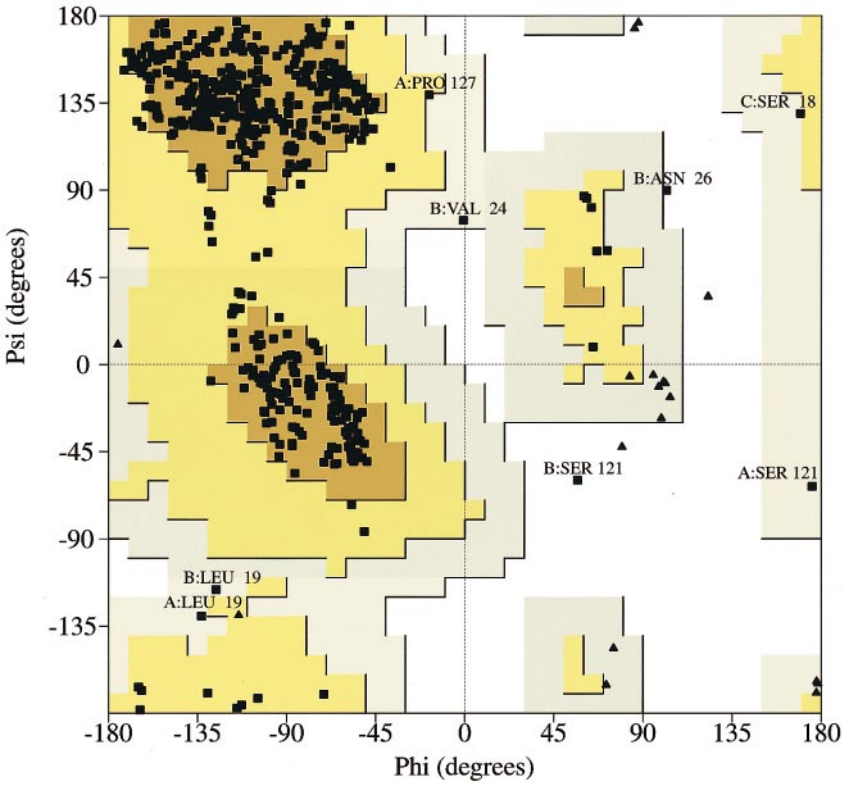


Figure 1. A Ramachandran plot (Ramachandran & Sasisekharan, 1968) of the peptide conformations within the three independent subunits of DYMV produced with PROCHECK (Laskowski *et al.*, 1993). Glycine residues are shown as triangles (▲). Disallowed regions are white and generously allowed regions are cream; together they contain seven non-proline and non-glycine residues (1.5%) as indicated by the labeled markers. These residues are located in the loops around the quasi 3-fold axis (residues 18 to 26) and in the FG-loops around the 5-fold or 3-fold axes (residue 121). The rmsds in the geometry of the model are: for protein bonds, 0.009 Å; angles, 1.83°; dihedral angles, 18.8°; improper angles, 1.18°; for RNA bonds, 0.015 Å; angles, 2.38°; dihedral angles, 39.3°; improper angles, 2.93°.

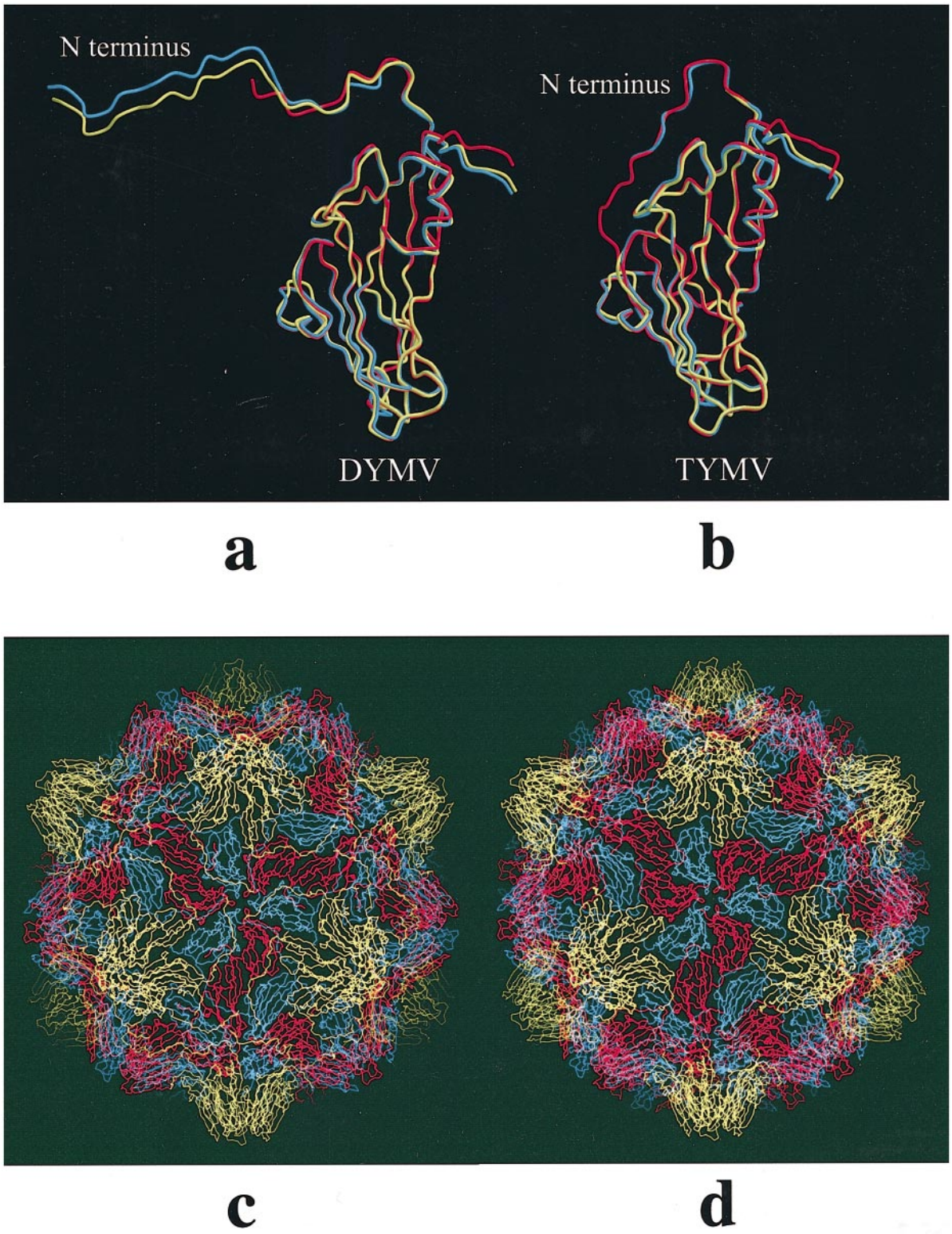


Figure 2. Superposition of the A and C subunits onto the B subunit of (a) DYMV, and (b) TYMV. The coloring scheme used here and throughout the Figures for both viruses is: A subunits, yellow; B subunits, blue; and C subunits, red. Subunits are superimposed by the main-chain atoms C^{α}, C and N of the β-strands of the β-barrels (a fit of 252 atoms). The rmsd are: for DYMV, 0.40 to 0.46 Å; for TYMV, 0.35 to 0.52 Å; for corresponding subunits between DYMV and TYMV, 0.42 to 0.46 Å. The N termini are the only conformationally distinct regions between the two viruses. (c) The DYMV and (d) the TYMV full capsid structure displayed as C^{α} tracings and viewed down the 3-fold axis. The N termini of the A subunits are especially pronounced in the DYMV structure.

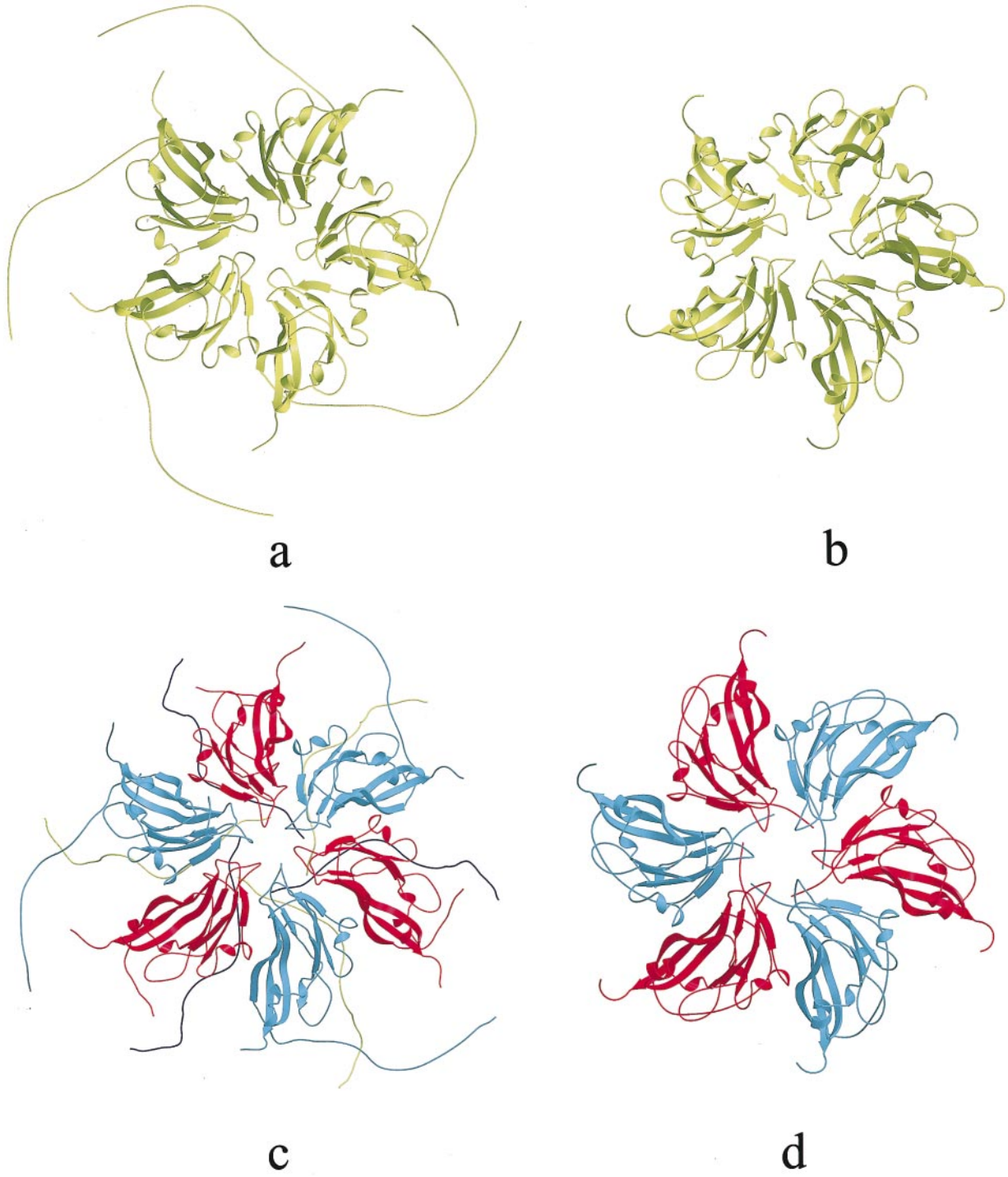


Figure 3 (*legend opposite*)

Figure 3(a) and (b), to make pentameric capsomeres. B and C subunits cooperate as three pairs, as shown in Figure 3(c) and (d), to make hexameric capsomeres having exact 3-fold and quasi 6-fold symmetry. The three types of subunits form trimeric units possessing quasi 3-fold symmetry (Figure 3(e) and (f)).

The DYMV particle is approximately 280 Å in diameter with a mean protein diameter of 258 Å (Figure 4). The exterior topography has peaks at the 5-fold and 3-fold (quasi 6-fold) axes between which are valleys, 29 to 35 Å deep, at the quasi 3-fold axes. On the interior surface of the capsid there are conical cavities around the 5-fold and 3-fold axes with depths of 28 Å and 19 Å and diameters at their openings of 60 Å and 78 Å, respectively. The viral protein shell is 15 to 20 Å thick. The peaks at the 5-fold and 3-fold axes have

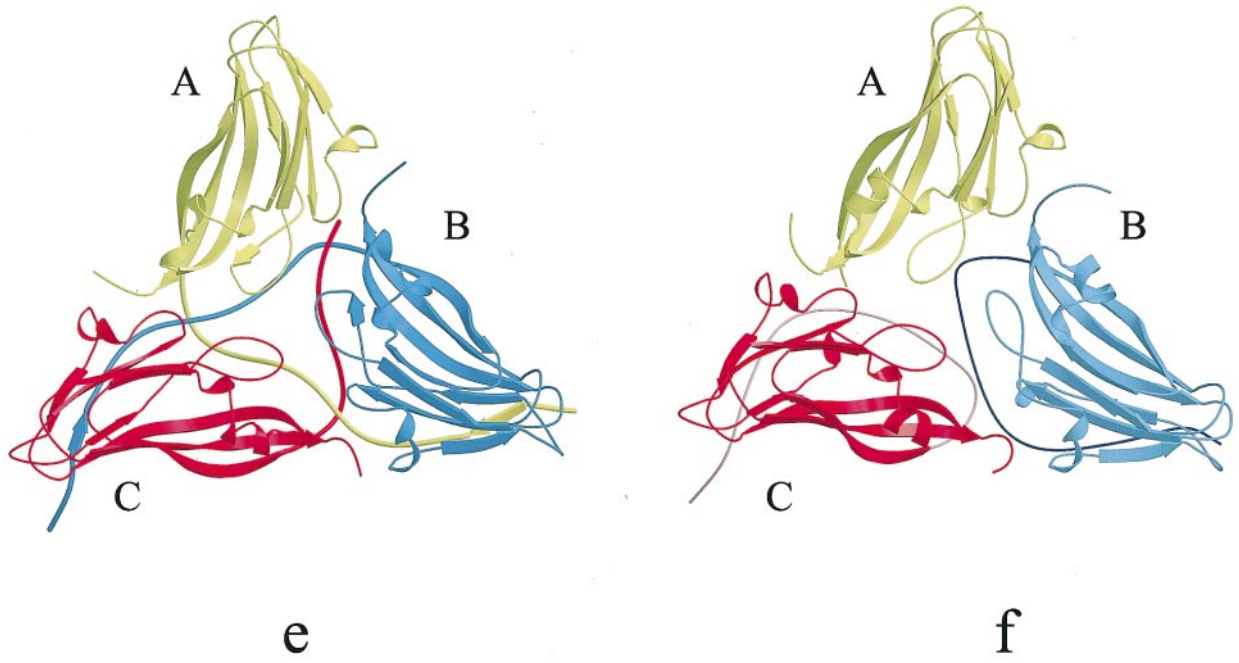


Figure 3. (a) DYMV and (b) TYMV pentamers of A subunits viewed down their 5-fold axes. The extended N termini of the DYMV subunits represent a prominent departure; otherwise, the pentamers are very similar. (c) DYMV and (d) TYMV hexamers of alternating B and C subunits viewed down their 3-fold axes. The TYMV hexamer is self-contained, whereas the DYMV hexameric unit has N termini from adjacent A and B subunits forming an annulus inside its cavity. (e) DYMV and (f) TYMV trimeric icosahedral asymmetric units. The dispositions of the N termini of the DYMV subunits result in a stronger binding within the trimeric unit than can be created by the self-associating dispositions of the N termini in the TYMV subunits.

craters or indentations of 13 Å and 20 Å, respectively. The pores at the 5-fold, 3-fold and quasi 3-fold axes are less than 1 Å in diameter, too small for the passage of ions or water without movement of the Val122 or Thr145 side-chains that close off these passages. These dimensions are very similar to the corresponding dimensions in TYMV. The residues restricting the pores in TYMV have slightly bulkier side-chains.

The protein subunit structure

The topology of the DYMV subunit is the canonical eight-stranded, anti-parallel β-barrel typical of $T = 3$ spherical plant viruses. All residues of subunits A and B were visible in electron density maps, while residues 1 to 13 of the C subunits were not. In TYMV, by contrast, the amino termini of subunits B and C were visible, while residues 1 to 26 of the A subunits were not. Similarly to TYMV, electron density maps of PhMV revealed all residues of subunits B and C but not the first nine residues of subunit A.

Figure 2(a) illustrates the similarity between the A, B and C subunits within DYMV and TYMV by superposition of backbone atoms C^{α}, C and N of the β-strands (a fit of 252 atoms). Likewise, side-by-side comparison of the β-barrel portions of DYMV and TYMV demonstrates their similarity. Major differences in subunit structures between DYMV and TYMV are confined to the amino-terminal ends. DYMV subunits have extended amino termini (residues 1 to 30) similar in many respects to the amino-terminal extensions found in satellite tobacco mosaic virus (Larson *et al.*, 1993a,b, 1998). The N termini of B and C subunits of TYMV, on the other hand, lie along the inner surface of their respective β-barrels, resulting in compact subunits. The B and C subunits of PhMV are very similar to those of TYMV but visible residues of the N-terminal polypeptide of subunit A are situated on the inside surface of the B subunit of a neighboring ABC trimer. Therefore, in three different tymoviruses there are three different dispositions of the N termini of the A subunits.

Given the unanticipated differences in dispositions of amino termini in the viruses, we re-examined difference electron density maps (Figure 5) to ascertain whether some portion of the DYMV subunit interactions might resemble those observed in TYMV, and *vice versa*. Indeed, weak electron density consistent with alternative dispositions was present in maps for both viruses. This suggests that either a minor fraction of DYMV virions utilize entirely the arrangement of amino-terminal ends that predominate in TYMV (and *vice versa*), or most particles of both viruses are a mixture of the two arrangements with one predominating in one virus, and the second in the other. The former seems most plausible, since once the disposition of

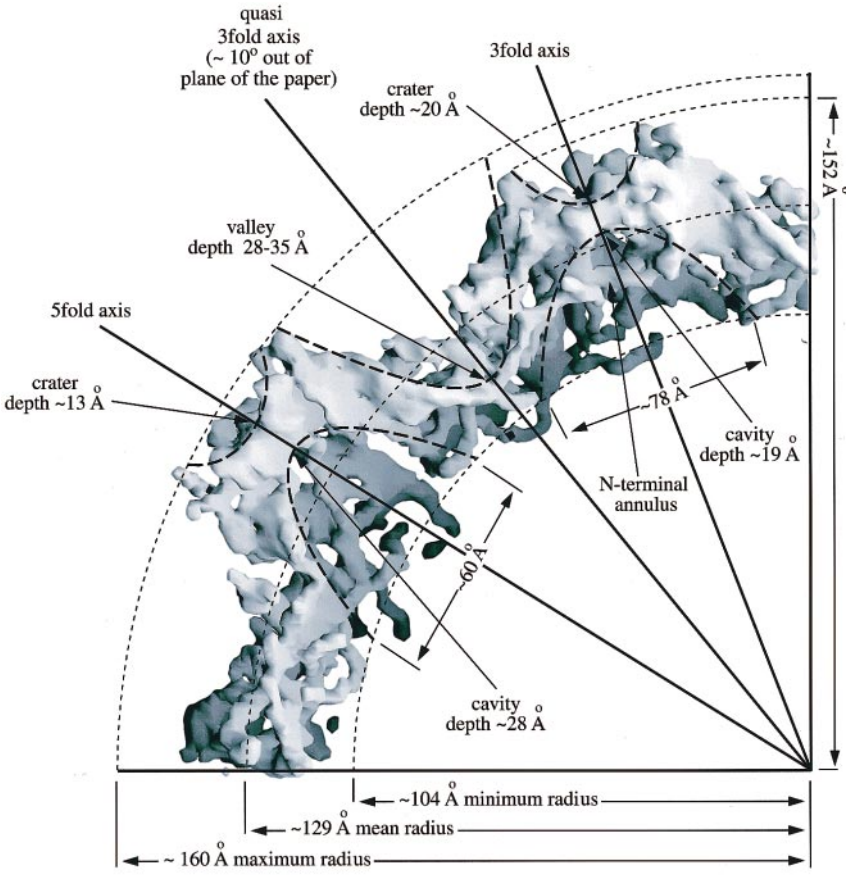


Figure 4. A sectional view of the DYMV viral capsid produced by the program GRASP using the atoms in a 40 Å slice of the capsid model. The topographical features are illustrated and labeled. The pores at the 5-fold and 3-fold axes, formed by Val122 side-chains, are ~0.25 Å in diameter. The pores at the quasi 3-fold axes, formed by Thr145 side-chains, have diameters of ~0.80 Å. The quasi 3-fold axis is out of the plane of the paper by 10°.

N termini (either according to DYMV or TYMV) is established, deviations from it would result in spatial conflicts.

Figure 6 is a diagram of the B subunit showing the sequence, the intramolecular hydrogen bonding and secondary structure. There are no intramolecular salt-bridges. There are two helical segments at residues 58 to 64 in the CD loop and 105 to 110 in the EF loop. The first has $\langle\phi\rangle$, $\langle\psi\rangle$ angles of -66° and -33°, respectively, suggestive of a regular α-helix, but the hydrogen-bonding scheme is not regular, having both $i \leftarrow i+3$ and $i \leftarrow i+4$ modes and missing some hydrogen bonds altogether. The second helix is a distorted 3_{10} helix with $\langle\phi\rangle$, $\langle\psi\rangle$ angles of -93° and -2°, respectively, and missing some hydrogen bonding along the segment. The β-strands in the CHEF sheet are regular, whereas the D and G strands in the BIDG sheet have β-bulges that disrupt the regular strand-to-strand hydrogen bonding. Occurrences of tight turns are listed in Table 1. These are characterized by $i \leftarrow i+3$ hydrogen bonding for type I or II turns.

There are two regions within the three icosahedrally distinct subunits of DYMV that are significantly different in conformation. The first region, defined by residues 23-27 of the quasi 3-fold loop, has a rmsd from the mean of corresponding ϕ and ψ angles of 59°, indicating a significant difference in the main-chain conformation between subunits A, B and C. Figure 7 shows this region to have the largest *B* values of any region in the molecules. These regions have no intersubunit contacts and only in subunit A is this region involved in RNA interactions. Therefore, the lack of packing constraints and the possible alternative N-terminal conformation explain the higher *B* values for this region. The RNA interactions with subunit A explain the lower *B* values for that subunit.

The second region is the FG loop. The type I turn exhibited in the FG loop of subunit B (Table 1) does not exist in subunits A and C. In fact, the rmsd from the mean of corresponding ϕ and ψ angles is 52°, with individual rms deviations >100°. This loop has the second-largest *B* values in the molecules (Figure 7). The FG loops form pores at the 5 and 3-fold axes through close contacts. The *B* values (especially for the A and C subunits) suggest that a somewhat disordered situation arises from the packing of these residues around the respective symmetry axes rather than a well-defined structure resulting from stable packing arrangements.

The mean *B* values of the A, B and C subunits are 28, 24 and 28 Å^2, respectively. Figure 7 illustrates the variation of *B* values with respect to residue number. Besides the two regions mentioned above, the N and C termini have elevated *B* values. Values for the C termini are explained by their extension into solvent on the exterior of the par-

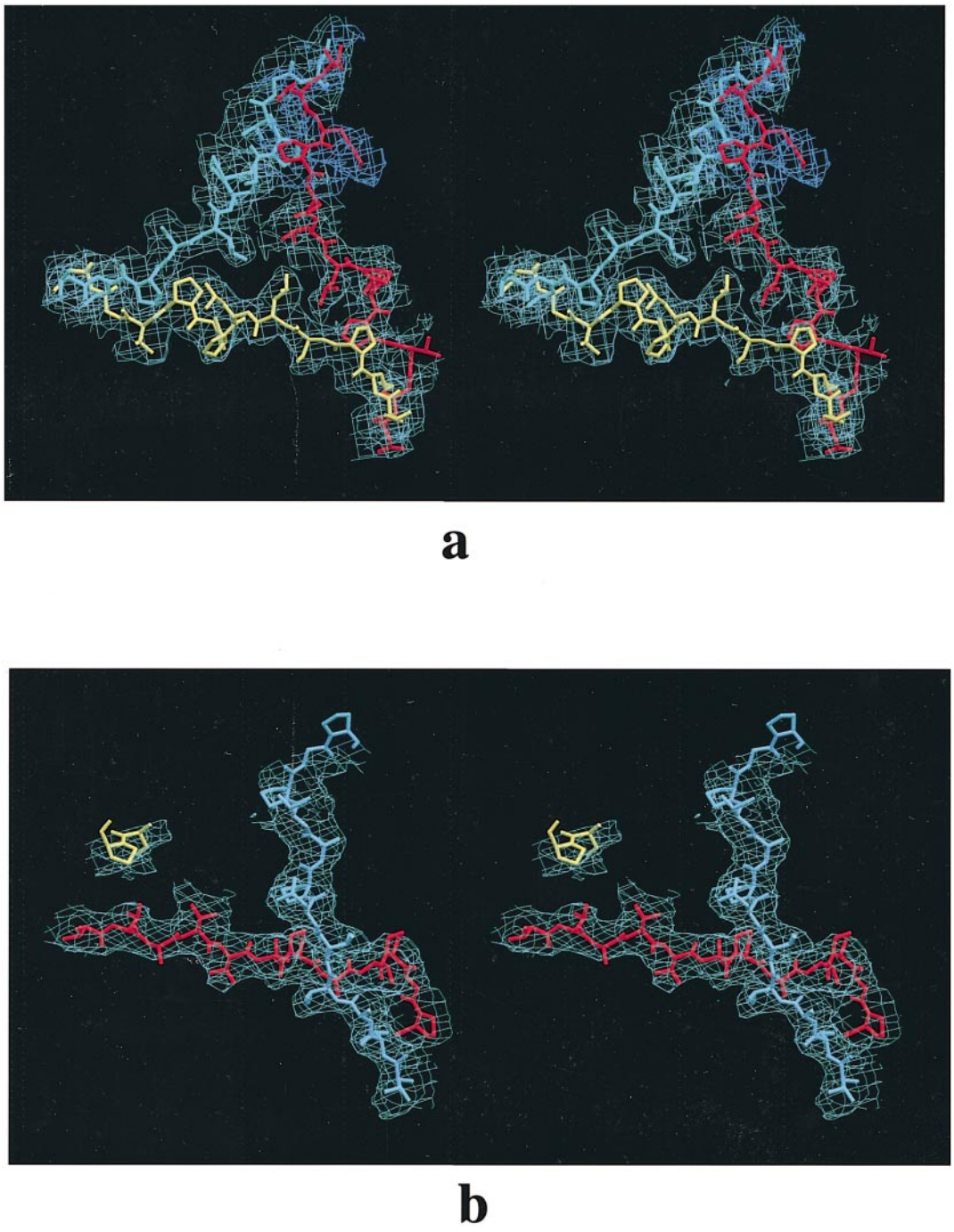


Figure 5. The N termini of (a) DYMV and (b) TYMV are viewed along the quasi 3-fold axis of the respective ABC trimers superimposed on the density of an $(F_o - F_c)$ exp $i\alpha_c$ omit map with residues 15 to 26 of all subunits omitted from the models. These maps demonstrate the validity of the differences seen in the dispositions of the N termini of the two viruses. The maps in (a) are contoured at 1.2σ (cyan) and 0.8σ (light blue). The map in (b) is contoured at 2σ.

ticle. *B* values for the N-terminal tails are not easily explained in the case of the A and B subunits, since they form a β-hexamer around the 3-fold axis. On the other hand, the N-terminal tail of the C subunit is disordered and unobserved in electron density maps; thus, the *B* values of the residues leading into the disordered region are extremely large. All four loops of the A subunit around the 5-fold axis have elevated *B* values when compared with adjacent residues. Of these, only the FG loop is involved in intersubunit contacts; the other loops form the crest of the 5-fold crater (Figure 4) and are, therefore, exposed to solvent. Perhaps the lower values exhibited by the same loops of the B and C subunits are due to interparticle contacts made by HI loops of these subunits, which account

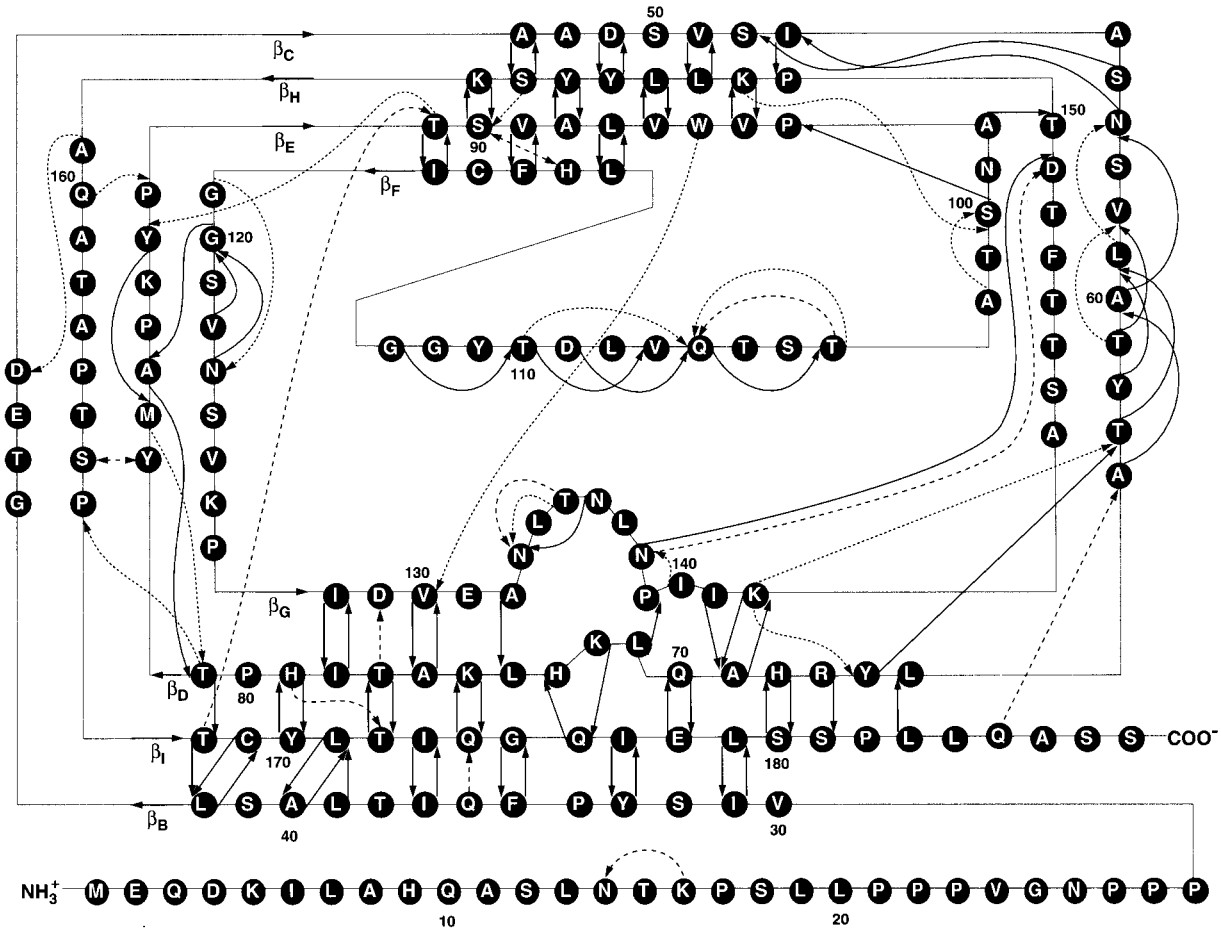


Figure 6. A diagram of the B subunit which, as noted in the legend to Figure 2, is essentially the same as the A and C subunits. Main-chain hydrogen bonds are shown as continuous lines. Hydrogen bonds between main-chain and side-chain atoms are shown as short broken lines and side-chain to side-chain hydrogen bonds are indicated with long broken lines.

for about 40 % of the loops. These contacts would reduce the effects of solvent exposure of the loops.

With regard to the disorder suggested by the regions of elevated *B* values, there were ten residues with discrete static disorder that could be modeled as alternative conformations of side-chains. These residues are the following: His9, Ser57, His73 of subunit A; Leu7, His9, Ser50, Lys72, His73 of subunit B; and Lys86, Gln185 of subunit C. In addition, entire residues Ser187 and Ser188 of the A and B subunits were modeled with two conformations as well as residue Ala186 of the B subunit. Occupancies for multiple conformers were determined by trial and error; *Q* values were adjusted by 0.05 (subject to the constraint $\Sigma Q_i = 1.0$) until the mean *B* values of the two conformers were approximately equal.

Water structure

Ordered solvent structure was modeled with 285 unique molecules of water. Their average temperature factor was 47 Å^2 with a range of 2-106 Å^2. Water molecules were assigned an occupancy of 1.0 unless placed in special positions (such as 2-fold axes) or were in conflict with partially occupied alternative conformations.

Table 2 contains an analysis of the ordered water molecules that can be classified as either structural or non-structural. Of the 285 water molecules per asymmetric unit, 223 are in contact only with protein atoms of a single capsid subunit, principally on the exterior, and six are in the second hydration shell. These 229 molecules constitute the non-structural water distribution.

As in STMV (Larson *et al.*, 1998), nearly 20 % of the ordered water molecules provide bridges between distant hydrogen-bonding groups. There are 27 water molecules in subunit interfaces of the ABC trimer, 65 water molecules in B-C interfaces of hexamers, and a total of 15 water molecules in A-A interfaces of pentamers.

RNA model

In the analysis of TYMV, although a general distribution of encapsidated RNA was evident in low-resolution (>5 Å) difference electron density maps,

Table 1. DYMV tight turns

Residues	Location	ϕ_2 (deg.)	ψ_2 (deg.)	ϕ_3 (deg.)	ψ_3 (deg.)	d(N → O) (Å)	$\langle d \rangle$ (Å)	Type
52-55	β_C-CD loop	−58	−40	−67	−16	3.08	3.07	I
S-I-A-S		−54	−39	−68	−12	2.95		
		−58	−44	−64	−15	3.17		
53-56	β_C-CD loop	−67	−16	−117	16	3.33	3.26	I
I-A-S-N		−68	−12	−112	−3	3.34		
		−64	−15	−104	3	3.10		
63-66	CD-loop-β_D	−54	−35	−79	−8	2.92	2.97	I
T-A-L-Y		−65	−25	−87	0	3.01		
		−56	−32	−86	−4	2.97		
81-84	β_D-DE loop	−66	−6	−95	−15	3.08	3.04	I
T-Y-M-A		−63	−7	−102	−5	2.98		
		−63	−6	−102	−9	3.06		
97-100	β_E-EF loop	−60	−27	−90	−16	3.02	3.01	I
P-A-N-S		−59	−23	−88	5	2.97		
		−60	−25	−85	4	3.03		
103-106	EF loop	−52	−25	−90	4	2.81	2.83	I
T-S-T-Q		−54	−28	−85	3	2.87		
		−54	−26	−85	3	2.80		
110-113	EF loop	−54	131	83	−6	2.99	2.90	II
T-Y-G-G		−51	120	98	−12	2.83		
		−56	121	95	−6	2.88		
								Not a tight
120-123	FG loop	175	−63	−105	−172	>3.50		turn
G-S-V-N		57	−60	−80	−25	3.28	3.28	I
								Not a tight
		−112	162	−52	142	>3.50		turn
133-136	β_G Bulge	−74	7	−100	−12	3.01	3.08	I
N-L-T-N		−76	10	−101	−25	3.18		
		−77	11	−107	−6	3.06		
Total no.						25	3.03	

Each tight turn grouping lists values for the A, B and C subunits in that order in the rows of the tight turn.

no discrete element of well-defined RNA was identified. In DYMV this was not the case. In particular, two fragments of RNA were seen and were modeled as generic seven-uridine and two-uridine fragments (designated RNA-A and RNA-B, respectively; see below). Figure 8 illustrates the RNA models superimposed on the electron density of an $F_o - F_c$ omit map with the respective fragments omitted. There was no density for the base ring of the first nucleotide nor for the ribose and base rings of the seventh nucleotide of RNA-A. The density for nucleotides 1-3 is particularly poor, as reflected in the temperature factors (Table 3). In a similar way there was no density for the base ring of the last nucleotide of the smaller fragment, RNA-B. The missing moieties are not included in the Figure. A trial-and-error procedure for determining the occupancy *versus* *B*-factor optimum was employed in which occupancies and *B* factors were incremented through a range and the optimal values taken as those that resulted in the minimum R_{free}. The resulting values of *Q* and *B* are given in Table 3.

Table 3 contains selected torsion angles for the RNA models. The geometry is generally random structure for both RNA fragments except, perhaps, for nucleotides 5 and 6 of RNA-A, which have torsion angles similar to those of helical RNA. In fact,

Table 2. Summary of water structure by interaction type

Interaction type	Number	Residue numbers
Second hydration sphere water molecules	6	2001-2006
First hydration sphere of subunit A	83	2101-2185
First hydration sphere of subunit B	71	2201-2271
First hydration sphere of subunit C	69	2301-2369
Bridging between subunit A and subunit B	20	2401-2420
Bridging between subunit A and subunit C	4	2501-2504
Bridging between subunit B and subunit C	21	2601-2621
Bridging between two A subunits	2	2147,2153
Bridging between two B subunits	0	-
Bridging between two C subunits	4	2701-2704
Bridging between protein and RNA	3	2801-2803
Bridging between subunit B and two C subunits	1	2901
Bridging between subunit B and two A subunits	1	2421

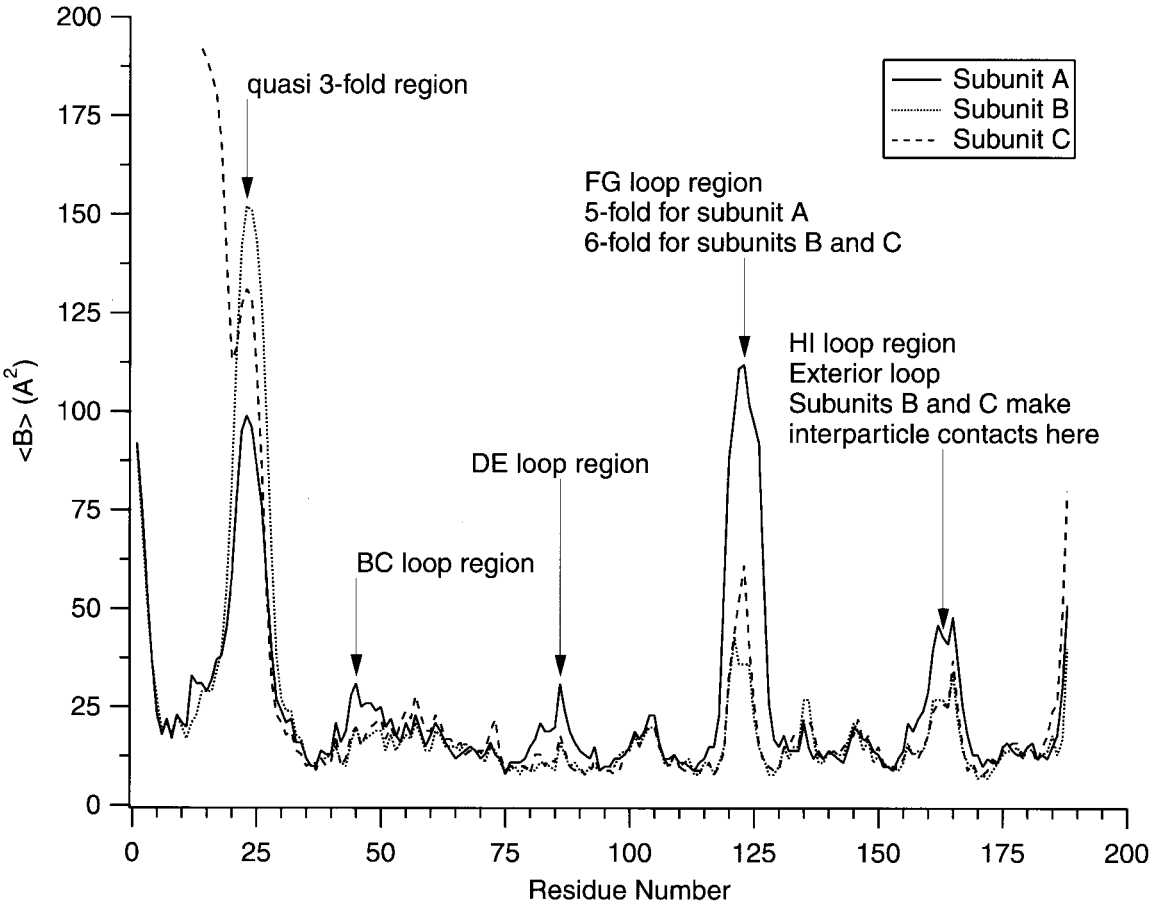


Figure 7. Graph of the mean main-chain temperature factors as a function of residue number illustrating the highly mobile or disordered regions of each subunit. It is clear that the N and C termini, the segments around the quasi 3-fold axes and the four loops (BC, DE, FG, and HI) at the top of the 5 and 3-fold protrusions exhibit the greatest mobility.

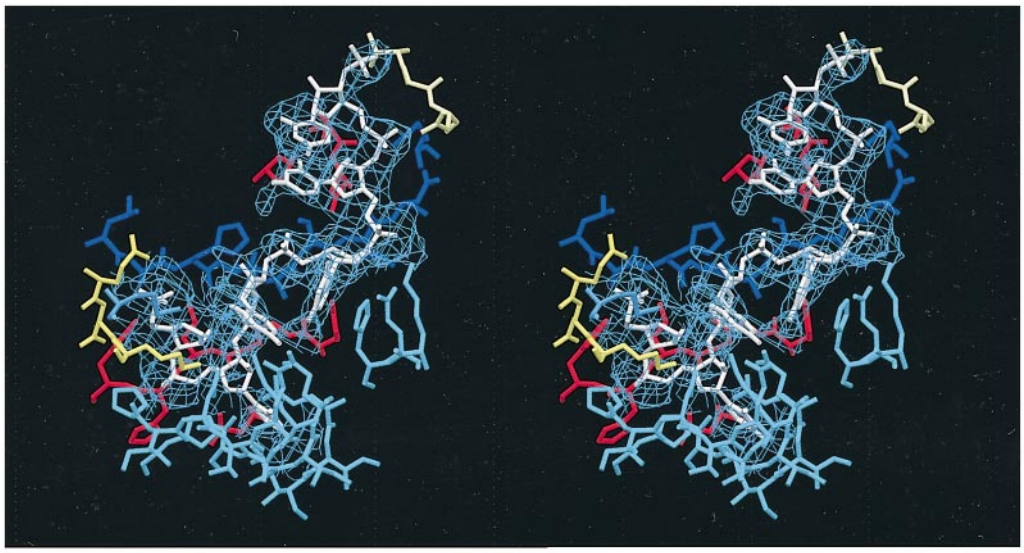
the original model for this strand was a fragment of one strand of the double-helical RNA from STMV that was readily placed into the observed density of nucleotides 5, 6 and 7. There is a kink at nucleotide 4 resulting in β and ε torsion angles that are very distinct from those of the other nucleotides. In general, the backbone torsion angles are not very uniform from nucleotide to nucleotide. In particular, α_2, β_1 and β_4, γ_2 and γ_6, and ε_4 deviate substantially from the mean values, and ζ angles are highly diverse.

Protein-protein interactions

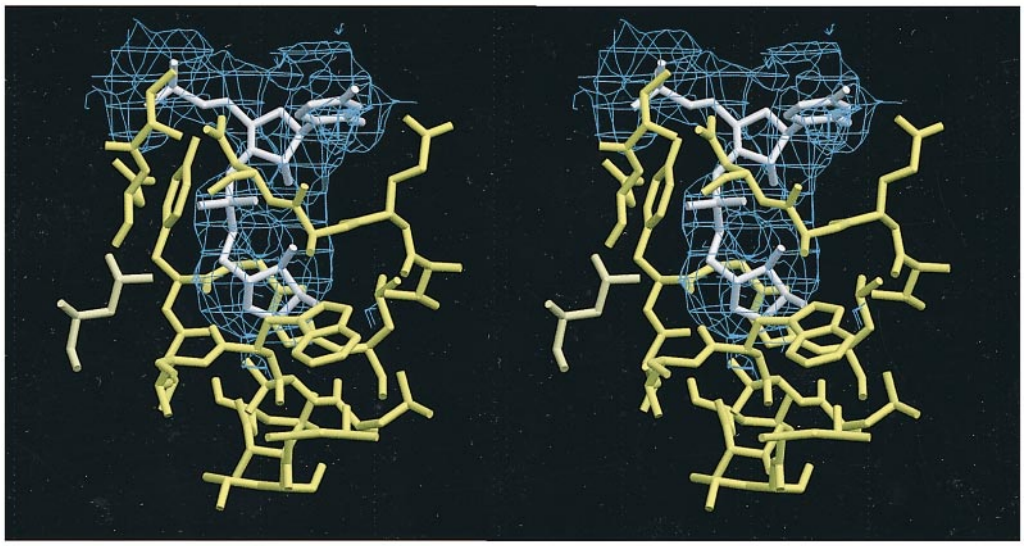
Table 4 lists the accessible surface area (ASA) of protein monomers based on conformers of highest occupancies calculated by the Lee & Richards (1971) algorithm. The ASA is dominated by hydrophobic residues (37%) and neutral hydrophilic residues (33%) with 12% of the ASA pertaining to aromatic residues and 17% to charged hydrophilic residues (similar to TYMV). Subunits of A have contact with nine other subunits (only five in TYMV), B subunits with 12 (only seven in TYMV) and C subunits have contact with nine other subunits (disregarding the invisible N-terminal residues) (only seven in TYMV). Total buried ASA by all contacting protein subunits of an A, B or C subunit is in the range 45-48%; 46-52% is buried when ordered RNA is included and 61-65% when ordered water molecules are included.

There are three major groupings of subunits forming the icosahedral shell of the virus: the ABC trimer (composed of one subunit each of the A, B and C subunits), the pentamer (composed of five A subunits), and the hexamer (composed of six alternating B and C subunits).

The ABC trimer, shown in Figure 3(e) and (f), forms a valley at the quasi 3-fold axis between the 5 and 3-fold protrusions. The N-terminal polypeptides cross near this axis, the A subunit terminus directed toward the B subunit, the B subunit terminus to the C subunit, and the C subunit terminus toward the A subunit in a manner similar to the A and B termini. This buries ~25% of the ASA of each participating subunit, an average of 2988 Å^2 per subunit. The total buried ASA is 8963 Å^2 for the group. This involves 67 hydrogen bonds and 535 intermolecular van der Waals contacts (contact



a



b

Figure 8. Stereoviews of the visible RNA fragments superimposed on density of $(F_o - F_c)$ exp $i\alpha_c$ omit maps (RNA omitted) with the protein environment shown. (a) The seven-nucleotide RNA fragment RNA-A superimposed on map density contoured at 1σ around nucleotides 1 and 2, and 2σ around nucleotides 3-7. (b) The two-nucleotide fragment RNA-B superimposed on density contoured at 1.5σ.

distance $\leqslant 4$ Å). There are 13 water molecules bridging A and B subunits, four water molecules between the A and C subunits and ten between the B and C subunits of an ABC trimer for an average of nine per subunit. TYMV, by comparison, has a total buried ASA of 4893 Å^2 upon formation of its trimer (1631 Å^2 per subunit) or only half that of DYMV, primarily due to lack of interaction of the N termini with neighboring trimer subunits.

Pentamer formation (Figure 3(a) and (b)) buries ~13% of ASA of each participating subunit and a total of 8235 Å^2 for the group (1647 Å^2 per subunit). A total of 40 hydrogen bonds and 740 intermolecular van der Waals contacts are made. There

Table 3. Summary of the occupancies, temperature factors and selected torsion angles in the RNA fragments

Nucleotide	*Q*	*B* PO_4 (Å^2)	*B* ribose (Å^2)	*B* base (Å^2)	χ (deg.)	α (deg.)	β (deg.)	γ (deg.)	δ (deg.)	ε (deg.)	ζ (deg.)
RNA-A											
1	0.65	135	195	na	na	−152	77	118	72	−161	−51
2	0.65	50	130	125	−154	−22	127	171	90	−160	−129
3	0.65	180	300	140	−147	−158	151	126	87	−156	39
4	0.65	105	95	105	−114	141	92	120	81	82	−92
5	0.65	85	40	100	−170	−155	162	129	78	−147	−56
6	0.65	45	80	100	−146	−100	180	75	89	−122	−4
7	0.65	105	na	na	na	na	na	na	na	na	na
RNA-B											
1	0.80	145	295	230	−119	−110	−166	0	74	177	−124
2	0.80	190	100	na	na	−156	−113	−3	80	na	na
Helical RNA torsion angles					−167	−81	183	52	81	−146	−75

na, not applicable.

are three bridging water molecules between adjacent A subunits of a pentamer, an average of three per subunit. TYMV has approximately the same total buried ASA of 8205 Å^2.

Hexamers (Figure 3(c) and (d)) are formed from alternating B and C subunits around 3-fold axes. Approximately 13 % of the ASA per subunit is buried for a total of 8961 Å^2 (1494 Å^2 per subunit). TYMV has almost 65 % more surface area buried in the formation of hexamers (14742 Å^2). A total of 45 hydrogen bonds and 591 intermolecular van der Waals contacts are made; 13 water molecules bridge B and C subunits for an average of 6.5 per subunit. In TYMV, N termini from the B and C subunits form an annulus that encircles the 3-fold axis (Canady *et al.*, 1996), while the N termini of the pentameric A subunits are invisible. In DYMV, amino acid residues 1 to 17 of B subunits replace those from C subunits (in TYMV), and amino acid residues 1 to 17 of B subunits are in turn replaced in DYMV by amino acid residues 1 to 17 of A subunits. Therefore, in DYMV, the annulus is again formed about the 3-fold axis, but by subunits that are not in the hexamer, as illustrated in Figure 9. The associated ABC trimer of each B and C subunit forming the hexamer provides one amino-terminal strand to the annulus; the A subunits provide strands from the trimers associated with the B subunits, and the B subunits provide strands from the C subunit trimers. In DYMV, the pentameric and hexameric capsomeres are, therefore, crosslinked by imposition of A and B subunit amino-terminal tails into the annulus of the hexamer from subunits next to the hexamer. The annulus, 10-11 Å in diameter, is about 6 Å inside the 3-fold pore. Each strand of the annulus forms three main-chain hydrogen bonds, as in a small, six-stranded β-barrel.

RNA-protein interactions

With both fragments of RNA, binding to protein through hydrogen bonds to either phosphate or ribose groups of the RNA appears not to be sequence-specific. The density seen for base rings fails to suggest any specific nucleotide sequence. The rings are directed away from the protein surface in general. Three putative water molecules could bridge the protein and the O2′ atoms of ribose moieties, one involved with RNA-A and two with RNA-B.

The RNA-A fragment, with respect to the A, B, and C subunits, is shown in Figure 10 and generally lies along the interface of BC dimers of a hexamer, near the N terminus of a second B subunit. The last nucleotide interacts with the N terminus of one A subunit near the quasi 3-fold axis of ABC trimers and the first interacts with the N terminus of a second A subunit near the annulus about the 3-fold axis. Therefore, five subunits (two A, two B, and one C) bind this fragment. This suggests that binding sites for RNA are either not available until protein subunits form the ABC trimer, or that the visible RNA elements serve as nucleation points to promote the aggregation of A, B and C subunits into a trimer arrangement. With the exception of some bases, most of the RNA strand is well defined, suggesting a fairly strict association with viral capsid, as illustrated in Figure 8.

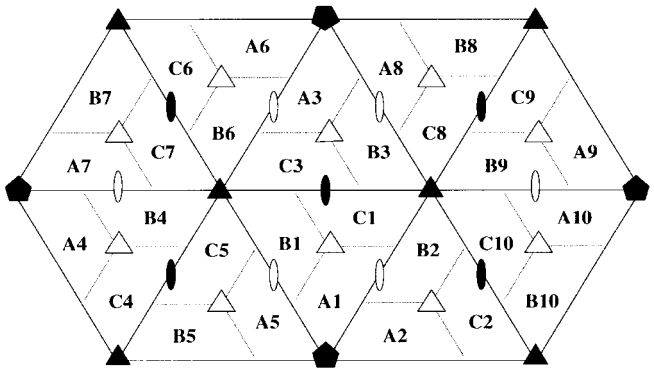
The RNA-B fragment is located inside the 5-fold cavity along the interface of neighboring A subunits, shown in Figure 10(a).

Particle packing

Virus particles make contacts only along unit cell body diagonals, which are coincident with icosahedral 3-fold axes. Therefore, only B and C subunits make interparticle interactions. There are 17 contacts less than 4 Å (possibly including one hydrogen bond between Ala161 N and Thr162 OG1, distances = 3.08 Å and 3.12 Å) involving residues Ala161, Thr162 and Thr165 of the HI loop between a B subunit of one particle and a C subunit of another particle, which results in a total of 102 (6 × 17) contacts between particles. These interactions account for only 630 Å^2 of buried surface area. Only one water molecule may bridge contact-

Table 4. Accessible and buried protein surface areas in DYMV and TYMV

Subunits neighboring the reference subunits A1, B1 and C1 are labeled according to the diagram. Exact rotation axes are identified by solid symbols and quasi rotation axes by open symbols. Using a 4 Å contact distance, subunit A1 is in contact with subunits A2, A5, B1, B2, B3, B5, C1, C3 and C5; subunit B1 is in contact with subunits A1, A2, A3, A5, B2, B3, B4, B5, B6, C1, C3 and C5; and, subunit C1 is in contact with subunits A1, A2, A3, B1, B2, B3, C3, C8 and C10.



	DYMV			TYMV		
	A1	B1	C1	A1	B1	C1
Accessible surface area (ASA)	12497	12484	10877	8407	10224	10381
Buried ASA						
2-fold (or quasi) (B2, A5, C3)	649	628	664	808	789	830
2-fold (or quasi) (C5, B3, A2)	64	147	61	0	0	0
quasi 3-fold (B1, C1, A1)	1769	1669	1156	569	1014	737
quasi 3-fold (C1, A1, B1)	1270	1829	1746	823	701	1049
5-fold or quasi 6-fold (A2, C5, B3)	785	739	714	836	1181	1217
5-fold or quasi 6-fold (A5, C3, B2)	832	732	729	805	1171	1147
3-fold (-, B4, C8)	-	7	26	-	45	65
3-fold (-, B6, C10)	-	17	23	-	32	56
quasi 6-fold (B3, A2, A3)	228	165	88	0	0	0
quasi 6-fold (B5, A3, -)	179	227	-	0	0	-
quasi 6-fold or 5-fold (C3, B2, -)	83	71	-	0	0	-
5-fold (-, B5, -)	-	55	-	-	0	-
Total buried ASA	5613	5987	4918	3811	4787	4937
Trimer formation	2877	3348	2738	1411	1698	1765
Pentamer formation	1647	-	-	1641	-	-
Hexamer formation	-	1459	1482	-	2376	2401
Total buried ASA by grouping						
Trimer		8963			4874	
Pentamer		8235			8205	
Hexamer		8823			14331	

ing subunits for a total of six intervening water molecules.

Implications for virion assembly and RNA packaging

The close structural similarity of DYMV and TYMV, essentially randomly chosen representatives, implies that most, if not all, of the tymoviruses have virtually identical capsid structures and that variations will be minor. On the other hand, the diversity in annulus formation seen in the two viruses, the ability of the amino-terminal ends to switch and swap interactions among subunits suggests a mode of structural versatility not previously suspected. Amino-terminal strand dispositions may derive simply from energetic con-

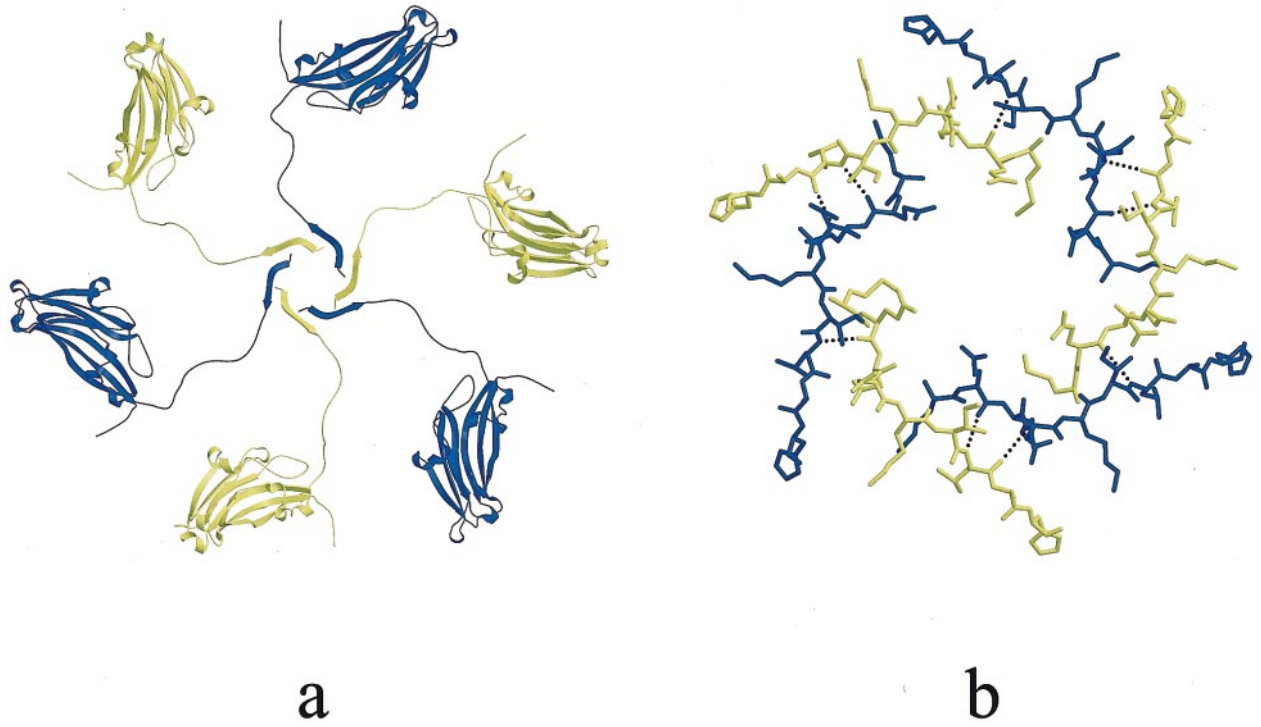


Figure 9. (a) The three A and three B subunits in DYMV that form the annulus around the 3-fold axis. (b) A close-up view of the annulus illustrating the three main-chain hydrogen bonds (shown as black broken lines) involving each strand of this six-stranded β-barrel-like structure. Residues 1 to 9 of the six N termini are shown.

siderations arising from differences in quaternary interactions ascribable to variations in amino acid sequence, or the explanation could be more elaborate. They could, for example, be determined by the amino acid compositions and resulting structures, and by the nature of the RNA (its sequence, secondary structure, or tertiary conformation) with which protein associates. RNA could specify local capsid conformation and, therefore, detailed virion architecture, as appears to be the case for satellite tobacco mosaic virus (Larson *et al*., 1998).

The appearance of RNA near the 60 quasi 3-fold axes (540 nucleotides in all) suggests some significant association between nucleic acid and protein. This was not self-evident, as most biophysical studies of other tymoviruses, in particular TYMV, suggested there to be little protein-nucleic acid interaction. This was implied as well by the occurrence for both TYMV (Huxley & Zubay, 1960) and DYMV (Walters & Scott, 1968, 1972) of ''empty heads'', which lacked RNA entirely.

The possibility of mixed modes of capsomere assembly and association in the tymoviruses may be pertinent to the process of decapsidation of nucleic acid upon infection. A popular idea is that RNA may be released from tymoviruses by displacement of a pentameric capsomere from the virion followed by ejection of the nucleic acid (Adrian *et al*., 1992; Keeling & Matthews, 1982; Katouzian-Safadi & Berthet-Colominas, 1983). Because pentameric capsomeres appeared to be significantly less firmly embedded in the capsid shell of TYMV than the hexameric capsomeres (Canady *et al*., 1996), this was an attractive hypothesis for TYMV. The more substantial interweaving and crosslinking of pentameric and hexameric capsomeres through interactions of the N termini of the A subunit would seem to make this a less appealing mechanism for DYMV. The possibility of mixed modes of construction with some small fraction of weakly held pentameric units (using the TYMV arrangement) embedded in a more regular, strongly fortified capsid structure (using the DYMV arrangement) might, however, provide an even better structural basis for an RNA escape mechanism. Weakly held units serve as pressure-release valves that allow nucleic acid to exit.

RNA segments bound at 60 icosahedrally equivalent points on the interior of the capsid further implies that the distribution of the condensed genomic RNA possesses at least some semblance of icosahedral symmetry. The sequences of the segments probably vary from quasi triad to triad, which accounts for the weak base density at some positions, but the RNA must, nonetheless, be structured in a pattern consistent with capsid symmetry.

In spite of the 540 (9/trimer × 60 trimers/particle) nucleotides seen bound to the interior of the capsid, interactions seem few in comparison with some other viruses, and this may have implications for RNA encapsidation and virus assembly. For

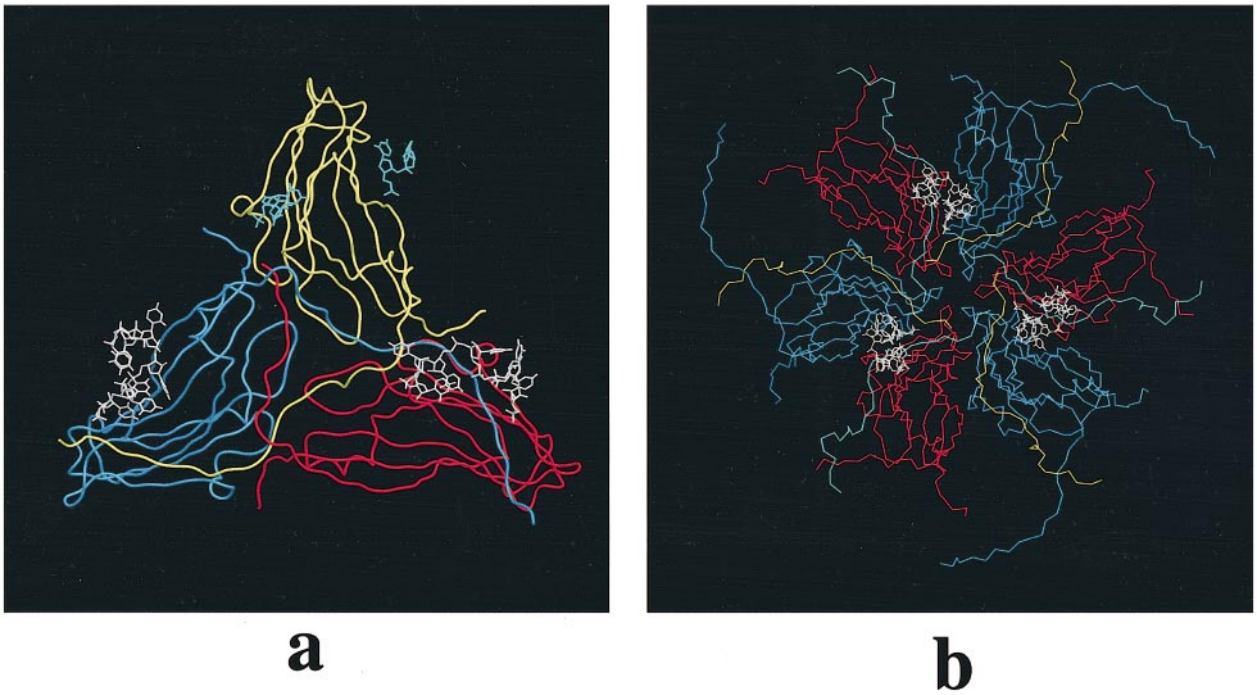


Figure 10. (a) The location of the RNA fragments with respect to the ABC trimer and (b), the location of RNA-A inside the hexameric cavity. RNA-A is shown as white, RNA-B as cyan. RNA-A lies along the inner surface between the annulus and the quasi 3-fold axis.

example, STMV has almost 60 % of its genomic RNA bound tightly by the capsid protein and almost certainly encapsidates its RNA by condensation of the nucleic acid with protein in a cooperative process. This seems unlikely in the case of the tymoviruses thus some other mechanism likely applies. The occurrence of empty heads suggests the same.

Two scenarios seem plausible. In one, the RNA is drawn into the nearly completed virion through a port, which is then closed by a capsomere hatch. A second possibility is that the RNA, through secondary and tertiary interactions, assumes a condensed form similar to its encapsidated conformation. The RNA may then present binding sites to capsid protein and serve as an ''assembly nucleus'' that initiates, directs, and/or catalyzes assembly. This seems a more attractive idea, as it might explain how some empty heads are formed, simply by loss of the nucleant during later assembly, growth of the virus capsid being driven principally by protein-protein interactions.

Materials and Methods

Production of DYMV

Great Northern shell beans (Henry Field, Shenandoah, Iowa) were grown in a climate-controlled greenhouse (10 °C-30 °C). An inoculum solution of DYMV (Desmodium yellow mottle tymovirus from American Type Tissue Culture, Rockville, MD, catalog number PV-155) was made by adding 0.5 ml of 0.05 M Tris-HCl (Fisher BP152-1, pH 7.2) to the desiccated material. When the plants were six to eight inches tall, the upper surface of one of the primary cotyledons of each plant was lightly dusted with carborundum (Fisher C192-500) and then inoculated by gentle rubbing using a cotton swab dipped in the DYMV inoculum. Plants were grown until flowering had just begun and then harvested and stored at −20 °C.

Isolation of DYMV

The preparation of DYMV was by a modification of the STMV procedure (Valverde & Dodds, 1987) and that for TYMV (Canady *et al.*, 1995). All procedures were carried out at 4 °C and all liquids, centrifuges and rotors were pre-chilled to 4 °C.

To 200 g of frozen tissue, 300 ml of 0.02 M Tris-HCl (pH 7.8) and 3 ml of 2-mercaptoethanol (Sigma) were combined and blended in a Waring blender, taking precautions that the tissue did not warm. The slurry was filtered through four layers of cheesecloth, and the solid material discarded. The filtrate was centrifuged at 10,500 ***g*** for 15 minutes, the liquid saved and the pellet discarded. The supernatant volume was measured and an equal volume of 1:1 (v/v) *n*-butanol/chloroform (Fisher) was added slowly with stirring. The mixture was centrifuged for 15 minutes at 10,500 ***g***. The aqueous phase was collected and centrifuged at 104,000 ***g*** for 1.5 hours in a Ti45 rotor. The supernatant was discarded and the pellet exposed to 6 ml of 0.02 M Tris-HCl (pH 7.8) overnight to soften. The pellet was resuspended and centrifuged for ten minutes at 3000 ***g***, and the pellet discarded. The supernatant was centrifuged at 186,000 ***g*** for 1.5 hours and the supernatant discarded. The pellet (clear, amber-tinted) was softened in 6 ml of 0.02 M Tris-

HCl (pH 7.8) with 0.04% (w/v) NaN_3 (Sigma) and resuspended. Centrifugation at 3000 *g* for ten minutes removed debris. Purified DYMV was stored at 4 °C. Coomassie blue staining following SDS-PAGE showed one band at 21,000 Da. The A_{260nm} to A_{280nm} ratio was 1.30. An estimate of virus concentration was made by assuming that a 1 mg/ml solution of virus would have an A_{260nm} of 7.0.

Crystallization

Purified DYMV at a concentration of 3 mg/ml was used to screen for crystals using Crystal Screen and Crystal Screen II (Hampton Research, Laguna Hills, CA). Trials were deployed using Cryschem Plates (Charles Supper Co., Natick, MA) covered with clear tape and maintained at room temperature. Each trial used a 0.5 ml reservoir and 3 µl each of the buffer and the DYMV concentrate. Small cuboidal crystals were observed after one month with 2.0 M sodium formate ($NaCHO_2$, Fluka), 0.1 M sodium acetate ($NaC_2H_3O_2$) at pH 4.6. By optimizing the conditions, large crystals, up to 0.7 mm on a side, were grown using either the sitting or the hanging-drop method (McPherson, 1999). The pH optimum was found to be 4.8 in both cases, while the optimum sodium formate concentration ranged from 1.8-2.5 M for the hanging-drop, or 2.0-2.2 M for the sitting-drop method. Crystals were first observed after one month but continued to grow up to six months. An average drop contained three to seven crystals. The cubic crystals were robust and easily manipulated with little damage from handling.

Diffraction studies

Data collection was carried out at the Stanford Synchrotron Radiation Laboratory using beamline 7-1 and a radiation wavelength at 1.08 Å. A MAR Research imaging plate detector system (X-ray Research GmbH, Hamburg, Germany), using a 300 mm imaging-plate diameter and a crystal-to-detector distance of 350 mm, was used. Data were collected using an exposure dose of 1300 counts per second and an oscillation angle of 0.5° for both native and derivative crystals. For larger native crystals, higher-resolution data were obtained using an oscillation angle of 0.25° and a dose exposure of 2500 counts per second. A total of 11 native crystals were used for both low and high resolution, with an average of five images per crystal.

Data reduction

Images were reduced using DENZO and SCALEPACK (Otwinowski & Minor, 1997). The images were processed with a mosaicity of 0.15. The crystals were of cubic space group $P4_232$ with $a = b = c = 348.5$ Å, with two particles in the unit cell. The asymmetric unit contains 15 subunits or 8.3% of the virus particle. The virions are centered in the unit cell at 0,0,0 and 1/2, 1/2, 1/2. A total of 574,527 observations were reduced to a unique data set of 178,705 reflections with $R_{sym} = 0.107$ and $\langle I/\sigma_I \rangle = 7.3$. The data set was 91.3% complete to 2.7 Å.

Structure solution and refinement

At the commencement of the structure determination, the sequence of the DYMV protein coat had not been determined. However, having determined the structure of TYMV (Canady *et al.*, 1996), we proceeded to use its refined model as the starting model for DYMV. Accordingly, the model was placed at the origin of the DYMV unit cell with three mutually perpendicular 2-fold axes of the icosahedral particle coincident with the cell axes. The initial correlation coefficient based on E^2 was 0.41 for 6-40 Å data where:

$$\text{correlation coefficient} = (\langle E_o^2 E_c^2 \rangle - \langle E_o^2 \rangle \langle E_c^2 \rangle)/(\langle E_o^4 - \langle E_o^2 \rangle^2 \rangle \langle E_c^4 - \langle E_c^2 \rangle^2 \rangle)^{1/2}$$

Rigid-body refinement against data to 4.5 Å followed by conjugate-gradient minimization (CGM) using XPLOR3.1 and XPLOR3.8 (Brünger *et al.*, 1987; Brünger, 1991, 1992) and extension of the resolution to 4.0-40 Å resulted in an R value of 0.299 ($R_{free} = 0.317$). At this point a $2F_o - F_c$ map revealed broken and confused density around the quasi 3-fold axis of the ABC trimer and a poor fit of residues 18 to 28 of the B and C subunits. Subsequently, amino acid residues 18 to 28 of each subunit of the TYMV model were omitted from the phase calculations until the correct sequence was obtained. Further refinement by simulated annealing (SA) and CGM gave $R = 0.258$ ($R_{free} = 0.278$) for 4.0-40 Å data.

Still without the sequence, model phases were extended to a resolution of 3.0 Å by molecular averaging and phase extension. Using this averaged map, the side-chains of tymovirus non-conserved residues were modified to fit the visible side-chain density. Portions of the model coinciding with choppy or missing density in the map had their occupancies set to zero. Refinement was continued by SA and CGM, utilizing bulk solvent corrections and extending the resolution to 2.7-40 Å with a 4 σ_F cutoff, resulting in $R = 0.240$ ($R_{free} = 0.253$). At this point, the sequence of the coat protein was provided by Dr Adrian Gibbs (personal communication) and the model was modified accordingly. Continued refinement with SA and CGM resulted in $R = 0.231$ ($R_{free} = 0.239$). Using successive omit maps, the omitted residues 18-28 of the ABC trimer were constructed to reflect the change of direction of the N termini of DYMV. With this essentially correct chain trace, the model refined to $R = 0.199$ ($R_{free} = 0.209$). From this point, successive cycles of model rebuilding, bulk solvent corrections, water modeling from difference maps, and refinement by SA and CGM produced a model with 267 water molecules and 14 residues with alternative conformations ($R = 0.153$; $R_{free} = 0.162$). Analysis of the water structure against difference omit maps resulted in the discovery of two RNA fragments and residues 14 to 17 of the N terminus of the C subunit that had been modeled as ordered water. Accordingly, polyuridine RNA models for seven and two-nucleotide fragments were built into the density with the removal of conflicting water molecules. Uridine nucleotides were used in these models simply because they were readily available. Admittedly, cytosine nucleotides would have been a better choice based on the RNA composition of 37% cytosine. However, nothing would really be gained here by converting the uridine to cytosine. Residues 14 to 17 of subunit C were also fitted into the density and conflicting water molecules eliminated. Additional water molecules were added to the model to bring the total to 285. The occupancy, Q, and group B values (the groups were base ring, ribose ring and PO_4^{2-}) for the RNA fragments were optimized by trial and error against R_{free}.

Computer programs

Geometrical parameter files used in the refinement were the following: a slightly modified PARHCSDX.PRO for protein (Engh & Huber, 1991), PARAM19.SOL for solvent, DNA-RNA-ALLH3NEW.PARAM for RNA, and PARAM.HBN for hydrogen bonding. The protein shell was constrained to icosahedral symmetry. The 5-fold averaging of electron density maps and phase extension were accomplished with the PHASES program suite of Furey (1990) using an atomic model mask. Model building was performed with the program O (Jones & Kjeldgaard, 1994). PROCHECK (Laskowski *et al.*, 1993) and XPLOR3.8 were used for evaluating the quality of the final and intermediate models. RAVE (Jones & Kjeldgaard, 1994) was used for averaging final low-resolution difference maps using a solid spherical mask. MAPMAN (Jones & Kjeldgaard, 1994) was used to manipulate maps for O and for RAVE. Figure 2(c) and (d) was produced with O. Figures 5, 8 and 10 were produced with SETOR (Evans, 1993). Figures 2(a) and (b), 3 and 9 were created with MOLSCRIPT (Kraulis, 1991) and rendered with RASTER3D (Bacon & Anderson, 1988; Merritt & Murphy, 1994). GRASP (Nicholls *et al.*, 1991) was used to produce Figure 4, and the program IGOR (WaveMetrics, Inc., 1998) for Figure 7. Figure 1 was produced using PROCHECK (Laskowski *et al.*, 1993).

RCSB Protein Data Bank accession numbers

Model coordinates and structure amplitudes have been submitted to the RCSB Protein Data Bank (code 1ddl).

Acknowledgments

This research was supported by a grant from the National Aeronautics and Space Administration. The authors thank Anne MacKenzie and Adrian Gibbs of the Australian National University for the sequencing of the RNA and, hence, the protein as well. We thank the San Diego Supercomputer Center for grants of time on the Cray C90 and T90 supercomputers.

References

Adrian, M., Timmins, P. A. & Witz, J. (1992). *In vitro* decapsidation of turnip yellow mosaic virus investigated by cryo-electron microscopy: a model for the decapsidation of a small isometric virus. *J. Gen. Virol.* **73**, 2079-2083.

Bacon, D. J. & Anderson, W. F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219-220.

Brünger, A. T. (1991). Simulated annealing in crystallography. *Annu. Rev. Phys. Chem.* **42**, 197-223.

Brünger, A. T. (1992). The free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature,* **355**, 427-474.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic *R* factor refinement by molecular dynamics. *Science,* **235**, 458-460.

Canady, M. A., Day, J. & McPherson, A. (1995). Preliminary X-ray diffraction analysis of crystals of turnip yellow mosaic virus (TYMV). *Proteins: Struct. Funct. Genet.* **21**, 78-81.

Canady, M. A., Larson, S. B., Day, J. & McPherson, A. (1996). Crystal structure of turnip yellow mosaic virus. *Nature Struct. Biol.* **3**, 771-781.

Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect. A,* **47**, 392-400.

Evans, S. V. (1993). SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graph.* **11**, 134-138.

Furey, W. (1990). PHASES. Abstract, *American Crystallographic Association, series 2*, **18**, 73.

Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. (1978). Tomato bushy stunt virus at 2.9 Å resolution. *Nature,* **276**, 368-373.

Hirth, L. & Givord, L. (1988). Tymoviruses. In *The Plant Viruses* (Koenig, R., ed.), chapter 3, pp. 163-212, Plenum, New York.

Huxley, H. E. & Zubay, G. (1960). The structure of the protein shell of turnip yellow mosaic virus. *J. Mol. Biol.* **2**, 189-196.

Jones, A. T. & Kjeldgaard, M. (1994). *O-The Manual, Version 5.10*, Uppsala University Press, Uppsala.

Katouzian-Safadi, M. & Berthet-Colominas, C. (1983). Evidence for the presence of a hole in the capsid of turnip yellow mosaic virus after RNA release by freezing and thawing. *Eur. J. Biochem.* **137**, 47-55.

Keeling, J. & Matthews, R. E. F. (1982). Mechanism for release of RNA from turnip yellow mosaic virus at high pH. *Virology,* **119**, 214-218.

Kleywegt, G. J. & Brünger, A. T. (1996). Checking your imagination: applications of the free *R* value. *Structure,* **4**, 897-904.

Kleywegt, G. J., Bergfors, T., Senn, H., LeMotte, P., Gsell, B., Shudo, K. & Jones, T. A. (1994). Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-*trans*-retinoic acid and a synthetic retinoid. *Structure,* **2**, 1241-1258.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Krishna, S. S., Hiremath, C. N., Munshi, S. K., Prahadeeswaran, D., Sastri, M., Savithri, H. S. & Murthy, M. R. N. (1999). Three-dimensional structure of Physalis mottle virus: implications for the viral assembly. *J. Mol. Biol.* **289**, 919-934.

Larson, S. B., Koszelak, S., Day, J., Greenwood, A., Dodds, J. A. & McPherson, A. (1993a). Double helical RNA in satellite tobacco mosaic virus. *Nature,* **361**, 179-182.

Larson, S. B., Koszelak, S., Day, J., Greenwood, A., Dodds, J. A. & McPherson, A. (1993b). Three dimensional structure of satellite tobacco mosaic virus at 2.9 Å resolution. *J. Mol. Biol.* **21**, 375-391.

Larson, S. B., Day, J., Greenwood, A. & McPherson, A. (1998). Refined structure of satellite tobacco mosaic virus at 1.8 Å resolution. *J. Mol. Biol.* **277**, 37-59.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.

Lee, B. & Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379-400.

Luzzati, V. (1952). Traitement statistique des erreurs dans la determination des structure cristallines. *Acta Crystallog.* **5**, 802-810.

Matthews, R. E. F. (1991). *Plant Virology*, pp. 231-237, Academic Press, New York.
McPherson, A. (1999). *Crystallization of Biological Macromolecules*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallog. sect. D,* **50**, 869-873.
Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association-insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281-296.
Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Macromolecular Crystallography* (Carter, J. W., Jr & Sweet, R. M., eds), pp. 307-326, Academic Press, San Diego.
Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **23**, 283-438.
Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallog. sect. A,* **42**, 140-149.
Scott, H. A. & Moore, B. J. (1972). A comparison of certain properties of Desmodium yellow mottle and turnip yellow mosaic viruses. *Virology,* **50**, 613-614.
Speir, J. A., Munshi, S., Wang, G., Baker, T. S. & Johnson, J. E. (1995). Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryoelectron microscopy. *Structure,* **3**, 423-431.
Valverde, R. A. & Dodds, J. A. (1987). Some properties of isometric virus particles which contain the satellite RNA of TMV. *J. Gen. Virol.* **68**, 965-972.
Walters, H. J. & Scott, H. A. (1968). A virus isolate from Desmodium related to turnip yellow mosaic. *Phytopathology,* **58**, 1071.
Walters, H. J. & Scott, H. A. (1972). Host range and some properties of Desmodium yellow mottle virus. *Phytopathology,* **62**, 125-128.

Edited by T. Richmond

(Received 5 November 1999; received in revised form 30 March 2000; accepted 21 June 2000)