Structure and assembly of *Sesbania mosaica* virus

H. S. Savithri\(^1\) and M. R. N. Murthy\(^2\,*\)

---

*Sesbania mosaica* virus (SeMV) is a ss-RNA (4149 nt) plant sobemovirus isolated from farmer’s field around Tirupathi, Andhra Pradesh. The viral capsid (30 nm diameter) consists of 180 copies of protein subunits (MW 29 kDa) organized with icosahedral symmetry. In order to understand the mechanism of assembly of SeMV, a large number of deletion and substitution mutants of the coat protein (CP) were constructed. Recombinant SeMV CP (rCP) as well as the N-terminal rCP deletion mutant D\(\Delta\)N65 mostly formed smaller particles consisting of 60 protein subunits. Although particle assembly was not affected due to the substitution of aspartates (D146 and D149) that coordinate calcium ions by asparagines, the stability of the resulting capsids was drastically reduced. Deletion of residues forming a characteristic B-annulus at the icosahedral 3-folds did not affect the assembly of VLPs. Mutation of a single tryptophan, which occurs near the icosahedral five-fold axis to glutamate or lysine, resulted in the disruption of the capsid leading to soluble dimers that resembled the quasi-dimer structure of the native virus. Replacement of positively charged residues in the amino terminal segment of CP resulted in the formation of empty shells. Based on these observations, a plausible mechanism of assembly is proposed.

Keywords: Coat protein dimer, icosahedral particles, mechanism of assembly, *Sesbania mosaica* virus.

Introduction

Viruses are obligate parasites that have proteinaceous capsids that encapsidate and protect their cognate genetic material. Small isometric viruses have a precise molecular structure. Therefore, their structures can be determined by single-crystal X-ray diffraction studies. Structural studies on several viruses at or near atomic resolution have provided detailed information on the particle size, radial distribution of nucleic acid and protein, morphology of particles, three-dimensional structure of coat protein, capsid architecture, distribution of residues on the outer and inner surfaces, molecular interactions between protein subunits and role of RNA–protein and metal ion mediated protein–protein interactions in the stability of the capsids. Although structures of several viruses are known, the precise mechanisms of their capsid assembly are largely unknown. This is due to the fact that virus particle assembly is highly co-operative and hence it is difficult to isolate or characterize intermediates of the assembly pathway.

It is well known that only 60 subunits can be organized with exact symmetry to form a capsid such that the environments of all subunits are identical. Coat proteins (CP) of viruses such as satellite tobacco necrosis virus consist of only 60 subunits with exact icosahedral symmetry and are referred to as \(T = 1\) viruses. However, several plant and animal virus capsids consist of 180 chemically identical CP subunits. These \(T = 3\) viruses consist of 12 pentameric and 20 hexameric protein subunits\(^1\). Although the 180 polypeptides of \(T = 3\) capsids are identical, they occur in three slightly different bonding environments. Therefore, these capsids contain 60 copies each of A, B, C type subunits that occupy quasi-equivalent positions. The architectures of \(T = 1\) and \(T = 3\) viruses are shown in Figure 1. \(T = 3\) viruses contain two types of dimers A/B and C/C. The C/C dimers are related by an exact 2-fold symmetry (icosahedral 2-fold) while A/B subunits are related by a quasi 2-fold symmetry. In contrast, \(T = 1\) viruses have only one type of dimers related by icosahedral 2-fold symmetry.

During the past two decades, we have conducted a large number of studies on the structure and assembly of a small spherical plant virus, Sesbania mosaic virus (SeMV). SeMV belongs to the genus *Sobemovirus* and infects *Sesbania grandiflora* belonging to Fabaceae. The virus was discovered in farmer’s fields near Tirupathi, Andhra Pradesh. We have determined the genomic sequence\(^2\) and the three-dimensional structure of the native virus at 3.0 Å resolution\(^3\). SeMV genome is covalently linked at the 5′-end to a viral coded genome linked protein (VPg) and encodes four potential open reading frames. The CP, coded by the 3′-terminal open reading frame, is expressed by a subgenomic messenger. The native virus particles are remarkably stable with a \(T_m\) of 91°C. Fluorescence studies have shown that even at 6 M GuHCl, native virus particles do not undergo dissociation or denaturation\(^5\). Removal of bound calcium ions by treatment with EGTA or EDTA leads to swelling or slight expansion of SeMV particles. The swollen particles are

---

\(^*\)For correspondence. (e-mail: mrn@mbu.iisc.ernet.in)
Figure 1. Organization of protein subunits in $T=1$ and $T=3$ icosahedral viruses. All the 60 subunits in $T=1$ capsids are in identical surroundings whereas in $T=3$ particles, three distinct bonding arrangements are found. The icosahedral asymmetric unit consists of three chemically identical subunits in quasi-equivalent environments (A, B and C). Small differences might exist in the conformation of the three polypeptides that might be crucial for the correct assembly of $T=3$ particles.

Figure 2. SeMV coat protein structure revealed by X-ray crystal structure determination at 3.0 Å resolution illustrated as a cartoon diagram. Subunits A and B of the icosahedral asymmetric unit have very similar conformation (Figure 1a). The conformation of the C subunit (Figure 1b) shows significant differences from those of A and B subunits, particularly in the amino terminal segment of the polypeptide. Protein subunits consist of three distinct structural features. A large arginine rich segment at the N-terminus (N-ARM) is disordered. This is followed by a segment ordered only in one-third of the subunits (C subunits) and form a characteristic hydrogen bonded structure at the quasi 6-fold (icosahedral 3-fold) axes (see Figure 4). The remainder of the C-terminal polypeptide folds in to an eight-stranded antiparallel $\beta$-barrel with ‘jelly roll’ topology. This fold is also found in a large number of unrelated icosahedral viruses.

less stable, suggesting that divalent metal ion-mediated interactions are essential for the stability of the particles. Further incubation of the swollen virus in 4 M LiCl disrupts the particles into CP dimers. The swollen virus and the dissociated dimers are sensitive to treatment with proteases due to the exposure of the amino terminal segment, which is rich in arginines and lysines.

A careful analysis of the three-dimensional structure of the virus revealed segments of the polypeptide that might be important for the assembly of virus particles. The SeMV CP expressed in E. coli assembles into virus-like particles (VLPs) that are nearly identical to the native virus particles. With the view of understanding the mechanism of assembly, we have constructed a large number of CP mutants and overexpressed them in E. coli. In this review, we briefly describe the structural and stability studies carried out on the wild type and mutant VLPs and CP subunits and discuss their implication to SeMV particle assembly.

Structure of SeMV and its implication to assembly

The structure of native SeMV particles was determined at 3.0 Å resolution. As in other $T=3$ viruses, the capsid consists of 180 coat protein subunits (CP, 29 kDa). The CP adopts a jellyroll sandwich fold consisting of eight antiparallel $\beta$-strands connected by loops and a few helices (Figure 2). Similar polypeptide folds are found in several other plant and animal viruses. The asymmetric unit of the capsid is composed of three chemically identical subunits (named A, B and C) arranged in quasi-equivalent environments (Figure 3). The A subunits form pentamers at the icosahedral 5-fold axes while the B and C subunits from hexamers at the icosahedral 3-fold (quasi 6-fold) axes in the capsid. An icosahedral 2-fold relates two C type subunits while a quasi 2-fold relates A and B
subunits. The amino terminal arms of the C subunits are ordered from residue 44 while in the A and B subunits they are ordered from residue 73. The ordered amino terminal arms of C type subunits form a β-annulus-like structure at the quasi 6-fold axes (residues 48–59; Figure 4). The β-annulus and the dimeric interactions between icosahedral 2-fold related protein subunits form a continuous scaffold connecting all C subunits of the capsid. Therefore, it was thought that the β-annulus is instrumental for the error-free assembly of \( T = 3 \) particles. The asymmetric unit of the icosahedral particle also contains three calcium ions (Figure 3, ref. 3), accounting for a total of 180 calcium ions on the capsid. These ions related by quasi 3-fold symmetry are located at the A–B, B–C and C–A subunit interfaces. The calcium is coordinated in an octahedral geometry with six ligands (Figure 5). Two of the ligands are the carboxylate oxygens of D146 and D149 from one subunit and three others are from the neighbouring subunit (carbonyl oxygen atoms of Y205, carboxyl oxygen of N267 and C-terminal carboxyl oxygen of N268). The sixth ligand, represented by a weak electron density in most of the structures determined so far, is a molecule of water hydrogen bonded to S116. The amino acid sequence constituting the calcium binding motif (CBS) is highly conserved across sobemoviruses and other related viruses suggesting that the calcium binding and the associated ligands might be essential for the integrity or assembly of the capsids. A fourth ion in the icosahedral asymmetric unit is on the quasi 3-fold axis and its identity has not been clearly established. The disordered segment of the amino terminal arm contains an arginine rich motif (N-ARM). The positive charges in the 180 ARMs of the capsid could neutralize nearly a third of the negative charges of RNA phosphates. This motif is present in several viruses, although the precise position of its occurrence in the polypeptide is variable. The role of this arginine rich motif, if any, in the assembly of SeMV needed to be established.

The three-dimensional structure revealed four aspects of the structure that may be crucial for the assembly of SeMV: (i) the arginine rich motif, (ii) the β-annulus, (iii) calcium binding motif and (iv) inter-subunit interactions.

Figure 3. The organization of the three subunits A, B and C of the icosahedral asymmetric unit. The polypeptide folds are shown as ribbon diagrams. The amino terminal segment ordered only in C subunits is at the left lower edge of the figure. Spheres correspond to ions bound to the capsids. The ions bound at the inter-subunit interfaces are calciums. The nature of the ion bound on the quasi 3-fold axis is not known.

Figure 4. β-annulus: The amino terminal arms of three C subunits (shown in different grades of grey) related by the icosahedral 3-fold axis are involved in an intricate pattern of hydrogen bonding. These interactions, along with the interactions across the icosahedral 2-fold related C subunits link all C subunits of the capsid. This is a feature shared by several viruses.

Figure 5. The geometry of calcium binding. This geometry is common to both \( T = 3 \) and \( T = 1 \) capsids.
promoted by specific residues. We probed the importance of each of these interactions by site-directed mutagenesis followed by structure and assembly studies.

Structure and assembly of recombinant capsids of SeMV

Over expression of SeMV-CP gene in E. coli was shown to result in the assembly of $T=3$ VLPs resembling native particles. VLPs encapsidate CP mRNA and E. coli 23S rRNA. It may be noted that the size of the 23S RNA is comparable to that of the cognate viral RNA. The thermal stability of VLPs ($T_m \approx 87^\circ C$) was only slightly less when compared to that of the wild type particles ($T_m \approx 91^\circ C$). Treatment with EDTA reduced the sedimentation coefficient of VLPs and increased their sensitivity to treatment with proteases showing that the VLPs are also stabilized by calcium-mediated protein–protein interactions. GuHCl induced denaturation profile of VLPs showed that they were less stable when compared to the native virus although even here denaturation was not complete at 6 M GuHCl (ref. 4).

$T=3$ capsids of intact VLPs and VLPs formed by the N-terminal deletion mutant (CP-NA22) could be crystallized. The crystals of these intact and mutant VLPs were isomorphous. A substitution mutant of the intact CP, CP-P53A, in which a conserved proline at position 53 was substituted by alanine (CP-P53A) also assembled into $T=3$ capsids and was crystallized. This proline occurs in the $\beta$-annulus at which there is a characteristic kink that might be important for the formation of the annulus structure. The structures of rCP, CP-NA22 and CP-P53A were determined at 3.6, 5.5 and 4.1 Å respectively.

The structures of these recombinant capsids closely resemble that of the wild-type particles. Inter-subunit interactions and binding of calcium ions at the inter-subunit interfaces of the A, B and C subunits are also similar in these capsids. As in the native capsid, calcium ions adopt an octahedral coordination. At the quasi 6-fold axis, the $\beta A$ arms of the three C subunits interact to form the $\beta$-annulus as in the native capsids. In CP-P53A, the substitution of the conserved proline by alanine at position 53 does not seem to affect either the bending or conformation of the $\beta$-annulus. The structures of these capsids indicate that differences in the nucleic acid that is encapsidated, deletion of short segments from the amino terminus and substitution of a conserved proline at the $\beta$-annulus do not affect the assembly of VLPs.

Assembly of truncated CP to particles of smaller diameter and their structures

Deletion of longer segments from the amino terminus resulted in the formation of VLPs of smaller diameters. Polypeptides obtained by the deletion of 31 (CP-NA31) and 36 residues (CP-NA36) from the amino terminus assembled mostly into particles of smaller diameter of ~20 nm. The smaller particles had only 60 copies of subunits ($T=1$ capsids; Figure 1). Close examination of assembly products formed by CP-NA36 in an electron microscope revealed not only particles of ~20 nm diameter but also particles that were intermediate in size to those of $T=1$ and $T=3$ particles. The sedimentation coefficient of this component was also intermediate to those of 20 and 30 nm particles. Therefore, these VLPs are likely to possess pseudo $T=2$ morphology with 120 subunits. It was shown that the deletion of the N-terminal 65 amino acid residues (CP-NA65) resulted in the exclusive formation of $T=1$ particles. CP-NA65 $T=1$ particles were slightly less stable than those of $T=3$ particles with a $T_m$ of 83°C and stability up to 5 M GuHCl. As in the case of $T=3$ particles, EDTA treated $T=1$ particles were less stable and half of the particles dissociated at 2 M GuHCl suggesting the role of calcium ions in the stability of smaller particles also.

The structure of the $T=1$ particles formed by CP-NA65 was determined at 3.0 Å resolution. This recombinant protein possesses the sequence corresponding to the S-domain of the native $T=3$ icosahedral particles and lacks the $\beta$-annulus, the $\beta A$ strand (residues 67–70) and the arginine rich motif (ARM; residues 28–36; ref. 10). The polypeptide fold of the subunit closely resembles that of the S-domain of the native virus. The structure of the icosahedral dimeric unit in the $T=1$ structures closely resembles the structure of the quasi-dimer of the $T=3$ structure. The recombinant particles bind calcium ions in a manner indistinguishable from those of native capsids. Dimers resembling icosahedral CC dimers of $T=3$ capsids are not observed in the $T=1$ capsids (Figure 1).

In the native virus particles, the N-terminal 43 residues are disordered in all the three subunits of icosahedral asymmetric unit. Therefore, it was anticipated that deletion of these N-terminal residues would not affect assembly into native like $T=3$ particles. Surprisingly, CPAN31 and CPAN36 mutants assembled mostly into $T=1$ particles similar in structure to particles assembled from N65 protein. The ‘unseen’ or disordered amino terminal segment of the coat protein seems to exert a strong influence on the particle assembly. This suggests that the interaction between the amino terminus and RNA may be crucial for initiating $T=3$ particle assembly.

Recombinant capsids with deletion or substitution of calcium ligands

The three-dimensional structure of SeMV shows that the two carboxy terminal residues contribute to calcium binding. Deletion of two residues the carboxy terminal did not prevent assembly of $T=3$ particles. However, the stability of the resulting capsids (CP-CA2) was drastically reduced.
reduced. Although it was possible to crystallize CP-CA2, the crystals did not diffract X-rays, perhaps reflecting the lack of rigid structure of these capsids. Polypeptides obtained by the deletion of these two residues in NΔ65 (CP-NΔ65-CA2) could assemble into VLPs of approximate diameter 20 nm. These VLPs were less stable with a denaturation profile similar to that of CP-NΔ65 treated with 2 mM EDTA suggesting that calcium ions are not bound to these VLPs. Similarly, CP-NΔ65-D146N-D149N, where the calcium ligands Asp146 and Asp149 were replaced by Asn assembled into homogeneous particles of ~20 nm diameter similar to those of CP-NΔ65. However, these particles did not bind calcium. The structure of this mutant revealed particles that were slightly expanded. In contrast to the uniform particles observed with the calcium ligand substitution mutants of the truncated polypeptide CPNΔ65, similar mutations in the full length polypeptide (CP-D146N and CP-D146N-D149N) resulted in particles of heterogeneous morphology. These observations demonstrate that calcium ions are more crucial for the stability of the larger T = 3 particles when compared to T = 1 particles, although in neither case they are mandatory for assembly

N-terminal ARM and β-annulus deletion mutants

The structure of SeMV and other related viruses had suggested that the amino terminal segment of the polypeptide is likely to play a major role in the assembly of virus particles. In order to address the role of these segments in virus assembly, a number of mutants were constructed. Substitution of a few arginine residues (CP-R32-36E) or all arginine residues (CP-R28-36E) to Glu in the basic ARM led to less stable, empty T = 3 capsids (CP-R28-36E), suggesting that the positively charged ARM is essential for the neutralization and encapsidation of the nucleic acid and not for T = 3 particle assembly.

Although the β-annulus structure found in several icosahedral viruses has been regarded as a crucial element in the assembly of virus particles, its importance was not experimentally examined earlier. In SeMV, β-annulus is formed by hydrogen bonding interactions between residues 48–52 from one C-subunit and residues 55–59 of the neighbouring three-fold related C-subunit (Figure 4). To examine the importance of the β-annulus, the residues involved were partially (CP-ΔA48-53) or fully deleted (CP-ΔA48-59). Crystal structure of the mutant rCP-ΔA48-59 which lacked the whole segment forming the annulus was determined. Contrary to expectations, in this mutant, the capsid assembly and stability were unaffected. The structure revealed that indeed the β-annulus is absent. Therefore, capsid assembly can proceed without the formation of β-annulus and its formation is probably a consequence of the assembly.

Mutational studies on interfacial residues of SeMV coat protein

SeMV particles are constructed from two types of dimers: one related by icosahedral 2-fold and the other related by quasi 2-fold. Disassembly of SeMV leads to the formation of coat protein dimers. It was of interest to determine the structure of isolated dimeric units and compare the inter-subunit packing to the corresponding contacts in T = 1 and T = 3 particles. With this view, the key interactions between subunits in the T = 3 virus were disrupted by mutation of crucial residues at these interfaces. Tryptophan 170 (W170), which occurs at a site close to the icosahedral 5-fold (Figure 6), and which constitutes a cluster of large hydrophobic residues at the 5-fold interface was targeted for mutagenesis. The mutations were carried out both in rCP and CP-NΔ65.

Gel filtration studies showed that CP-W170E, CP-W170K, CP-NΔ65-W170E, CP-NΔ65-W170K mutants were indeed expressed as dimers. The purified CP dimers were prone to protease degradation, probably because of the exposure of the arginine rich N-terminus. This was further confirmed by the observation that CP-NΔ65-W170E and CP-NΔ65-W170K, where the N-terminus is absent, did not show any signs of degradation. Studies using CD showed that CP-NΔ65-W170E was the most stable among the mutants with a melting temperature of about 65°C. The other mutants had a Tm around 45°C, which was similar to that of the dimers obtained by lithium chloride induced dissociation. Crystal structure of the dimer resulting from W170K mutation was determined. The structure of the isolated dimers was found

Figure 6. Cluster of tryptophan residues around the icosahedral 5-fold axes. Mutation of these residues to lysine or glutamic acid leads to the disruption of the capsids. The resulting protein is in dimeric form.
Figure 7. A plausible model for the assembly of SeMV. Initially the protein subunits (with disordered amino termini) assemble into a pentamer of A/B dimers. Interaction of the flexible amino terminal ARM with RNA leads to the formation of CC dimers and ordered β-annulus. Interactions with RNA is disrupted in the CP R28-36E mutant, which forms only empty \( T = 3 \) particles. If the amino terminus is truncated, only \( T = 1 \) or pseudo \( T = 2 \) particles are formed. Calcium binds to the assembling or assembled particles resulting in a slight reduction in the particle radius.

to resemble the A/B dimer and not the C/C dimer of the \( T = 3 \) virus particles. It also revealed a number of structural changes that take place, especially in the loop and interfacial regions, during the course of assembly.

SeMV assembly

A model for the assembly of SeMV particles (Figure 7) could be proposed on the basis of the large number of studies summarized in this review: structures of \( T = 3 \) and \( T = 1 \) particles formed by intact protein and its deletion and substitution mutants, calcium ligand and interfacial residue mutants, and isolated dimeric structure of the coat protein. The analysis of the structures of these mutant capsids suggests that although calcium binding contributes substantially to the stability of the particles, it is not mandatory for their assembly. Similarly, the structure of the \( \beta \)-annulus deletion mutant provides clear evidence that the \( \beta \)-annulus is not necessary for the formation of the \( T = 3 \) capsids. In contrast, the presence of a large fraction of the amino terminal arm including sequences that precede the \( \beta \)-annulus appears to be indispensable for the error free assembly of \( T = 3 \) particles. When mutations that disrupt particle assembly such as W170K are introduced, the CP is expressed in a dimeric form that resembles the dimer of \( T = 1 \) particles and the quasi symmetric A/B dimer of \( T = 3 \) particles. Therefore, it is likely that a pentamer of A/B dimers in which the \( \beta \)-annuli are disordered could initiate the assembly of the capsids at the icosahedral five-fold. Further assembly proceeds in the presence of RNA. Interaction of the ARM with RNA imposes order in the amino terminal segments of dimers added to the 10-mer complex leading to the formation of the \( \beta \)-annulus and C/C dimers. Subsequent addition of CP dimers leads to the formation of swollen \( T = 3 \) particles. The particles become compact after the addition of calcium ions at the intersubunit interfaces.


ACKNOWLEDGEMENTS. We thank H. S. Subramanya, K. Gopinath, P. S. Sateeshkumar, G. L. Lokesh, T. D. S. Gowri, C. S. Vijay, V. Saravanazh, Subashchandrabose Chinnatambi, M. Bhuvaneswari, V. Sangita, Anju Papacchan, P. Gayathri and Smita Nair for their contributions. We also thank the Department of Science and Technology and the Department of Biotechnology, Government of India for financial support.