Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy

Jeffrey A Speir, Sanjeev Munshi, Guoji Wang, Timothy S Baker and John E Johnson*

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907–1392, USA

Background: RNA–protein interactions stabilize many viruses and also the nucleoprotein cores of enveloped animal viruses (e.g. retroviruses). The nucleoprotein particles are frequently pleomorphic and generally unstable due to the lack of strong protein–protein interactions in their capsids. Principles governing their structures are unknown because crystals of such nucleoprotein particles that diffract to high resolution have not previously been produced. Cowpea chlorotic mottle virions (CCMV) are typical of particles stabilized by RNA–protein interactions and it has been found that crystals that diffract beyond 4.5 Å resolution are difficult to grow. However, we report here the purification of CCMV with an exceptionally mild procedure and the growth of crystals that diffract X-rays to 3.2 Å resolution.

Results: The 3.2 Å X-ray structure of native CCMV, an icosahedral (T=3) RNA plant virus, shows novel quaternary structure interactions based on interwoven carboxy-terminal polypeptides that extend from canonical capsid β-barrel subunits. Additional particle stability is provided by intercapsomere contacts between metal ion mediated carboxyl cages and by protein interactions with regions of ordered RNA. The structure of a metal-free, swollen form of the virus was determined by cryo-electron microscopy and image reconstruction. Modeling of this structure with the X-ray coordinates of the native subunits shows that the 29 Å radial expansion is due to electrostatic repulsion at the carboxyl cages and is stopped short of complete disassembly by preservation of interwoven carboxyl termini and protein–RNA contacts.

Conclusions: The CCMV capsid displays quaternary structural interactions that are unique compared with previously determined RNA virus structures. The loosely coupled hexamer and pentamer morphological units readily explain their versatile reassembly properties and the pH and metal ion dependent polymorphism observed in the virions. Association of capsomeres through inter-penetrating carboxy-terminal portions of the subunit polypeptides has been previously described only for the DNA tumor viruses, SV40 and polyoma.

Structure 15 January 1995, **3**:63–78

Key words: *bromoviridae*, cowpea chlorotic mottle virus, quasi-equivalence, T=3, virus structure, X-ray crystallography

Introduction

Many intensively studied viruses and nucleoprotein cores of enveloped animal viruses (for example, retroviruses) are unstable and pleomorphic. The principles governing their structures are generally unknown due to the lack of high resolution studies. These particles are stabilized predominantly by RNA–protein interactions. They are unstable when exposed to even mild physical denaturants (such as ultracentrifugation) or chemical denaturants. It is likely that particle heterogeneity, inadvertently induced during preparation of these viruses, has prevented the formation of well-ordered crystals for study at high resolution by X-ray crystallography. Viruses of this type are, however, among the most accessible for assembly studies because subunit tertiary structure is not disturbed when the virion disassembles. Cowpea chlorotic mottle virions (CCMV) are typical of the particles described above, and when purified with an exceptionally mild procedure, can form crystals that diffract X-rays to 3.2 Å resolution [1].

CCMV is a member of the bromovirus group of the *Bromoviridae* family. Tripartite genomes, encapsulated in polyhedral or bacilliform particles, distinguish the 31 member *Bromoviridae* family from other plant virus taxa [2]. The CCMV genome consists of three unique, single-stranded, positive-sense RNA molecules which are encapsidated separately. The capsid is composed of 180 chemically identical protein subunits (19 800 Da; 190 amino acids) which form a 286 Å diameter icosahedral shell with T=3 quasi-symmetry [3–6]. Image processing of negatively-stained arrays of native CCMV particles showed the presence of discrete hexamers (with perfect six-fold symmetry to the resolution of the image) and pentamers, with rather tenuous connections between these morphological units in the T=3 lattice [5,6].

CCMV was the first icosahedral virus reassembled *in vitro* from purified viral protein and RNA to form infectious particles [7] and its assembly has been the subject of biochemical and biophysical investigations for over 25 years.

*Corresponding author.

Native CCMV is stable around pH 5.0 and sediments at 88S. At pH 7.0 and low ionic strength (μ<0.1) the particles undergo a concerted transition to a species with a sharp, 78S peak in sedimentation velocity centrifugation analysis [8]. These particles are a swollen form of the virus where the average size of the particle has increased by roughly 10%. This form has the same protein to RNA ratio as native virus and the transition is reversed if the pH is lowered to 5.0, or if the Ca^{2+} or Mg^{2+} concentration is raised to 0.05 M. Increasing the pH to 7.5 and the ionic strength (μ>0.4) results in particle disassembly to a less discrete, ~40S ribonucleoprotein species and dimeric protein subunits [9–12]. Under well-defined conditions of pH, ionic strength, and divalent metal ion concentration, CCMV capsid protein or capsid protein and RNA reassemble to form T=1, T=3, swollen T=3, and T=7 icosahedral particles, multi-shelled T=3 and T=7 particles, sheets, tubes, rosettes, and a variety of laminar structures [8,9,13,14].

While first observed for bromegrass mosaic bromovirus [15], discrete swollen states controlled by pH and metal ion concentration have been observed in many other plant viruses [8,16–18] and the structures of tomato bushy stunt virus (TBSV) in the native and swollen states have been determined by X-ray crystallography at 2.9 Å [19,20] and 8 Å [21], respectively. It has been suggested that, for those plant viruses that swell, the expanded form is an intermediate in the disassembly and infection process. Similar capsid changes occur in animal viruses as demonstrated by the compelling evidence that a structural transition occurs in the poliovirus capsid when it binds to its cognate receptor [22] and it has been suggested that subunit movements in the picornavirus may be similar to those observed in TBSV [23].

In spite of the tendency for TBSV, turnip crinkle virus (TCV) and southern bean mosaic virus (SBMV) to form swollen T=3 particles that are similar to those observed for bromoviruses, there is no evidence for the formation of sheets, tubes and other polymorphic protein assemblies described for CCMV. The high resolution X-ray structures of TBSV [19], TCV [24] and SBMV [25,26] showed extensive regions of inter-subunit association and, in contrast to the CCMV structure solved by electron microscopy, that the subunits forming hexamers are a trimer of dimers that do not have six-fold symmetry. The T=3 insect nodaviruses [27–29] have a particle organization strikingly similar to TBSV, TCV and SBMV. In contrast to CCMV, all the T=3 and T=4 [30] viruses previously studied by crystallography form robust particles that are stabilized by extensive protein–protein interactions and that can withstand severe purification procedures and form well-ordered crystals diffracting beyond 3.0 Å resolution. These data suggest that the structure of bromovirus capsids are different from those observed in other T=3 plant and insect viruses. We now report the 3.2 Å structure of native CCMV and a detailed fit of the subunit atomic coordinates to the 28 Å electron density map of swollen CCMV, determined by cryo-electron microscopy (cryoEM) and image reconstruction.

Results

CCMV dimensions and morphology

The CCMV structure, displayed both as a Cα coordinate model and as a diagram in Fig. 1, is the first determined for a member of the *Bromoviridae*. The capsid is formed of distinct, protruding pentameric and hexameric morphological units (referred to below as pentamers and hexamers, respectively, or collectively as capsomeres). The outer capsid surface extends to a maximum radius of 142 Å at the pentamers and to a minimum of 120 Å at the two-fold symmetry axes. The interior radius of the protein shell averages 104 Å, but portions of ordered protein at the icosahedral three-fold symmetry axes penetrate to radii of 95 Å. The morphology of the virion is that of a truncated icosahedron (Figs 1a and 1b) with three chemically identical gene products, labeled A, B, and C, in each icosahedral asymmetric unit. All other T=3 plant and insect viruses whose structures have been determined at high resolution adopt geometry closer to that of the rhombic triacontahedron (Fig. 1c).

Subunit fold and orientation

All CCMV subunits adopt the canonical virus β-barrel fold, with residues 27–49 and 179–190 extending from the barrel at the amino termini and carboxyl termini and playing major roles in the formation of particle quaternary structure (Fig. 2a). Subunits from only one other class of virus, the DNA tumor papovaviruses (polyoma and SV40) (Fig. 2b), have similar extensions at the carboxyl termini. These extensions play a major role in forming the viral quaternary structure in papovaviruses as they do in CCMV [31,32]. Residues 1–26 in CCMV are not visible in the electron density map. These are predominately basic residues (nine arginine plus lysine residues; Fig. 3), which are internal, and interact with RNA [33,34].

The β-strands of each subunit run nearly parallel to the nearest five-fold or three-fold (quasi six-fold) icosahedral rotation axis. Thus the β-barrels appear to 'stand on end' relative to the curvature of the capsid and this allows for tight clustering of the subunits into distinct capsomeres. The intracapsomere contacts are mostly hydrophobic and involve the EF insert (referring to the amino acids between two strands of a β-sheet as the insert between those strands; Figs 2a and 3) and βH strand of one subunit and the FG insert, DE insert, βD and βI strands of the adjacent subunit. Several of these interactions display quasi symmetry in making similar interactions at A-A5, B2-C, and C-B5 subunit interfaces (Figs 1a and 1b; see discussion).

Hexamer morphological units

The B and C subunits cluster around the icosahedral three-fold (quasi six-fold) axes and form an arrangement with nearly perfect six-fold symmetry (rms deviation for

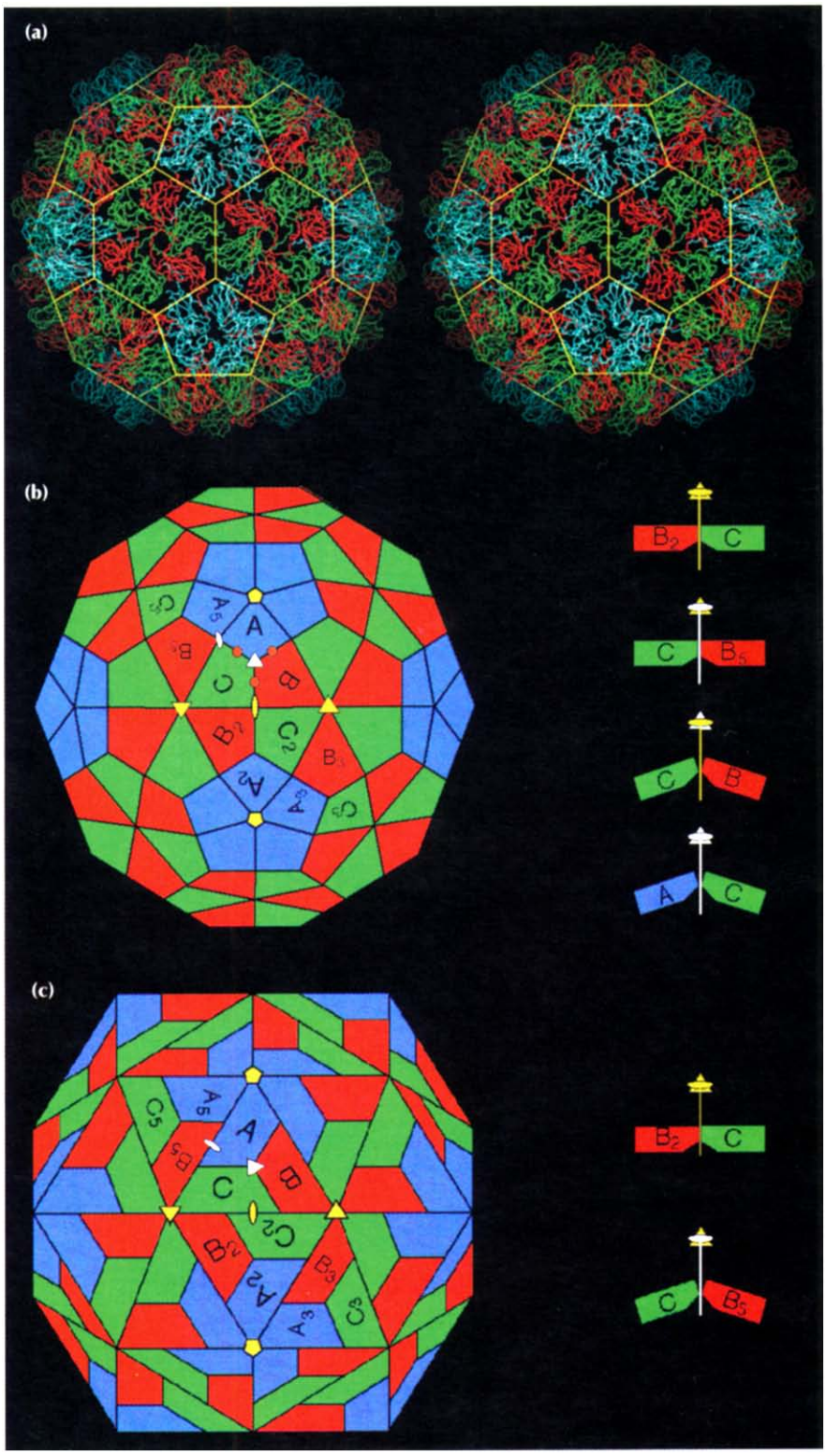

Fig. 1. Structure of the CCMV capsid and its novel geometrical features. **(a)** Stereoview of the protein shell as a Cα tracing. Color coding is defined in (b). The yellow cage represents the edges of a truncated icosahedron. **(b)** A truncated icosahedron model [38] shown in the same orientation as in (a). Positions of icosahedral rotation axes are marked by yellow symbols (pentagons: five-fold rotation axes; triangles: three-fold rotation axes; ovals: two-fold rotation axes). The central triangle with one five-fold (top) and two three-fold axes (lower left and right) at its vertices and containing polygons labeled A, B and C defines the icosahedral asymmetric unit. The polygons represent chemically identical protein subunits and within this area they occupy slightly different geometrical (chemical) environments and this is indicated by differences in their coloring. Polygons with subscripts are related to A, B and C by icosahedral symmetry (i.e. A to A_5 by five-fold rotation). The apparent three-fold rotation axis at the center of the asymmetric unit [vertices of cage in (a)] is not exact (quasi three-fold axis: white triangle) as it relates icosahedral three-fold axes (quasi six-fold axes) to a five-fold axis outside of its local environment. Similarly, polygons labeled A and B_5 form a quasi two-fold axis (white oval). Putative calcium-binding locations in one asymmetric unit of CCMV are marked by brown circles. Interactions between B_2-C and between C-B_5 polygons are defined by 180° dihedral angles (side view at top right) whereas bends similar in magnitude occur at the B-C and C-A polygon interfaces (138° and 142°, respectively) (side view at bottom right). **(c)** A rhombic triacontahedron model [38] shown in the same orientation and labeled identically to the truncated icosahedron model. The A, B and C polygons are co-planar within each asymmetric unit. Two such asymmetric units are co-planar by icosahedral two-fold symmetry giving the rhombic solid its prominent diamond shaped facets. The interaction between B_2 and C remains planar (a 180° dihedral angle between two-fold related asymmetric units, top right) whereas the interaction of C and B_5 adopts a 144° dihedral angle (bottom right). Hexamers are therefore best described as trimers of dimers.

the Cα protein atoms in residues 27–179 when rotated by 60° is 0.571 Å). Residues 27–190 were visible in each of these subunits (Figs 2 and 3). The six amino-terminal arms (residues 27–49) of the B and C subunits, which converge at the quasi six-fold axes, are clearly the major determinant of hexameric capsomere formation. The arms intertwine to form a hexameric tubular structure, (the β-hexamer, Fig. 4) made up of six, short parallel β-strands (residues 29–33) which close a channel running along the center of the quasi six-fold axes between the interior and exterior of the virus. The side-chain portions of residues Gln29, Val31 and Val33 fill the inner volume of the cylinder by contacting six-fold related residues in the adjacent β-strands. Side-chain oxygens of Gln29 residues hydrogen bond with the main-chain nitrogens of adjacent Gln29 residues, making a circular

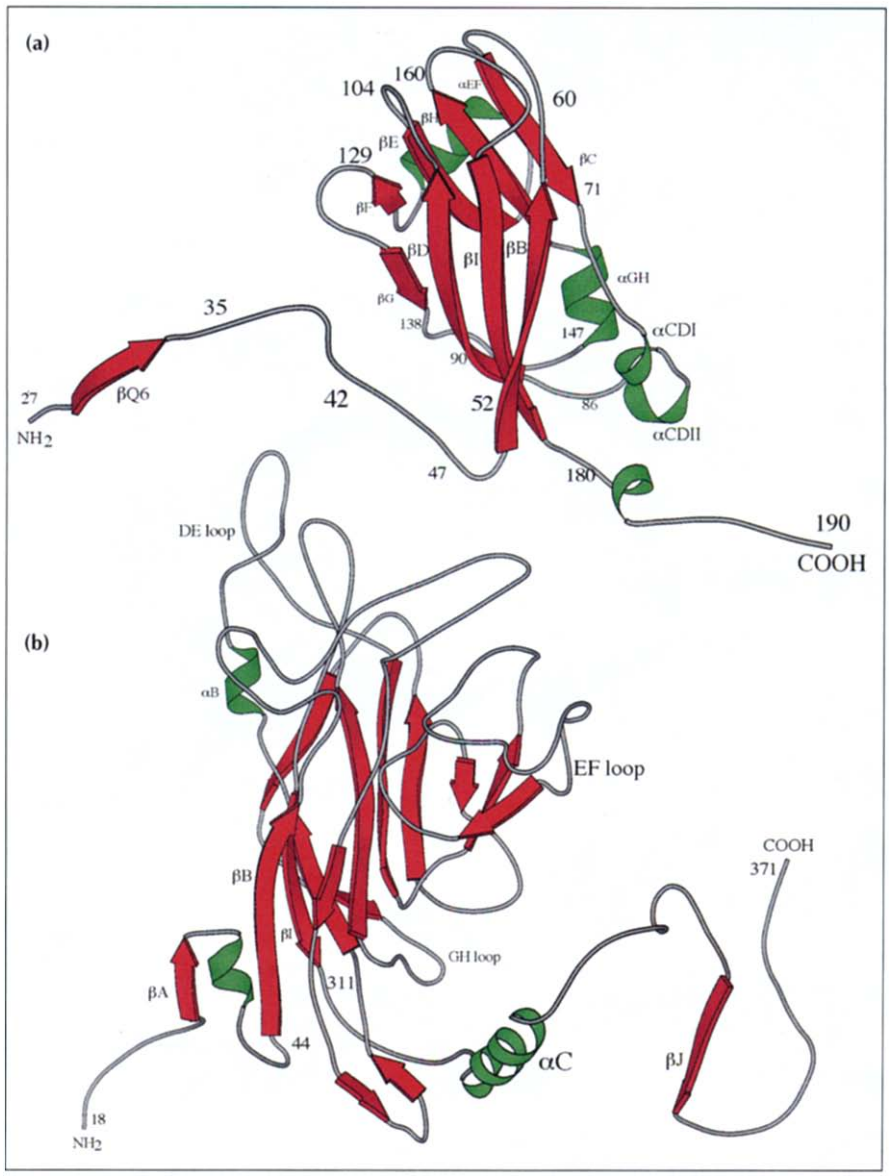

Fig. 2. Ribbon diagrams showing the tertiary structure of the **(a)** CCMV and **(b)** polyoma virus VP1 capsid subunits (VP1 drawn at smaller scale). Virus exteriors are at the top of each diagram. Selected residues and secondary structure elements are labeled. (Polyoma virus coordinates were kindly provided by Thilo Stehle and Steven Harrison, Harvard University.)

ring of interactions. The valine residues stack upon one another inside the β-tube forming a circle of hydrophobic bonds. This hydrophobic environment is surrounded by the regular hydrogen-bonding network of the β-strands and separated from the particle interior by the Gln29 side-chain atoms.

Pentamer morphological units

The A subunits cluster as pentameric capsomeres at the five-fold symmetry axes. The 41 amino-terminal residues are not visible in the electron density. Lys42 makes the closest approach to the icosahedral five-fold axis, but it does not interact with the symmetry-related lysines. The difference in the number of observable residues in the A subunits, compared with the B and C subunits, does not reflect any difference in the actual primary structures, but rather differences in the portions of the polypeptides that obey icosahedral symmetry. The difference in local environments does not affect the tertiary structures of the individual subunits, which have virtually identical β-barrels (rms deviation for the superposition of all atoms in residues 52–179 of the A, B and C subunits is 0.735 Å).

Dimer interactions

The primary contact between capsomeres is facilitated by the extended carboxy-terminal regions of the polypeptides (residues 181–190) which run nearly tangential to the virus surface. The carboxy-terminal arms of all 180 subunits point outward from their capsomeres of origin and across icosahedral two-fold or quasi two-fold axes to interact with subunits in adjacent capsomeres. The amino-terminal arm of the 'invaded', two-fold related subunit, clamps the interpenetrating carboxy-terminal arm between itself and the invaded β-barrel. The clamping is reciprocated across the two-folds and forms the major (dimeric) interaction between subunits ($A\text{-}B_5$ and $C\text{-}C_2$, see Fig. 1b) in the capsid. This interaction most probably corresponds with the assembling unit of the virus, because disassembled capsid protein is a soluble

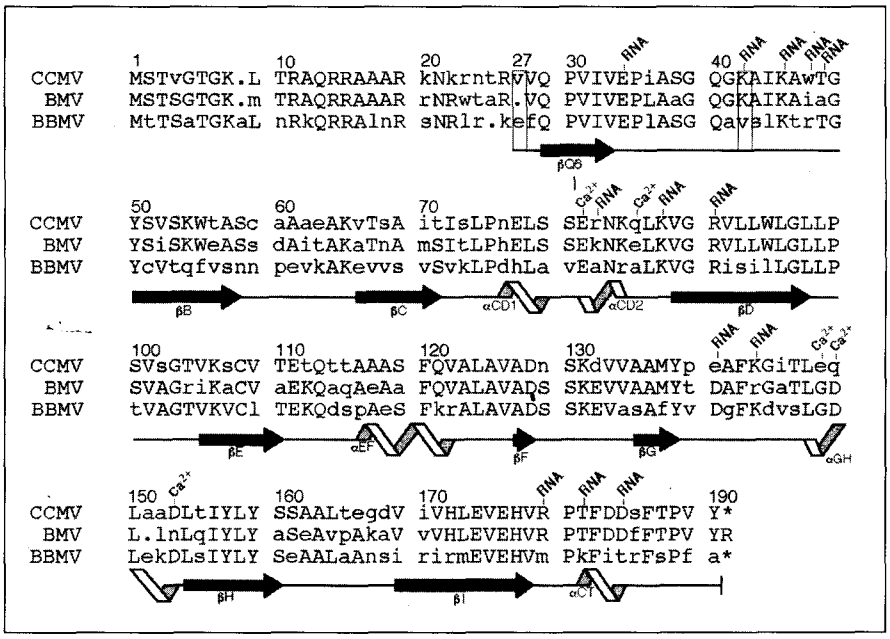

Fig. 3. Sequence alignment of the bromovirus group capsid proteins and designation of the corresponding polypeptide secondary structure (CCMV, cowpea chlorotic mottle virus; BMV, bromegrass mosaic virus; BBMV, broad bean mosaic virus) [4,57,58]. Residue numbering at the top of each sequence block is for the CCMV protein with the left-most digit over the column to which the number corresponds. All other labels are based upon analysis of the high resolution CCMV capsid structure. Upper-case letters represent residues which are identical in at least two of the three sequences. The capsid proteins of CCMV and BMV share 70% identity whereas CCMV and BBMV are only 48% identical. Depending on the local capsid environment, ordered electron density for residues 27 (first box) to 190 (B and C subunits) or for residues 42 (second box) to 190 (A subunits) is observed. Residues which are involved in RNA binding and calcium ion coordination (Ca^{2+}) are labeled. Secondary-structure elements determined by the program PROCHECK [59] are represented at the bottom of each sequence block. Extended black arrows and twisting ribbons represent β-sheet and α-helix secondary structure, respectively. These symbols are placed under the residues to which the represented structure has been assigned, and each labeled according to their location in the CCMV protein tertiary fold (Fig. 2a). Solid lines between β-sheets and α-helices have no regular secondary structure assigned. The dots indicate inserted gaps for purposes of alignment and the asterisks indicate that the equivalent terminal residues are not present in CCMV and BBMV.

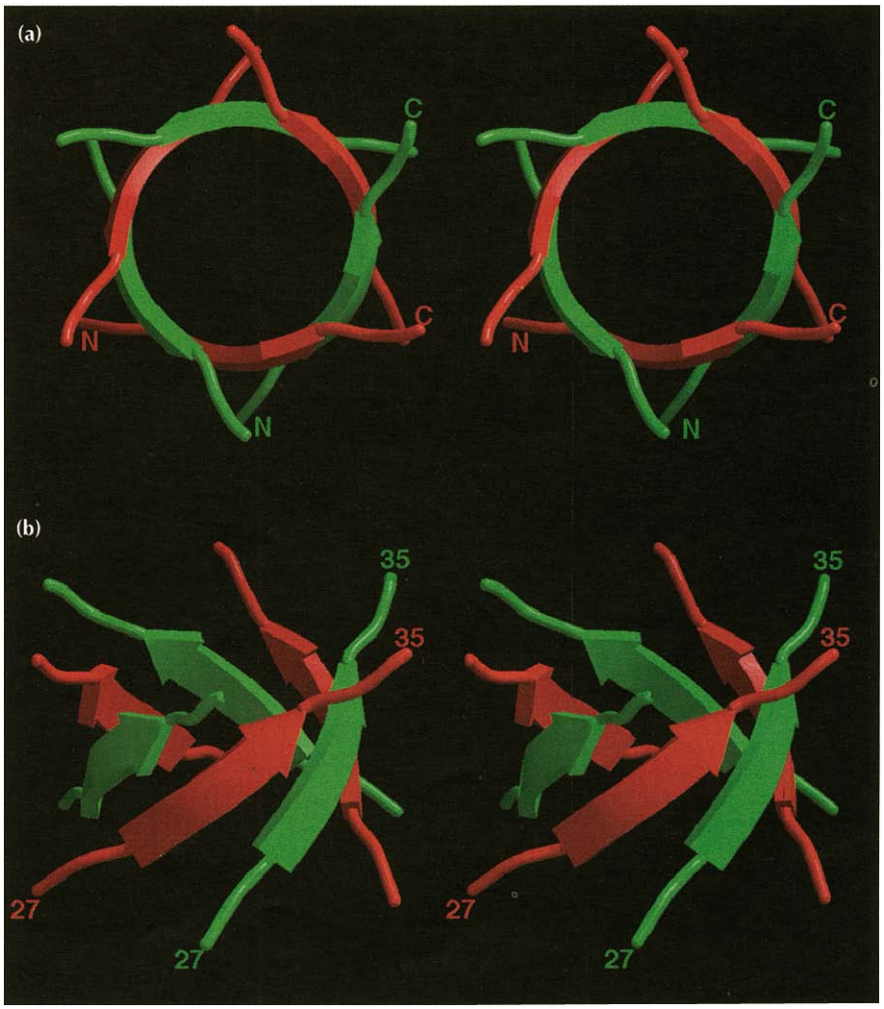

Fig. 4. Stereoviews of the β-hexamer shown as ribbon drawings that correspond to residues 27–35 of the B (red) and C (green) subunit amino termini. **(a)** The β-hexamer viewed directly down the quasi six-fold axis towards the particle interior. The amino and carboxyl termini are labeled for two of the six polypeptide renditions. Note that the strands labeled are not related by icosahedral symmetry. **(b)** The β-hexamer viewed approximately tangential to the virus surface. Residue numbers are given for the same polypeptides labeled above.

dimer [9,10]. Residues 36–54 of the amino terminus and various residues from the B-I-D-G β-wall (Fig. 2a) form the subunit clamp site and interact with residues 183–190 of the invading carboxyl terminus (Fig. 5). These interactions include both hydrophobic and hydrophilic contacts. Phe186, an essential amino acid, makes hydrophobic interactions with many different residues (Table 1) in a 'fist in cupped-hand' manner. The 'fist' (Phe186 side chain) fits into the 'palm' (B-I-D-G β-wall), and the 'fingers' (long aliphatic portions of the residues) reach around and past the 'fist' to the 'wrist' where some form charged or polar interactions. Phe186, Arg90, Glu174 and Glu176 are all conserved among the bromoviruses (Table 1) and of the five other residues making this interaction, four are conserved between CCMV and BMV (Fig. 3). Hydrophilic capsid dimer contacts involve residues in or near the single turn carboxy-terminal helix (residues 179–187). Asp184 forms a salt-bridge to Arg179 at the end of βI, and also hydrogen bonds with the main-chain nitrogen of Ala46. Neighboring Asp183 creates a main-chain hydrogen bond with Val52, and Lys54 on the βB strand hydrogen bonds with the Ser185 carbonyl oxygen. Overall, the molecular clamp maintains a high level of quasi symmetry. The quasi two-fold related subunits (A-B_5) have virtually identical contacts to those seen in the icosahedral dimers (C-C_2) which in turn leads to the truncated icosahedron morphology of the capsid (Fig. 1).

Table 1. Interactions between residue Phe186 on the extended A subunit carboxyl terminus and residues of the invaded B_5 subunit.

Residue in B_5	Location in fold	No. of contacts to Phe186 in A	Closest contact (Å)	Average contact (Å)
Ile44	amino terminus	15	3.6	4.5
Val52	βB	15	3.7	4.4
Arg90	βD	8	4.0	4.5
Val91	βD	1	4.7	–
Leu92	βD	12	3.9	4.4
Met137	βG	2	4.6	4.7
Glu174	βI	6	3.8	4.5
Glu176	βI	4	4.0	4.5

Contacts are defined as pairs of atoms ≤5.00 Å apart. Ile44 and Val52 hold positions directly before and after and parallel to the plane of the buried Phe186 phenyl ring creating numerous contacts.

Metal ion binding sites

Quasi three-fold interactions between neighboring capsomeres are limited to intersubunit, metal-mediated, carboxyl clusters that lie 8 Å from the quasi three-fold axes (Fig. 1b). These clusters contain three acidic residues (Fig. 3) (two from one subunit, one from the other) which are protonated because the crystals were grown at pH 3.3 in the presence of EDTA. Thus, although no divalent metal ions are present in the solved structure, the likely mode of protein–metal interaction is apparent if

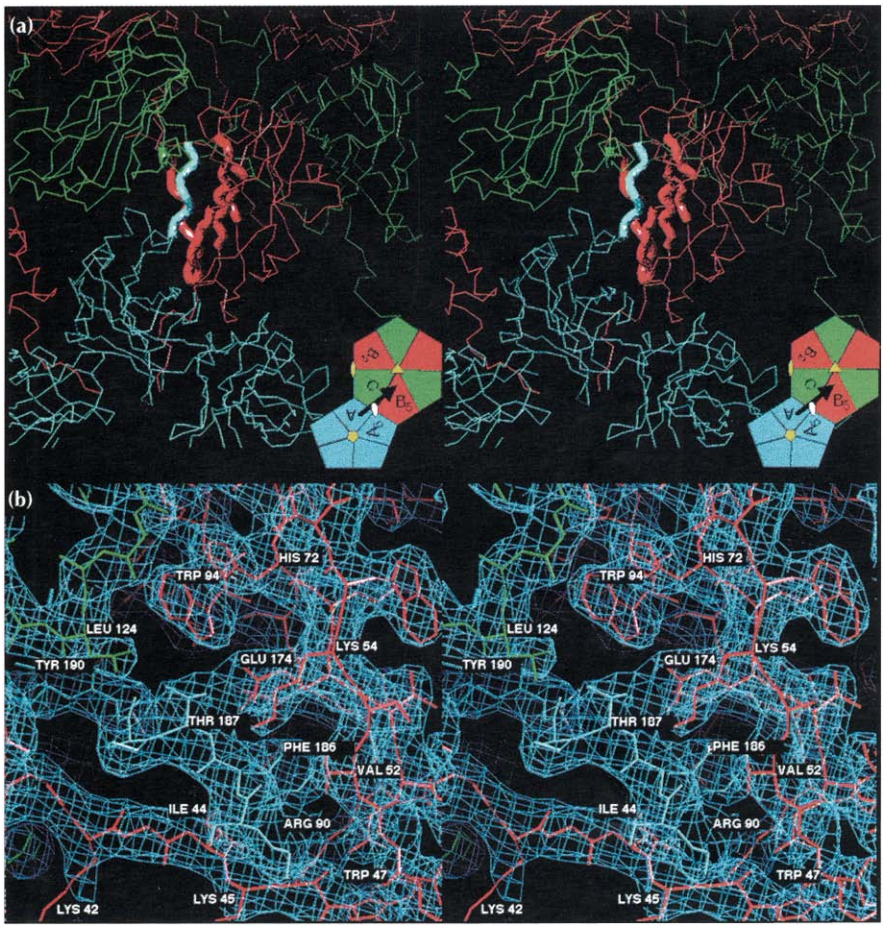

Fig. 5. Stereoviews of the A-B_5 quasi two-fold dimer contact responsible for binding pentamers to hexamers (see Fig. 1 for definitions and color coding). **(a)** An overview of the interaction showing the A subunits (blue) clustered around a five-fold axis and the B (red) and C (green) subunits clustered around a quasi six-fold axis. The bold tubes correspond to the regions of modeled electron density detailed in part (b). The inset is taken directly from Fig. 1b and displays the figure orientation relative to the particle symmetry axes. The black arrow represents the A subunit carboxyl terminus shown invading the B_5 subunit here, and the visual orientation for (b). **(b)** A detailed view of the 3.2 Å electron density (light blue cage) for the interaction of the carboxy-terminal portion of the A subunit (blue wire model, residues 184–190) with the 'clamp' region of the B_5 subunit (red wire model, residues 41–56, 91–95, 133–136, 171–176, the orientation is similar to that of the CCMV ribbon model in Fig. 2a). A small portion of the C subunit (green wire model, residues 123–125) is also visible. Hydrogen bonding of Thr187 by Glu174, the 'fist in hand' interactions of Phe186 (see text; Table 1), and the interaction of Leu124 of the quasi six-fold or five-fold related subunits with the Thr187 methyl group occur at each of the 180 clamp sites. This view is rotated 30° about the horizontal axis relative to (a). Views of the C-C_2 icosahedral dimer contacts would be indistinguishable from those shown here except for the replacement of the pentamer with a second hexamer.

putative Ca^{2+} ions are modeled at the A-B, B-C and C-A subunit interfaces (Table 2, Fig. 6). It is noteworthy that all CCMV coordinating residues stem from helical structure and, although the helices reside on separate polypeptides, they form a helix-loop-helix type motif as seen in the intracellular Ca^{2+} receptor calmodulin [35]. The slightly overestimated Ca^{2+}-ligand distances in Table 2 reflect the protonated state of the interface and absence of electrostatic attraction to a resident calcium ion. Glu81 appears to be an essential bidentate ligand conserved among all three bromoviruses (Fig. 3). The six ligands adopt approximate trigonal biprism geometry, however, water molecules (not visible at this resolution) may be additional ligands and some rearrangement of all ligands in the presence of Ca^{2+} is likely.

Table 2. Details of the putative calcium-binding site between the A and B subunits; one of 180 such sites in the CCMV capsid.

Subunit	Coordinating residue	Location in fold	Coordinating atoms	Distance to putative Ca^{2+}
B	Glu81	αCDII	Oε2	2.5 Å
B	Glu81	αCDII	O	3.5 Å
B	Gln85	αCDII	Oε1	2.4 Å
B	Glu148	αGH	Oε1	3.8 Å
A	Gln149	αGH	Oε1	3.6 Å
A	Asp153	αGH	Oδ2	2.9 Å

Protein–RNA interactions

In addition to the charged and disordered amino-terminal polypeptide, 13 ordered residues interact with RNA in each CCMV subunit (Fig. 3). Most of these interactions are ionic and involve lysine and arginine residues that extend into density that is well above background but is not readily modeled in a specific RNA conformation. One tryptophan residue, that lies close to the quasi three-fold related subunit interface, clearly extends the base stacking of two consecutive ribonucleotides and similar interactions occur at its three quasi-equivalent positions (Fig. 7). In these areas, electron density at approximately 2σ contour level reveals molecular boundaries that are well resolved and are separate from the capsid protein. The density can be modeled as partially helical, single-stranded RNA (selection of purine or pyrimidine was arbitrary). When compared with the A-B interface, density in the B-C and C-A interfaces was better resolved at higher contour levels. The planes of the nucleotide bases closest to protein contact the plane of the tryptophan side chains and this mimics the stacking interaction of a third or fifth nucleotide. At all three interfaces Trp47, Arg82, Glu140, Lys143 and Thr181 appear to interact with the RNA (Fig. 7). Glu140 and Lys143 of the same subunit, and Arg82 of an adjacent (quasi three-fold related) subunit are within a 3 Å radius of one another allowing Glu140 to form salt-bridges with both residues. These contacts obey local quasi three-fold symmetry, but the RNA structures which can be modeled beyond the Trp47 contacts (and also the unmodeled, bulk RNA density) exhibit conformational differences (Fig. 7). The modeled RNA is part of continuous density which becomes weaker and less well resolved as it approaches the quasi three-fold axes, but remains clearly visible above the background at these axes. The RNA forms three distinct bulges directly under the shell near the quasi three-fold axes where it interacts with Lys87 residues. Unlike the RNA density at the Trp47 contact, this density could not be accurately modeled. This density may represent RNA optimally positioned for release from the particle when it expands at alkaline pH (see below).

Structure of native particles determined by cryoEM and image reconstruction

The structure of native CCMV, derived from cryoEM and image reconstruction, was used to develop the initial model for calculating phases for the X-ray structure amplitudes at low resolution (see the Material and methods section). It was subsequently compared with the protein model derived from the 3.2 Å resolution X-ray map (Fig. 8a) and found to agree with remarkable fidelity.

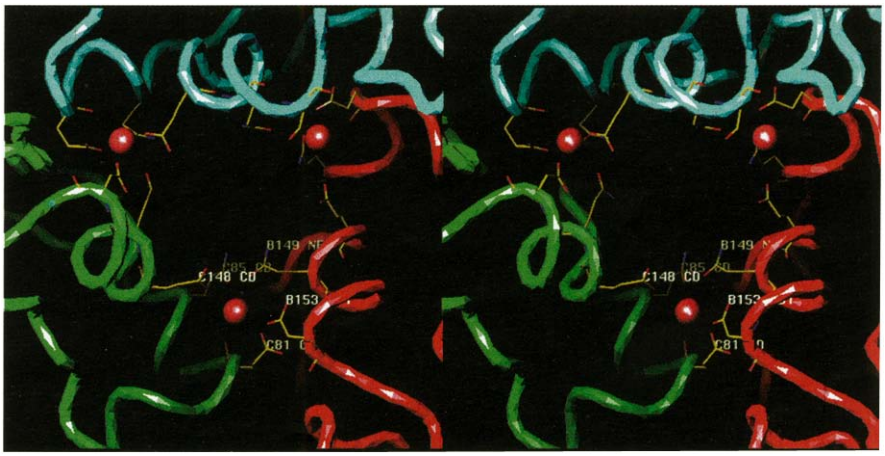

Fig. 6. The putative quasi three-fold related calcium-binding sites in the CCMV capsid. The view is looking from the exterior to the interior of the virus directly down a quasi three-fold axis. Continuous tubes, color coded as in Fig. 1b, trace the polypeptide Cα atoms. Modeled calcium ions are displayed as brown spheres at each subunit interface (refer to Fig. 1b for their relative locations in the overall capsid) where only the potential calcium coordinating residues are shown, color coded by atom type (nitrogen, blue; oxygen, red; and carbon, yellow). Residues at the B-C subunit interface are labeled (as in Fig. 3 which shows the locations of the residues that interact with metal ions) in white. Positions of the αCDII and αGH helices are at the bottom edges of the quasi three-fold opening, and lining the sides of the opening more towards the exterior of the particle, respectively.

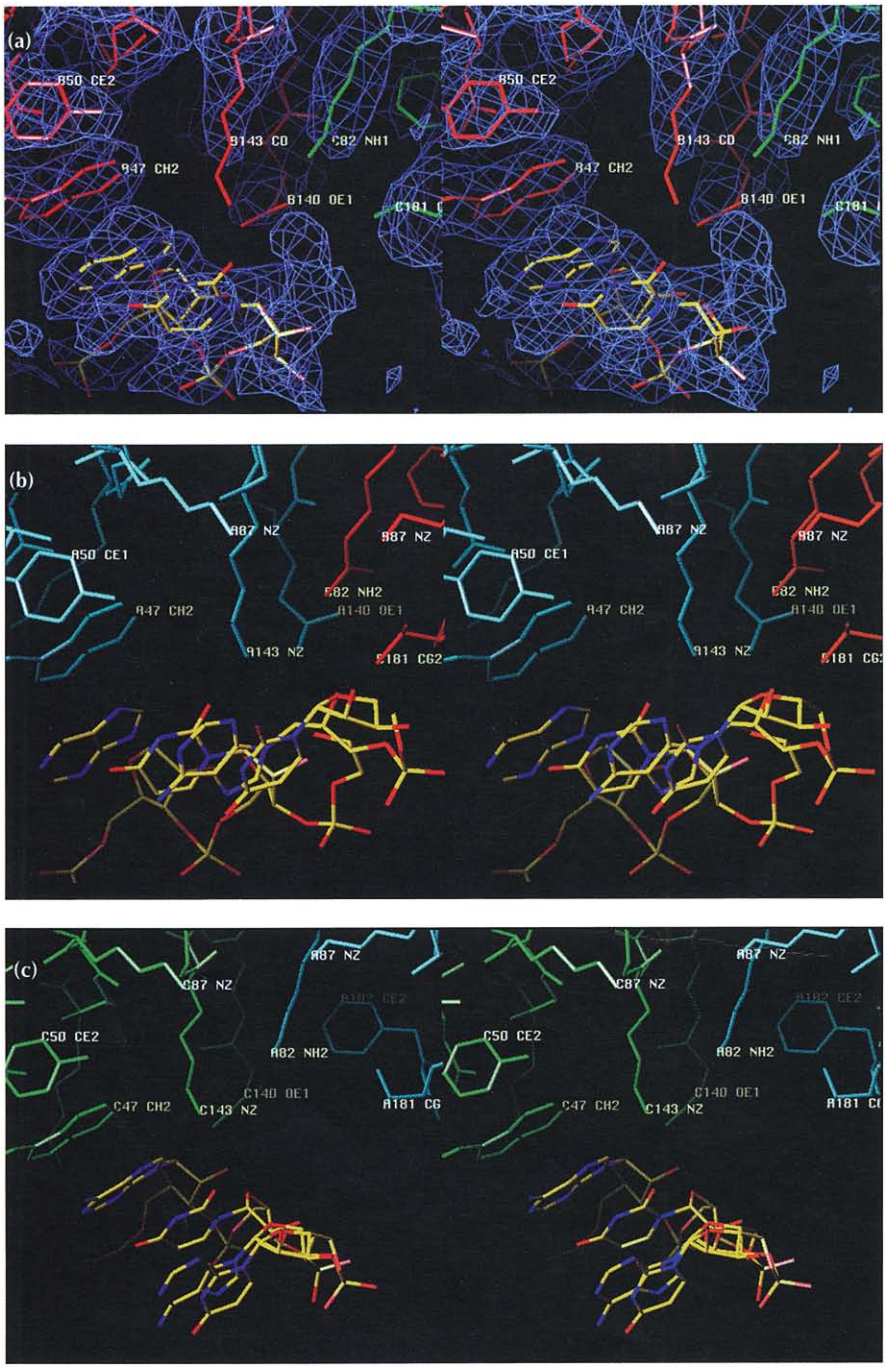

Fig. 7. Similarly oriented stereoviews of the protein–RNA interactions at the A (blue), B (red), and C (green) subunit interfaces. The views are tangential to the protein shell looking away from the quasi three-fold axis along the subunit interfaces. The RNA model is colored according to atom type (nitrogen, blue; oxygen, red; and carbon, yellow). Selected residues are labeled as in Fig. 3, which shows the location of residues that interact with RNA. **(a)** The B-C subunit interface. Electron density (blue) at approximately 2σ contour level is shown for this interaction, but is left out of the following parts of the figure for clarity. **(b)** The A-B subunit interface. **(c)** The C-A subunit interface.

The distribution of density assigned to RNA in the cryoEM map was also comparable to that of the X-ray map. The strongest RNA density was just under the B subunit near residues 47–49. Like the RNA density seen in the X-ray structure, the RNA density seen by cryoEM fills the B-C interface and extends towards the quasi three-fold axes. However, the quasi three-fold symmetry breaks down as only very weak density is found at the C-A and A-B interfaces. Density was also located below the five-fold and three-fold axes in the form of flat oval plates and three-fold symmetric tubes, respectively. Some of this density may correspond to the polypeptide amino termini (residues 1–27) that are disordered in the X-ray structure.

Structure of swollen particles determined by cryoEM and image reconstruction

The native capsid (Fig. 8a) opens and expands when transferred to an EDTA–phosphate solution buffered at pH 7.0 [8,13]. The structure of this swollen form of CCMV (Fig. 8b) was determined by cryoEM and image reconstruction to examine the structural basis for this polymorphism. An 'atomic model' of the swollen virus was built by individually fitting the A, B and C subunits from the X-ray model into the cryoEM density (Fig. 8b). Differences between the native and swollen atomic models were measured to quantitate the structural changes that had to be introduced to convert from the native to the expanded structure (Tables 3 and 4). The

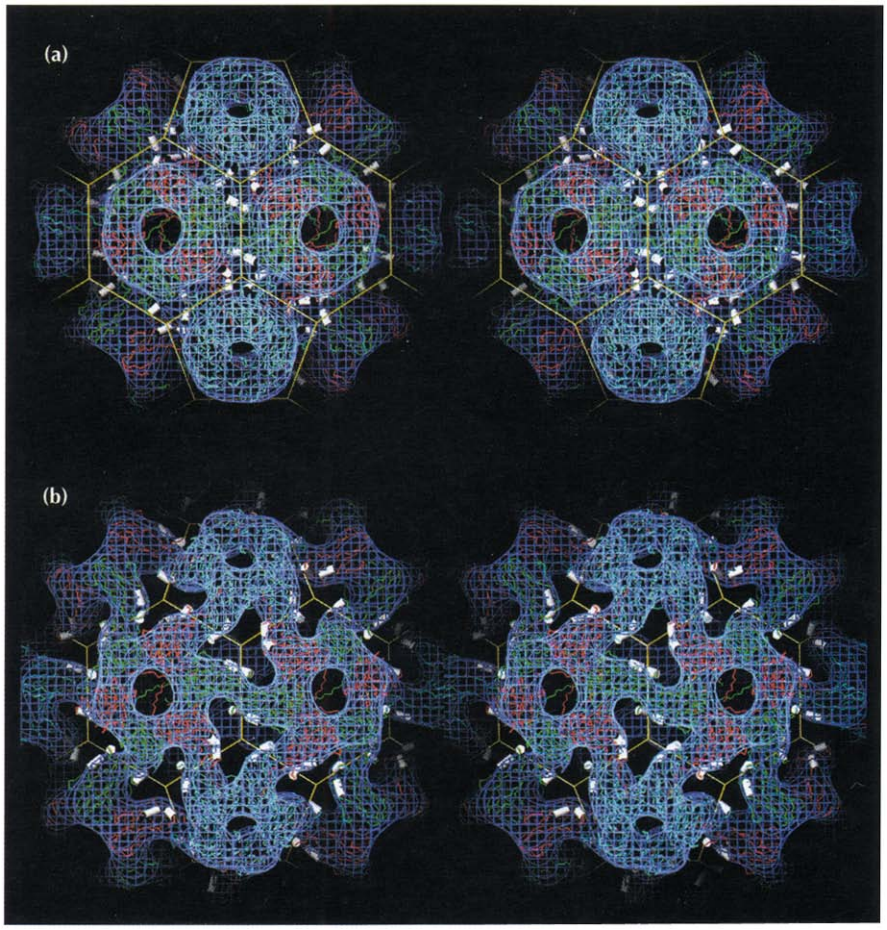

Fig. 8. Stereoviews of two forms of the CCMV capsid showing that its polymorphism is controlled by changes in pH and metal ion concentrations. Atomic coordinates determined from X-ray crystallography and the truncated icosahedron cage are displayed as in Fig. 1a with electron density (blue) determined by cryo-electron microscopy. The helices which carry calcium-binding residues (αCDII and αGH in Figs 2a, 3, and 6) are displayed as white cylinders. **(a)** The CCMV capsid at pH 4.5. The atomic model is displayed as originally built in the X-ray electron density. The excellent fit between the 3.2 Å X-ray model and the 23 Å EM reconstruction shows the level of compatibility between the two independent structure determination techniques. Note the three pairs of CDII and GH helices positioned around each quasi three-fold axis (vertices of the truncated icosahedron). Calcium ions (180 per capsid) are bound between morphological units by residues associated with the helices (Table 2, Fig. 6). **(b)** The swollen CCMV capsid at pH 7.5 in the absence of metal ions determined at 28 Å resolution. Atomic coordinates from the high resolution native structure were modeled to fit the density from the cryoEM reconstruction.

hexamers and pentamers remain essentially intact and unchanged, except that they move to higher radius along quasi six-fold and five-fold axes and therefore separate from each other at the quasi three-fold axes. Swelling may be triggered by electrostatic repulsion at the Ca^{2+}-binding sites near the quasi three-fold axes. When metal ions are removed and the pH is raised to 7.0, the repulsion of these negatively charged acidic residues expands the virus at the quasi three-fold axes. It is likely that the pK_a of the acidic residues is abnormally high owing to their close proximity and possibly the presence of phosphate residues from the RNA. Thus, swelling does not occur until pH 7.0, which is higher than the normal pK_a (~4.0) of aspartic acid and glutamic acid residues [13]. The separation between acidic residues that ligand to the Ca^{2+} ions (Asp153 and Glu81) in the native structure increases by ~5 Å in the metal-free, swollen form (Tables 3 and 4) and leads to the formation of large holes in the protein shell. Such holes form conduits for the radially expanded RNA genome, which lies below the openings (Fig. 8b). Most of the RNA is close to the quasi three-fold axes and in domains that are connected by strong rods of density that cross both quasi and icosahedral two-fold axes. This distribution of RNA, which is like a large ball-and-stick structure, where the balls are the strong density just under the quasi three-fold axes, closely follows the truncated icosahedron cage shown in Fig. 8b. Protein–protein interactions at the quasi three-fold related interfaces appear to be absent after expansion. The space vacated by the Lys87 residues on expansion creates a clear path for RNA to leave the particle and this may explain why, in the swollen form of CCMV, the RNA can act as a template for protein synthesis and as a substrate for RNase [8,36].

The greatest differences between the native and swollen CCMV models occur at the icosahedral and quasi two-fold axes (~12% expansion), whereas the smallest change occurs at the pentamer axes (~7% expansion) (Tables 3 and 4). Although the pentamers and hexamers retain their shape and relative icosahedral positions following expansion, all subunits within each capsomere tilt inwards by about 7°, leading to a widening at the base and some compression at the outside of each capsomere. This gives the capsomeres their slightly flattened appearance compared with those in the native virus (Fig. 8). The capsomeres also rotate by ~5° about their axes, producing more linear carboxy-terminal connections between capsomeres compared with the native particles. These structural changes maintain the contacts between capsomeres, but at higher radii, and account for the greater displacement of residues near both types of two-fold axes. Nevertheless, these dimer interactions remain largely intact (curving ropes of electron density between pentamers and hexamers in Fig. 8b). The carboxyl termini in the swollen model continue to be clamped by subunits in adjacent capsomeres. As observed in the native structure the carboxy-terminal interactions at the

Table 3. CCMV capsid expansion measurements.[a]

Subunit interface	Distance 1[b]		Distance 2[c]	
	Native	Swollen	Native	Swollen
B-C	8.3	14.2	10.4	17.7
C-A	7.9	11.7	10.6	16.5
A-B	8.0	14.1	10.8	18.2
Average	8.1	13.3	10.6	17.4
Expansion	5.28		6.88	

[a]All measurements are in Å. [b]Measured between Cα atoms of residues Asp153 and Glu81. [c]Measured between Cα atoms of residues Gln149 and Gln85.

Table 4. CCMV capsid expansion measurements.

Subunit	Radial enlargement[a]				
	5fold axis[b]	3fold axis[b]	2fold axes[c]	Interior[d]	β-hexamer[e]
Native A	142.4	–	119.9	105.5	–
Swollen A	152.9	–	133.4	117.0	–
Native B	–	139.7	119.8	103.5	95.4
Swollen B	–	152.2	135.1	117.4	103.1
Native C	–	139.7	119.4	103.4	95.4
Swollen C	–	150.7	134.1	116.4	101.3
Native av.	–	139.7	119.7	104.1	95.4
Swollen av.	–	151.4	134.2	116.9	102.2
Expansion	10.5	11.7	14.5	12.8	6.9

[a]All measurements are in Å. [b]Largest Cα atom radial value was for Glu63. [c]Largest Cα atom radial values was for Asn76 at both icosahedral and quasi-icosahedral axes. [d]Radial distance average for Cα atoms of residues 42–52. [e]Smallest Cα atom radial value was for Val27.

quasi two-fold and icosahedral two-fold axes are closely similar, although a thinner tube of density covers the icosahedral interaction (Fig. 8b). The model suggests that the essential features of the β-hexamer are retained in the swollen structure, but the parallel β-tube becomes partially disrupted when the subunits tilt.

Discussion

Geometry of *T*=3 virus particles

The structure of CCMV displays a subunit organization anticipated, but not previously found, in *T*=3 RNA plant and insect virus structures [19,20,24–29]. It is novel in that it follows, in detail, *T*=3 quasi-symmetry as envisioned by Caspar and Klug [37]. Their conceptual model for quasi equivalent viruses anticipated a sheet of hexamers interspersed with 12 pentamers arranged to form a closed shell with icosahedral symmetry. The chemical basis for such a construction was the requirement that a single gene product form both hexamers and pentamers. Quasi equivalent chemical interactions were assumed to be maintained when subunits were related by a 72° rotation or a 60° rotation, but the nature of the molecular switch that determines which is formed was not anticipated.

The morphology of CCMV is that of a truncated icosahedron [38] (Figs 1a and 1b) with three chemically identical gene products, labeled A, B and C, in each icosahedral asymmetric unit. The pentamers and hexamers precisely fit their theoretical locations in the truncated icosahedron and impart a strikingly different capsid morphology compared with other RNA viruses studied at high resolution. The quasi six-fold axes are exact 60° rotation axes in the local environment. The dihedral angles between the B_2-C and between the C-B_5 polygons are 180°. This causes all the quasi six-fold related polygons to be co-planar. Curvature (i.e. closure) of the solid takes place at and near the quasi three-fold axes formed by the edges of the planar pentamers and hexamers. The C-B and C-A polygon interfaces are bent between capsomeres to similar angles (~140°). This can be seen in the CCMV structure (Fig. 1a) as the quasi three-fold related subunits tilting away from the quasi three-fold axes (and each other) thus maintaining interactions with five-fold and quasi six-fold related subunits.

The rhombic triacontahedron [38] (Fig. 1c) best describes the geometry of all the *T*=3 plant and insect virus structures previously determined. The quasi three-fold related polygons in the rhombic solid form planar triangular units. Two of these triangles, related by an icosahedral two-fold axis, form a planar, diamond-shaped surface that is perpendicular to the two-fold axis. Rotation of the rhombic plane by the icosahedral three-fold axis generates the quasi six-fold axis, but this is actually a trimer of dimers where intersubunit contacts are distinctly different at the C-B_2 interface (flat) and the B_5-C interface (144° dihedral angle). Although the same chemical surfaces are adjacent at the two interfaces, the geometric interaction is totally different. These subunit–subunit interactions are altered by the presence (flat) or absence (bent) of a polypeptide [19,20,24–26] or RNA segment [27,29] in the internal part of the interface in *T*=3 plant and insect virus structures determined to date.

Comparison of CCMV with papovaviruses: the novel role of the carboxy-terminal region

The only virus structures solved to date with quaternary interactions similar to CCMV are the DNA tumor papovaviruses, SV40 and polyoma [31,32]. Capsids of bromoviruses and papovaviruses are best described as stable, modular, morphological units that are 'tied' together by carboxy-terminal extensions of the subunit polypeptide which integrate into the tertiary structure of a two-fold related subunit. This concept is best illustrated in a CCMV model in which a single hexamer has been removed (Fig. 9). The carboxy-terminal extensions of neighboring hexamers and pentamers are clearly visible in the space occupied by the hexamer. Nearly identical dispositions of carboxyl termini are found in the space occupied by the pentamers (not shown). Papovavirus shells are composed of 72 pentamers associated through long

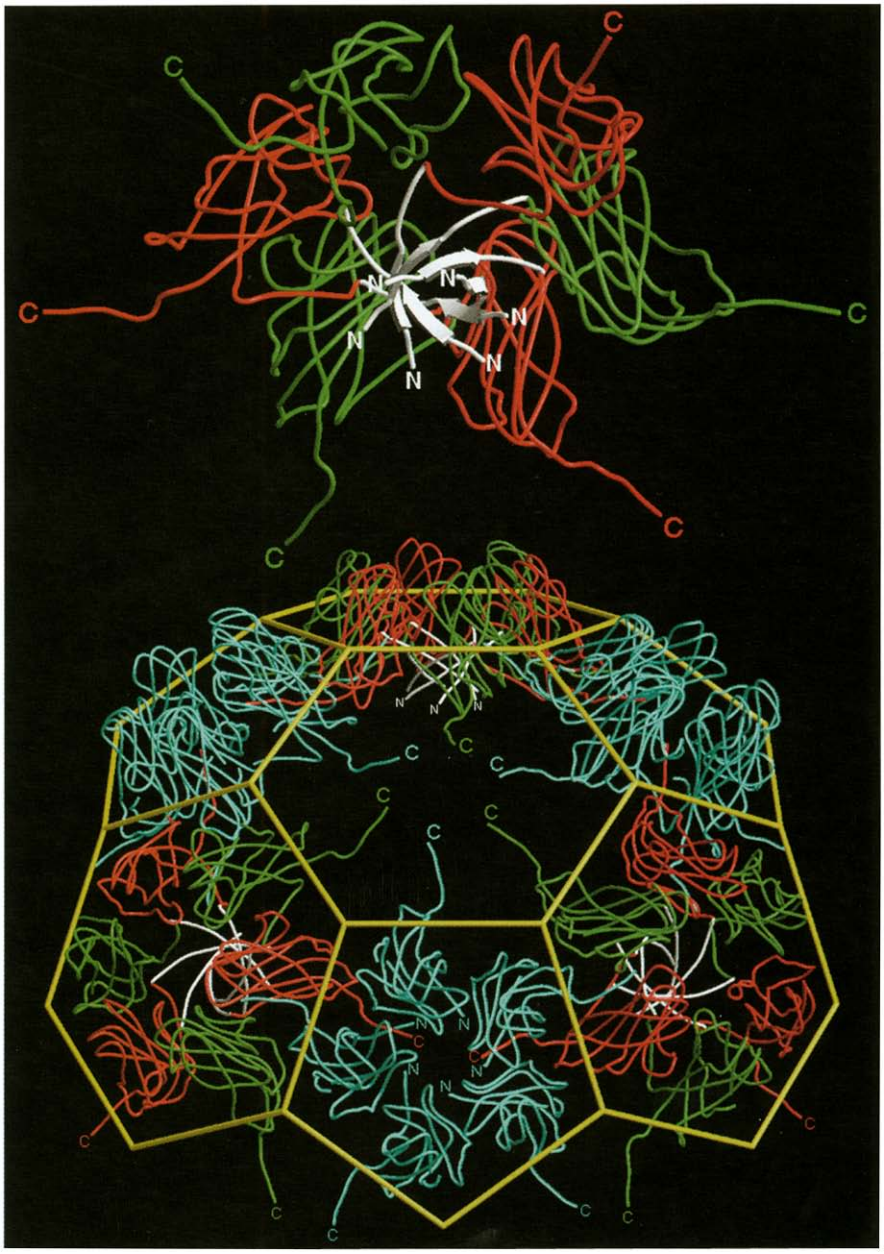

Fig. 9. An exterior view of a portion of the viral capsid close to a quasi six-fold axis. Continuous tubes represent subunit Cα tracings and are color coded as in Fig. 1. The yellow cage represents the idealized truncated icosahedron. Selected amino and carboxyl termini are labeled and color coded by subunit of origin. A hexameric morphological unit has been removed from the center of the contiguous shell and rotated about the horizontal axis to show on its underside the intracapsomere parallel β structure (white) that stabilizes this morphological unit (the β-hexamer, Fig. 4). The hole left in the shell displays the disposition of residues 181–190 of the six carboxyl termini which invade the absent hexamer and interact with the clamp as shown in Fig. 5. Likewise, the removed hexamer displays the same carboxy-terminal residues that invade neighboring pentamers and hexamers pointing outward around its periphery. The orientation of the bottom pentamer provides a view down a five-fold axis. Note the lack of ordered interactions between the A subunit amino termini near the axis (i.e. disorder after residue Lys42), however, they have sufficient structure to clamp incoming B subunit carboxyl termini.

carboxyl termini extending from the subunit module. Although these contacts are architecturally similar to those in CCMV, the papovavirus carboxyl termini have differences in the amount of ordered structure, dependent on their location in the capsid. Changes in the carboxy-terminal structure in papovavirus subunits allow a response to the type of pentamer–pentamer interaction required in the T=7 surface lattice in such a way that variations in contacts between morphological units are tolerated while maintaining specificity. Significant differences in the conformation of the carboxy-terminal residues (181–190) are also observed in the CCMV A, B and C subunits. These are determined by the local requirements to bind either adjacent pentamer (quasi two-fold related) or hexamer (icosahedral two-fold related) capsomeres. Unlike the papovaviruses, all the residues are well ordered in both environments in CCMV.

All subunits in each capsid type both 'clamp' and donate an extended carboxyl terminus. Carboxyl termini become wedged between sections of a single subunit's amino terminus and β-barrel in the target capsomere for CCMV and polyomavirus. In CCMV, an invading carboxyl terminus is clamped in the gap between residues 36–52 of the amino terminus and the B-I-D-G β-wall of the receptor subunit. In adjacent pentamers, the polyoma carboxy-terminal βJ-strand wedges between βA and βB of the invaded subunit to form an extended antiparallel A-J-B-I-D-G β-sheet interaction (both ribbons in Fig. 2 are oriented to view the area of carboxyl termini insertion). Although the polypeptide termini are used in identical manners, the clamp sites are not in identical environments (note the approximate 90° rotational difference between the CCMV and polyoma β-barrel orientations in Fig. 2).

Role of carboxy-terminal regions in other viruses

Carboxyl termini do help to stabilize quaternary structures in other viruses, but in ways that differ from CCMV and papovaviruses. The 100 residues at the carboxy-terminal end of TBSV and TCV (T=3 plant viruses) and the 20 carboxy-terminal residues of the VP1 and VP3 subunits of P=3 picornaviruses, poliovirus, rhinovirus and mengovirus also stabilize quaternary structure [19,20,24,39–41], (these are also known as pseudo T=3 because β-barrels forming VP1, VP2 and VP3 correspond positionally to A, C and B subunits, respectively in T=3 structures, each with a unique amino acid sequence). In TBSV and TCV, the residues fold into 10 β-strand protruding domains (P-domains) positioned on the exterior of a contiguous protein shell formed by β-barrels (S-domains). Surfaces of P-domain β-sheets form an extensive hydrophobic contact across two-fold or quasi two-fold axes creating distinct dimers situated on the outer viral surface. This P-dimer interaction is not found in SBMV, another T=3 plant virus. The SBMV capsid subunits and the S-domains of TBSV and TCV have closely similar tertiary structures (β-barrel folds), but SBMV has no P-domain. Without P-domains SBMV subunits assemble into a rhombic triacontahedron particle that is virtually identical to the shell domain of TBSV. SBMV also expands under metal-free, alkaline conditions like TBSV and TCV, thus raising questions about the structural role of the P-domain. The picornavirus carboxy-terminal extensions lie over the outside surface of adjacent subunits in the capsid protomer (the P=3 protomer: VP1=A, VP2=C, and VP3=B_5 in Fig. 1c). In each case the role of carboxyl termini in the formation and stabilization of the capsid is unique and differs in comparison with those found to be highly similar in CCMV, SV40 and polyomavirus. The external carboxyl termini described in this section are not integrated (clamped) into an adjacent capsomere subunit fold.

Role of ordered amino termini in particle formation

The formation of the β-hexamer stabilizes the hexameric morphological units. It is the first 'β-annulus'-type structure known to follow the predicted quasi symmetry. In previously studied T=3 plant viruses, the visible amino-terminal portions of three polypeptide strands (each contributed by C subunits) form a β-annulus centered on the icosahedral three-fold symmetry axes (quasi six-fold axes) [20,24,26]. Because only C subunit amino termini are involved in the β-annulus, they are trimeric (i.e. they follow the icosahedral three-fold symmetry) and form a T=1 structure without quasi equivalent counterparts in the A and B subunits, where the same sequence of amino acids is disordered [42]. The ordered amino-terminal portions also form a wedge that holds the C and B_2 subunits in the planar conformation in the rhombic triacontahedron. In the T=3 black beetle (BBV) and flock house (FHV) insect viruses the amino termini are disordered, except for a small stretch in the C-B_2 interface and therefore do not form β-annulus structures. Ordered duplex RNA fills the C-B_2 interface as well and makes several interactions with the coat proteins [27,29]. The same sequence of residues in the A and B subunits in both the plant and insect viruses does not follow icosahedral symmetry and is not visible in the electron density. In contrast the CCMV amino termini that form the β-hexamer are contributed by the B and C subunits, do not contact adjacent β-barrels, and do not directly affect the interfacial contacts between the two-fold and quasi two-fold related subunits.

A pentameric β-tube observed in the P=3 animal picornaviruses [39–41] is reminiscent of the CCMV β-hexamer, but it does not affect quasi symmetry. Five VP3 subunit amino termini converge at the icosahedral five-fold axes and make parallel β-strand interactions which, like the β-annuli at the plant virus three-folds, follow icosahedral symmetry. The picornavirus β-tubes stabilize a well-characterized pentameric assembly intermediate made up of five protomers $(VP0,VP1,VP3)_5$ [39,40], where VP0, VP1 and VP3 correspond positionally to C, A and B_5 in Fig. 1c.

Comparison of the expanded forms of CCMV and TBSV

The loosely coupled CCMV capsomeres can form an open, expanded structure above pH 7.0 in the absence of metal ions, or they can form a rigid, compact structure at acidic pH or in the presence of divalent metal ions (Fig. 8). Similar states have been observed in other plant viruses such as TBSV for which an 8 Å structure has been determined [21]. Stability of the swollen forms of TBSV and CCMV appears to be derived from different interactions as found for their native particles. The cryoEM-based atomic model of swollen CCMV displays all subunit carboxyl termini still bound by adjacent capsomeres and these interactions, as well as protein–RNA contacts, appear to stabilize the particle. Although it is impossible to make a definitive statement at this resolution and based solely on the model studies, there is no need to invoke the formation of additional ordered structure for the stabilization of CCMV as was observed in TBSV. Swollen TBSV is stabilized by residues in the amino termini of the A and B subunits; these become ordered only in the expanded particles and form extra β-sheet structure across the quasi two-fold axes, while icosahedral two-fold dimers are essentially unchanged [21]. Expansion in CCMV and TBSV is triggered by the ionization of acidic amino acids that normally bind Ca^{2+} or that are protonated at acidic pH. The resulting charge repulsion creates a large conformational change that results in the creation of openings that are ~20 Å in diameter at the quasi three-fold axes. Although the expanded structure of CCMV and the shell of TBSV are closely similar, the location of the metal-binding residues in the tertiary structures of TBSV and CCMV is different. TBSV binds two Ca^{2+} at each quasi three-fold interface with anionic ligands on the βC-βD insert, the βE strand, a connection to βF, and main-chain carbonyl oxygens contributed from single residues on the βG strand and the end of βH. CCMV Ca^{2+}-binding ligands are shown in Figs 3 and 5.

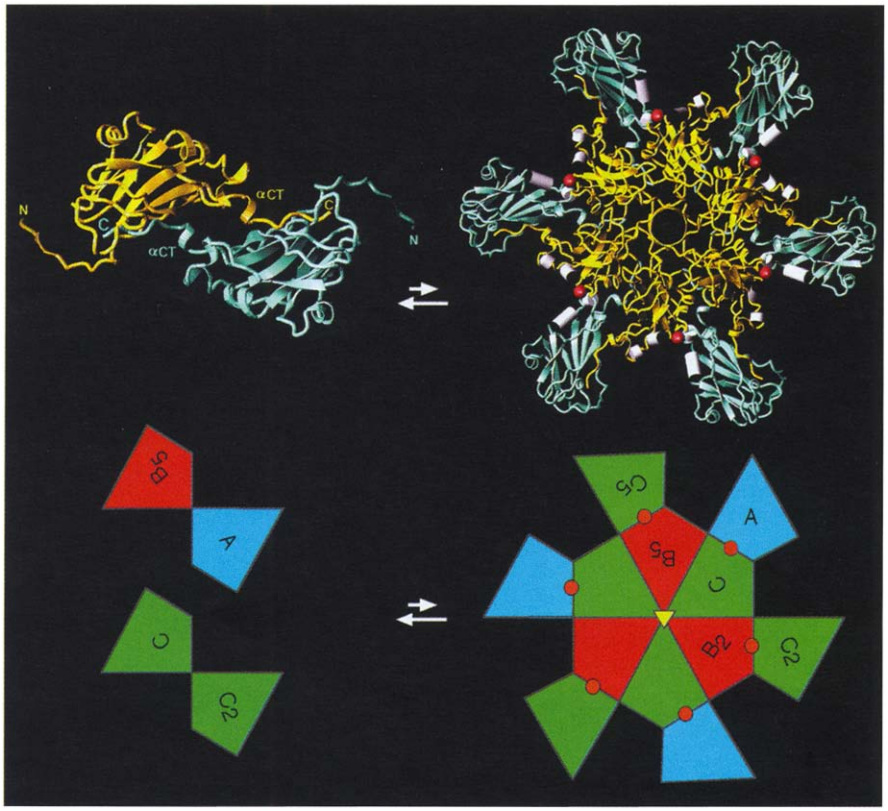

Fig. 10. Proposed intermediates in CCMV assembly. A dimer (shown left and viewed from the exterior directly down a particle two-fold axis) and a 12mer (shown right and viewed from the particle exterior directly down a quasi six-fold axis) are displayed both as ribbon diagrams and as polygons from Fig. 1. Color coding of the ribbons is a mix of either B and C, or A and C colors (gold and light blue, respectively) and denotes those relationships, displayed by the polygon coloring, that cannot be distinguished until a greater portion of the capsid is formed. Amino termini and carboxyl termini are labeled for the dimer ribbon showing the exchange and clamping of carboxyl termini across both icosahedral and quasi two-fold axes (displayed in the lower part of the figure as $C\text{-}C_2$ and $A\text{-}B_5$ polygons). Six such dimers form a 12mer driven by the preferred formation of the β-hexamer. Helices that coordinate calcium (shown as red spheres in the ribbons and polygons) are displayed as white cylinders at subunit interfaces which are also capable of binding RNA.

Structural basis of polymorphism in CCMV assemblies

The polymorphism observed in CCMV and other members of the *Bromoviridae* results from the strong tendency to form hexameric morphological units. The assembling unit is most likely a subunit dimer (Fig. 10a) stabilized by reciprocal 'clamping' of the eight carboxy-terminal residues (183–190). Integration of carboxyl termini in the two-fold related β-barrel fold creates the greatest number of non-bonded interactions in the capsid. Hexamers, composed of six dimers (Fig. 10a), would form next. They may exist in equilibrium with dimers in solution and probably form the nucleation site for particle formation. Hexamers have twice the number of interactions per subunit compared with pentamers. This difference is due to the absence of an ordered amino-terminal structure in the pentameric A subunits. There is no tendency for a 'β-pentamer' structure to form about the five-fold axes because the reciprocal carboxy-terminal and amino-terminal interactions that stabilize the dimer (Fig. 5) orient the polypeptide composed of residues 42–47 and lead to direct steric hinderance with residues in five-fold related β-barrels. As a result, residues before Lys42 are disordered in the capsid pentamers and this suggests that there are probably no pentamers in solution. It is likely that they only form during the assembly process.

The B-C and C-A subunit interfaces (Fig. 1b) are present in the nucleating 12mers. Residues in these interfaces are involved in the binding of metal ions and RNA. Binding of either will induce curvature at each interface prior to the addition of more dimers (Fig. 10b). Particle growth is dependent on the addition of dimers because two hexamers cannot directly assemble with each other. Addition of dimers to the 12mer form A-B, B-C and C-A interfaces. Each newly created interface binds RNA, and the resulting curvature drives the formation of pentamers to form a closed shell. Note that ordered RNA is present at each of these interfaces in the native particle (Fig. 7) as are the Ca^{2+}-binding ligands (Fig. 5). Various CCMV coat protein aggregates can be made under well-defined conditions *in vitro* in the absence of RNA and metal ions, but there is a great tendency to form structures with little or no curvature. Many of the laminar structures and tubes assembled appear to be sheets of hexamers that are either planar or gently curved to accommodate a packaging role for the protein, suggesting that, without inducement, pentamers are not formed. The presence of RNA and metal ions during assembly experiments invariably produces smaller spherical particles.

The importance of protein–RNA interactions in stabilizing CCMV and dictating its morphology is emphasized by comparing the visible RNA in CCMV with that observed in other structures. Ordered nucleic acid has been reported for six other virus structures [27,29, 43–46]. However, the structure described in this paper is the first report of ordered nucleic acid in a T=3 plant virus. The CCMV structure also displays the first RNA structure related by quasi icosahedral symmetry. The ordered RNA in the T=3 insect viruses FHV and BBV, and in the plant satellite tobacco mosaic virus (STMV), is part of double-stranded loops positioned at icosahedral two-fold axes. The only other plant virus structure in

which ordered RNA has been seen is bean pod mottle virus [43], where the RNA has taken up positions related by icosahedral three-fold symmetry. A total of 600 nucleotides can be modeled out of ~3000 packaged within each CCMV capsid.

The hexameric units of CCMV capsid protein, in the absence of metal ions, can coat nucleic acid and form shells of a size proportional to the RNA packaged, as can native virions of alfalfa mosaic virus (ALMV) and the isometric labile ringspot (ILAR) viruses, both of which are grouped in the *Bromoviridae* on the basis of genetic similarities to CCMV. We suggest that there are probably structural relationships among these viruses as well, although there is no detectable sequence similarity for the capsid subunits when bromoviruses are compared with ALMV and ILAR viruses. Neither ALMV or ILAR viruses bind Ca^{2+} and there is no evidence for particle swelling as observed in CCMV. The tendency to form pentamers may be minimal without the metal-mediated stabilization of the quasi three-fold interactions observed in CCMV. There must be biological advantages for the formation of particles that are not limited by a defined internal volume of an icosahedral shell. Comparison of virions displayed by members of the *Bromoviridae* demonstrates a likely evolutionary path for the development of pleomorphic particles that are not restricted with regard to their RNA content [2]. Particles of this type are frequently observed as nucleoprotein cores in animal viruses and it remains to be seen if there is any relation between the subunits forming these particles and those that form the pleomorphic plant virions.

Biological implications

Virions are nucleoprotein particles that have evolved to transfer their viral genomes among cells of a susceptible host and between the host organisms themselves. Particles must be sufficiently stable to protect the genome and yet be able to release it for replication and translation under appropriate conditions. Analysis of native cowpea chlorotic mottle virus (CCMV) at 3.2 Å resolution has revealed four structural features that probably explain the chemical basis for the formation of native and polymorphic assembly products: first, the presence of subunit (canonical virus jelly role structure) hexamers, stabilized by an unusual hexameric parallel β structure; second, a novel, carboxy-terminal-based, subunit dimer assembly unit; third, the site of protein–RNA interactions; and fourth, a cation-binding site that mediates subunit interactions.

The papovaviruses, polyoma and SV40, display networks of extended carboxyl termini in their quaternary structure that are comparable with CCMV. A striking similarity exists between the RNA plant virus and the DNA tumour virus at the point where the carboxyl termini are clamped in place by the amino-terminal region of the invaded polypeptide. Thus, not only is the association principle similar in the two virus groups, but so is the detailed interaction of the polypeptides. The intriguing similarity in quaternary structure between bromoviruses and papovaviruses suggests that the ubiquitous β-barrel fold of the subunits may not be the only structural feature conserved in viral evolution.

A common mechanism for achieving disassembly of plant viruses *in vivo* is an environmentally-induced particle expansion. Native CCMV virions swell by 10% when metal ions are removed with EDTA and the pH is raised to 7.0. We have been able to determine the structure of this particle at 28 Å resolution by cryo-electron microscopy and image reconstruction and model it with the coordinates of native CCMV subunits. Interestingly, the juxtaposition of negatively charged residues induces the formation of a 20 Å 'portal' for RNA release at the quasi three-fold axes. The structure is closely similar to the swollen form of tomato bushy stunt virus previously determined at 8 Å resolution by crystallography.

Materials and methods

Virus propagation, crystallization, and preliminary X-ray data analysis have been described [1]. Crystals of CCMV have $P2_12_12_1$ space group symmetry (a=381.3 Å, b=381.3 Å, c=408.6 Å) with four particles in the unit cell. A phasing model was developed on the assumption that CCMV subunits would have a β-barrel fold. The atomic coordinates for southern bean mosaic virus [26] were used to fit the density of an image reconstruction of CCMV based on electron micrographs of frozen hydrated virus. The final model was composed of 139 residues for each subunit, and the three subunits in the icosahedral asymmetric unit fit the density and had idealized quasi symmetry. Structure factors were computed for the model and cross-rotation function analysis at 15 Å resolution demonstrated good correlation with the observed data. The model orientation was consistent with self-rotation functions and the particle position was determined by sampling only the volume where no overlap could occur (based on the known particle dimension) with symmetry-related particles. Particle center positions were tested on a 2 Å grid by performing 5–10 cycles of molecular averaging and examining the correlation coefficient computed between the observed data and the data from the averaged, Fourier-transformed electron density. Although there were no crystallographic restrictions on the coordinates, the final molecular center was (0.0, 0.008, 0.0). Phases were determined to 4 Å resolution with the 60-fold non-crystallographic symmetry used for real-space averaging and phase extension [47,48]. A polyalanine model was constructed with the 4 Å electron density map and used to compute phases to 3.5 Å resolution. These phases were refined by real-space averaging. A detailed model, including side chains, was built into the 3.5 Å map and phases were computed to 3.2 Å resolution and these were then refined by real-space averaging. The model was improved on the basis of this density, which was computed with 631 446 reflections with an overall CC_{MR}=0.78 and R_{MR}=25.4% where

$CC_{MR}=\Sigma(<F_o>-F_o)(<F_c>-F_c)/[\Sigma(<F_o>-F_o)^2\Sigma(<F_c>-F_c)^2]^{1/2}$ and $R_{MR}=[\Sigma|F_o-F_c|/\Sigma F_o]\times100$ (F_o is the measured structure-factor amplitude and F_c is the structure-factor amplitude obtained by back transformation of the modified electron density). R_{mod}=31.4% after preliminary conjugate-gradient refinement of the 4339 protein atoms with X-PLOR where $R_{mod}=[\Sigma|F_o-F_c|/\Sigma F_o]$ (F_o is as defined above, F_c is the structure-factor amplitude computed from the atomic model) [49,50].

The cryoEM and image analysis procedures that were used to obtain the three-dimensional reconstructions of native and swollen CCMV were essentially similar to those previously reported [51–53]. The sample was maintained at near liquid nitrogen temperature in a Gatan 626 cryotransfer stage (Gatan Inc., Warrendale, PA) in a Phillips EM420 transmission electron microscope (Phillips Electronics Instruments, Mahwah, NJ). Micrographs were recorded under minimal dose conditions (<20e^- per Å^2), at an objective lens defocus of ~0.8 μm, and at an instrument magnification setting of ×49000. Micrograph selection and digitization, and particle image selection were performed as previously described [51–53].

Cross-common lines procedures [54,55] were used to select out a set of 16 particle images for native CCMV and 28 particle images for swollen CCMV from which the final three-dimensional reconstructions, with full 532 icosahedral symmetry [54] were computed. An R-factor calculation [56] indicated that the data were reliable to ~23 Å for native CCMV (40 particle images) and 28 Å resolution for the swollen CCMV (28 particle images). Each of these resolution values was estimated on the basis of a comparison (e.g., [51,53]) between two separate reconstructions that were computed from smaller and independently refined subsets of images.

The preliminary coordinates have been deposited in the Brookhaven Protein Data Bank.

Acknowledgements: We thank Michael Rossmann, Dan Marinescue, Robert Lynch, and the Purdue University Computing Center staff for use of the Intel iPSC/860 parallel processing system for real-space averaging of electron density; Robert McKenna and Mavis Agbandje for discussing methodology used in the feline panleukopenia virus crystal structure solution; Vijay Reddy for the use of his program to analyze icosahedral virus subunit contacts; Holland Cheng for assistance in converting the CCMV cryo-EM reconstruction electron density for viewing on graphics workstations; Jim Fox, Mark Young, Adam Zlotnick, Dino Moras, and Lars Liljas for insightful discussions on the CCMV crystal structure; and Sharon Fateley for help in the preparation of this manuscript. Figs 1a and 5–8 were generated with program O, Fig. 2 with MOLSCRIPT, and Figs 4 and 9 with Raster3D. This work was supported by National Science Foundation Grants 8817057A1-DMB (JEJ) and 9206305-MCB (TSB) and a grant from the Lucille P. Markey Foundation.

References

1. Speir, J.A., Munshi, S., Baker, T.S. & Johnson, J.E. (1993). Preliminary X-ray data analysis of crystalline cowpea chlorotic mottle virus. *Virology* **193**, 234–241.
2. Francki, R.I.B. (1985). The viruses and their taxonomy. In *The Plant Viruses.* (Fraenkel-Conrat, H. & Wagner, R., eds), pp. 1–18, Plenum Press, New York.
3. Lane, L.C. (1974). The bromoviruses. *Adv. Virus Res.* **19**, 151–220.
4. Dasgupta, R. & Kaesberg, P. (1982). Complete nucleotide sequences of coat protein messenger RNAs of brome mosaic virus and cowpea chlorotic mottle virus. *Nucleic Acids Res.* **10**, 703–713.
5. Baker, T.S. (1978). The packing of cowpea chlorotic mottle virus in crystalline monolayers. In *Proceedings of the Ninth International Congress on Electron Microscopy* **2**, 24–25.
6. Steven, A.C., Smith, P.R. & Horne, R.W. (1978). Capsid fine structure of cowpea chlorotic mottle virus: from a computer analysis of negatively stained virus arrays. *J. Ultrastruct. Res.* **64**, 63–73.
7. Bancroft, J.B. & Hiebert, E. (1967). Formation of an infectious nucleoprotein from protein and nucleic acid isolated from a small spherical virus. *Virology* **32**, 354–356.
8. Bancroft, J.B., Hills, G.J. & Markham, R. (1967). A study of the self-assembly process in a small spherical virus: formation of organized structures from protein subunits *in vitro*. *Virology* **31**, 354–379.
9. Adolph, K.W. & Butler, P.J.G. (1974). Studies on the assembly of a spherical plant virus. I. States of aggregation of the isolated protein. *J. Mol. Biol.* **88**, 327–341.
10. Verduin, B.J.M. (1978). Characterization of Cowpea Chlorotic Mottle Virus and its Assembly [Ph.D. Thesis]. University of Wageningen, Netherlands.
11. Fox, J.M., Johnson, J.E. & Young, M.J. (1994). RNA/protein interactions in icosahedral virus assembly. *Semin. Virol.* **5**, 51–60.
12. Adolph, K.W. & Butler, P.J.G. (1977). Studies on the assembly of a spherical plant virus. III. Reassembly of infectious virus under mild conditions. *J. Mol. Biol.* **109**, 345–357.
13. Bancroft, J.B. (1970). The self-assembly of spherical plant viruses. *Adv. Virus Res.* **16**, 99–134.
14. Jacrot, B. (1975). Studies on the assembly of a spherical virus. II. The mechanism of protein aggregation and virus swelling. *J. Mol. Biol.* **95**, 433–446.
15. Incardona, N.L. & Kaesberg, P. (1964). A pH-induced structural change in bromegrass mosaic virus. *Biophys. J.* **4**, 11–21.
16. Sorger, P.K., Stockley, P.G. & Harrison, S.C. (1986). Structure and assembly of turnip crinkle virus II. Mechanism of reassembly *in vitro*. *J. Mol. Biol.* **191**, 639–658.
17. Hsu, C.H., Sehgal, O.P. & Pickett, E.E. (1976). Stabilizing effect of divalent metal ions on virions of southern bean mosaic virus. *Virology* **69**, 587–595.
18. Rayment, I., Johnson, J.E. & Rossmann, M.G. (1979). Metal-free southern bean mosaic virus crystals. *J. Biol. Chem.* **254**, 5243–5245.
19. 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.
20. Olson, A.J., Bricogne, G. & Harrison, S.C. (1983). Structure of tomato bushy stunt virus IV. The virus particle at 2.9 Å resolution. *J. Mol. Biol.* **171**, 61–93.
21. Robinson, I.K. & Harrison, S.C. (1982). Structure of the expanded state of tomato bushy stunt virus. *Nature* **297**, 563–568.
22. Fricks, C.E. & Hogle, J.M. (1990). Cell induced conformational changes in poliovirus: Externalization of the amino terminus of VP1 is responsible for liposome binding. *J. Virol.* **64**, 1934–1945.
23. Filman, D.J., Syed, R., Chow, M., Macadam, A.J., Minor, P.D. & Hogle, J.M. (1989). Structural factors that control conformational transitions and serotype specificity in type 3 poliovirus. *EMBO J.* **8**, 1567–1579.
24. Hogle, J.M., Maeda, A. & Harrison, S.C. (1986). Structure and assembly of turnip crinkle virus. I. X-ray crystallographic structure analysis at 3.2 Å resolution. *J. Mol. Biol.* **191**, 625–638.
25. Abad-Zapatero, C., *et al.*, & Tsukihara, T. (1980). Structure of southern bean mosaic virus at 2.8 Å resolution. *Nature* **286**, 33–39.
26. Silva, A.M. & Rossmann, M.G. (1987). Refined structure of southern bean mosaic virus at 2.9 Å resolution. *J. Mol. Biol.* **197**, 69–87.
27. Fisher, A.J. & Johnson, J.E. (1993). Ordered duplex RNA controls capsid architecture in an icosahedral animal virus. *Nature* **361**, 176–179.
28. Hosur, M.V., *et al.*, & Rueckert, R.R. (1987). Structure of an insect virus at 3.0 Å resolution. *Proteins* **2**, 167–176.
29. Wery, J.P., Reddy, V.S., Hosur, M.V. & Johnson, J.E. (1994). The refined three-dimensional structure of an insect virus at 2.8 Å resolution. *J. Mol. Biol.* **235**, 565–586.
30. Johnson, J.E. *et al.*, & Baker, T.S. (1994). Comparative studies of *T*=3 and *T*=4 icosahedral RNA insect viruses. *Arch. Virol. (suppl)* **9**, 497–512.
31. Stehle, T., Yan, Y., Benjamin, T.L. & Harrison, S.C. (1994). Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature* **369**, 160–163.
32. Liddington, R.C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T.L. & Harrison, S.C. (1991). Structure of simian virus 40 at 3.8 Å resolution. *Nature* **354**, 278–284.
33. Vriend, G., Verduin, B.J.M. & Hemminga, M.A. (1986). Role of the amino-terminal part of the coat protein in the assembly of cowpea chlorotic mottle virus. A 500 MHz proton nuclear magnetic resonance study and structural calculations. *J. Mol. Biol.* **191**, 453–460.

34. Sacher, R. & Ahlquist, P. (1989). Effects of deletions in the amino-terminal basic arm of brome mosaic virus coat protein on RNA packaging and systemic infection. *J. Virol.* **63**, 4545–4552.
35. McPhalen, C.A., Strynadka, N.C.J. & James, M.N.G. (1991). Calcium-binding sites in proteins: a structural perspective. *Adv. Protein Chem.* **42**, 77–144.
36. Brisco, M., Hull, R. & Wilson, T.M.A. (1986). Swelling of isometric and of bacilliform plant virus nucleocapsids is required for virus-specific protein synthesis *in vitro*. *Virology* **148**, 210–217.
37. Caspar, D.L.D. & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1–24.
38. Williams, R. (1979). *The Geometric Foundation of Natural Structure.* Dover Publications Inc., New York.
39. Rossmann, M.G., *et al.*, & Vriend, G. (1985). Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* **317**, 145–153.
40. Hogle, J.M., Chow, M. & Filman, D.J. (1985). Three-dimensional structure of poliovirus at 2.9 Å resolution. *Science* **229**, 1358–1365.
41. Luo, M., *et al.*, & Palmenberg, A.C. (1987). The atomic structure of Mengo virus at 3.0 Å resolution. *Science* **235**, 182–191.
42. Harrison, S.C. (1984). Multiple modes of subunit association in the structures of simple spherical viruses. *Trends Biochem. Sci.* **9**, 345–351.
43. Chen, Z., *et al.*, & Johnson, J.E. (1989). Protein–RNA interactions in an icosahedral virus at 3.0 Å resolution. *Science* **245**, 154–159.
44. Larson, S.B., Koszelak, S., Day, J., Greenwood, A., Dodds, J.A. & McPherson, A. (1993). Double-helical RNA in satellite tobacco mosaic virus. *Nature* **361**, 179–182.
45. McKenna, R., *et al.*, & Incardona, N.L. (1992). Atomic structure of single-stranded DNA bacteriophage ϕX174 and its functional implications. *Nature* **355**, 137–143.
46. Tsao, J., *et al.*, & Parrish, C.R. (1991). The three-dimensional structure of canine parvovirus and its functional implications. *Science* **251**, 1456–1464.
47. Arnold, E., *et al.*, & Rossmann, M.G. (1987). The structure determination of a common cold virus, human rhinovirus 14. *Acta Crystallogr. A* **43**, 346–361.
48. Rossmann, M.G., *et al.*, & Lynch, R.E. (1992). Molecular replacement real-space averaging. *J. Appl. Crystallogr.* **25**, 166–180.
49. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458–460.
50. Brünger, A.T. (1993). *X-PLOR Version 3.1 Manual.* Yale University, New Haven, CT.
51. Baker, T.S., Newcomb, W.W., Olson, N.H., Cowsert, L.M., Olson, C. & Brown, J.C. (1991). Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional reconstruction. *Biophys. J.* **60**, 1445–1456.
52. Cheng, R.H., Olson, N.H. & Baker, T.S. (1992). Cauliflower mosaic virus: a 420 subunit (*T*=7), multilayer structure. *Virology* **186**, 655–668.
53. Dryden, K.A., *et al.*, & Baker, T.S. (1993). Early steps in reovirus infection are associated with dramatic changes in supramolecular structure and protein conformation: analysis of virions and subviral particles by cryoelectron microscopy and image reconstruction. *J. Cell Biol.* **122**, 1023–1041.
54. Fuller, S.D. (1987). The *T*=4 envelope of sindbis virus is organized by interactions with a complementary *T*=3 capsid. *Cell* **48**, 923–934.
55. Baker, T.S., Drak, J. & Bina, M. (1988). Reconstruction of the three-dimensional structure of simian virus 40 and visualization of the chromatin core. *Proc. Natl. Acad. Sci. USA* **85**, 422–426.
56. Winkelmann, D.A., Baker, T.S. & Rayment, I. (1991). Three-dimensional structure of myosin subfragment-1 from electron microscopy of sectioned crystals. *J. Cell Biol.* **114**, 701–713.
57. Ahlquist, P., Luckow, V. & Kaesberg, P. (1981). Complete nucleotide sequence of brome mosaic virus RNA3. *J. Mol. Biol.* **153**, 23–38.
58. Romero, J., Dzianott, A.M. & Bujarski, J.J. (1992). The nucleotide sequence and genome organization of the RNA2 and RNA3 segments of broad bean mottle virus. *Virology* **187**, 671–681.
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. Crystallogr.* **26**, 283–291.

Received: **25 Aug 1994**; revisions requested: **20 Sep 1994**; revisions received: **17 Oct 1994**. Accepted: **2 Nov 1994**.