Virus Research 273 (2019) 197755

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Elucidating the functional aspects of different domains of bean common mosaic virus coat protein

Sunil Kumar[a], Ruma Karmakar[b], Dushyant Kumar Garg[c], Ishu Gupta[a], Ashok Kumar Patel[a,*]

[a] Kusuma School of Biological Sciences, Indian Institute of Technology Delhi, New Delhi, 110016, India
[b] Centre for Rural Development and Technology, Indian Institute of Technology Delhi, New Delhi, 110016, India
[c] School of Biotechnology, Jawaharlal Nehru University, New Mehrauli Road, New Delhi, 110067, India

ARTICLE INFO

Keywords:
Coat protein
Virus assembly
N-terminal
RNA binding pocket
3′ untranslated region

ABSTRACT

The coat protein (CP) is the only structural protein present in the polyprotein of bean common mosaic virus. The well known characteristics of the CP are self-oligomerization and nucleic acid binding activity. The studies of the coat protein mutants revealed that the oligomeric property of CP solely depends on the amino-terminal residues and the nucleic acid binding domain present at the 194–202 residue position. The 3′UTR RNA of the virus showed high binding affinity with the RNA binding domain as compared to the 5′UTR RNA. Further, the intrinsic fluorescence study of the CP also suggested that the N- and C-terminal of CP contains a highly disordered region. The present study also illustrates that the coat protein contains a conserved RNA binding pocket among the potyviruses, but displays divergent oligomerization propensities due to the difference in residue at the N- and C-terminal.

1. Introduction

The coat protein (CP) is the most abundant structural protein, plays a key role during the life cycle of a virus by two crucial steps *viz.*, assembly and disassembly. During bean common mosaic virus (BCMV) virion formation, the CP monomer units oligomerize to assemble in helical manner around the viral RNA, which provides protection to viral RNA from the external environment (Urcuqui-Inchima, 2001). The assembled CP subunits can be dissociated at a high ionic strength and extreme pH, but they can be re-assembled again into flexuous rod-shaped structure *via* aggregating into a ring-like intermediate structure at an optimum pH (Anindya and Savithri, 2003; Jagadish et al., 1991). The interaction between CP subunits facilitate oligomerization through electrostatic interaction between the residues present at the amino- and carboxy-terminal ends of the CP (Kendall et al., 2008; Zamora et al., 2017).

All potyviruses genome consist of a single positive-sense ssRNA of ~10 kb in length which is encapsulated by approximately 2000 identical copies of coat protein (CP) (Anindya and Savithri, 2003). The genome of potyvirus encodes a single polyprotein of 340–370 kDa which later cleaves into 10 proteins *viz.* proteinase (P1), helper component proteinase (HC-Pro), proteinase (P3), 6-kDa peptide 1 (6K1), cylindrical inclusion (CI), 6-kDa peptide 2 (6K2), viral protein genome-linked (VPg), nuclear inclusion a (NIa), nuclear inclusion b (NIb) and CP (Revers and García, 2015). The polyprotein undergoes self-cleavage into ten mature proteins through proteolytic activity of P1, HC-Pro, and NIa-Pro. The proteolytic digestion of eight out of the ten polyprotein digestion is carried out by NIa-Pro, whereas, P1 and HC-Pro undergo self-cleavage at the P1/HC-Pro and HC-Pro/P3 junction, respectively (Worrall et al., 2015). The encapsidation of potyviral genomic RNA occurs inside the core of CP that forms characteristic flexuous rod shape structure, which is 11–20 nm in diameter with 680–900 nm length (Zamora et al., 2017).

The potyvirus coat protein (CP) is associated with different functions during the infectious cycle, such as viral RNA translation, replication, movement and encapsidation (Besong-Ndika et al., 2015). Virus structure formation begins with the CP-CP as well as the CP-RNA interaction. In addition to the virion packaging, another main function of CP is to carry a viral RNA from cell to cell and over long distances (Dolja et al., 1994). The replication as well as the translation mechanism of the viral RNA is regulated through the amount of CP present in a cell (Ivanov and Mäkinen, 2012). The viral RNA replication is suppressed at a higher CP concentration, whereas the RNA accumulation and translation is enhanced at a lower concentration in potyviruses (Ivanov et al., 2003).

Understanding the mechanisms through which the viruses wrap up their genomes into a stable nucleoprotein structure is a foremost prerequisite for elucidating the replication cycle in plants. The present

* Corresponding author.
E-mail address: ashokpatel@bioschool.iitd.ac.in (A.K. Patel).

https://doi.org/10.1016/j.virusres.2019.197755
Received 10 June 2019; Received in revised form 10 September 2019; Accepted 10 September 2019
Available online 13 September 2019

study, deals with the characterization of the different domains of BCMV coat protein through biochemical approach. This study addresses two major questions/research gaps, firstly, what is the oligomeric behavior of the CP and secondly, how the CP interacts with the RNA to form the BCMV virion structure.

2. Materials and methods

2.1. Virus source and coat protein gene isolation

French bean (*Phaseolus vulgaris*) seeds were sown in sterilized soil and kept in an insect-free growth chamber. Seedlings were raised in earthen pots at a frequency of 5 seeds per pot. Seedlings were monitored for disease incidence from primary leaf stage, first trifoliate to third trifoliate stage. Total RNA was extracted from 100 mg bean common mosaic virus (BCMV) infected leaves using TRIzol® LS Reagent (Invitrogen). The isolated RNA was used for cDNA synthesis using ImProm-II™ Reverse Transcription kit (Promega) according to the manufacturer's protocol. The coat protein gene was PCR amplified from cDNA using gene-specific primers (Table S1). The amplified coat protein gene was cloned into expression vector pET28a at multicloning sites using NdeI and HindIII restriction enzymes to obtain pET28a-CP construct. The coat protein gene in pET28a vector was confirmed by Sanger sequencing.

2.2. Purification of recombinant coat protein

The pET28a-CP clone was transformed into *Escherichia coli* BL21 (DE3) *Rosetta* cells (Novagen). A single colony of transformed cells was inoculated in 10 ml of Luria-Bertani medium (LB) containing 50 mg ml^{-1} kanamycin and 35 mg ml^{-1} chloramphenicol. The culture was grown at 37 °C for overnight. The overnight culture was inoculated to 1 L auto-induction media (1% Bacto-tryptone, 0.5% yeast extract, 25 mmol/L Na_2HPO_4, 25 mmol/L KH_2PO_4, 50 mmol/L NH_4Cl, 5 mmol/L Na_2SO_4, 2 mmol/L $MgSO_4$, 0.5% glycerol and 0.05% glucose) containing appropriate antibiotics and culture was induced by auto-induction method as described earlier (Studier, 2014). The cells were pelleted by centrifugation and stored at −80 °C until further use. Frozen cell pellets were resuspended in 50 ml of lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 1 mM dithiothreitol and 10 mM imidazole) containing protease inhibitor cocktail (Complete, EDTA free; Roche). The cells were lysed through sonicator for 30 min at 30% amplitude and cell debris was pelleted down by centrifugation at 10,000 × *g* for 30 min at 4 °C. The supernatant was passed onto a pre-equilibrated nickel-affinity column (His-Trap HP; GE Healthcare) with wash buffer (50 mM Tris pH 7.5, 300 mM NaCl, 10 mM imidazole and 1% Triton X-100). The bound N-terminal his-tagged coat protein was eluted at 0–60% of imidazole gradient corresponding to 250–300 mM imidazole. The eluted fractions were analyzed on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) for presence of coat protein. The eluted fractions were treated with PreScission protease for removal of his-tag. After PreScission protease digestion, fractions were passed through the Ni-NTA column for elimination of his-tag. The coat protein containing fractions were concentrated by ultrafiltration membrane (10 K, Millipore) and injected on a gel filtration column (Superdex 200; GE Healthcare) to analyze the characteristics of purified recombinant coat protein. For identification of coat protein, the protein bands corresponding to coat protein molecular weight were excised from SDS-PAGE which were trypsin-digested and the resultant peptides were analyzed with an ABI 4700 matrix-assisted laser desorption ionization–time of flight tandem mass spectrometer (MALDI-TOF/TOF MS). To further ascertain the oligomeric nature of the CP, its Native-PAGE analysis was carried out at two different concentrations (0.2 and 0.5 mg ml^{-1}). The samples were loaded onto a pre-run 6% native polyacrylamide gel in 1X TBE buffer (0.9 M Tris pH 7.5, 1.125 M boric acid and 20 mM EDTA). After electrophoresis at 100 V for 60 min, the gel was stained with coomassie brilliant blue and visualized after destaining using gel documentation system (Biorad).

2.3. Glutaraldehyde cross-linking

Cross-linking experiment was performed for analysis of coat protein subunit. The recombinant coat protein was incubated with 10 mM glutaraldehyde in reaction buffer (25 mM phosphate buffer pH 7, 100 mM NaCl) in a time dependent manner. The cross-linking reactions were stopped by 20 mM Tris − HCI buffer (pH 8.0) and reaction products were analyzed on 12% SDS–PAGE.

2.4. Equilibrium unfolding experiments

The CP was subjected to unfolding at varying concentration of denaturant guanidine hydrochloride (GdnCl). An 8 M stock of (GdnCl) was prepared in CP buffer (50 mM Tris, pH 7.5, 100 mM NaCl). The CP was incubated overnight with varying concentration of GdnCl in aforementioned buffer such that final protein concentration was 0.2 mg ml^{-1}. The structural signatures of the incubated samples were analyzed through following methods:

a) Fluorescence spectroscopy: All fluorescence experiments were carried out on Varian Cary Eclipse spectrofluorometer (Varian, Inc. USA) equipped with a Varian Cary temperature controller (Peltier multiple holder). The excitation wavelength was 295 nm for intrinsic tryptophan fluorescence, and spectra were collected over a wavelength range of 310–400 nm. The excitation and emission slits were 5 and 10 nm, respectively and scan speed was 100 nm min^{-1}. The protein concentration used was 0.2 mg ml^{-1} and the pathlength of the cuvette was 10 mm. The buffer background was subtracted from all spectra.

The extent of hydrophobic patches exposure was quantified by incubating the protein with 8-Anilino-1-naphthalenesulfonic acid (ANS). A stock solution of ANS was prepared in milliQ water and concentration was determined using molar extinction coefficient of 5000 M^{-1} cm^{-1} at 350 nm. Protein samples were incubated with ANS such that final protein and ANS concentration were 0.2 mg ml^{-1} and 40 μM, respectively. The mixture was incubated for 5 min in dark and fluorescence was measured over a wavelength range of 400–600 nm.

b) Far UV CD spectroscopy: Circular dichroism (CD) measurements were carried out on a J-815 CD spectropolarimeter (JASCO Corporation, Tokyo, Japan) equipped with a peltier multiple holder. Conformational changes in the secondary structure of protein were monitored in the region between 200 and 250 nm in a quartz cuvette (Starna, Essex, England) with a path length of 1 mm. The scanning speed, band width and data pitch were 50 nm min^{-1}, 1 nm and 1 nm, respectively. Two accumulations of each scan (within 600 H T voltage range) were averaged.

2.5. Purification of coat protein mutants

The coat protein mutants ΔN37CP, ΔNΔC37CP and ΔN37ΔC46CP were PCR amplified from wild type pET28a-CP clone using different sets of primers (Table S1). The amplified genes of mutant ΔN37CP, ΔNΔC37CP and ΔN37ΔC46CP were cloned in pET28a vector as mentioned above. The expression vector pET28a containing ΔN37CP, ΔNΔC37CP and ΔN37ΔC46CP genes were transformed into E. coli BL21 (DE3) *Rosetta* cells (Novagen) and were subsequently purified as mentioned in previous method. The purified recombinant ΔN37CP, ΔNΔC37CP and ΔN37ΔC46CP mutant proteins were loaded onto gel filtration column (Superdex 200; GE Healthcare) to analyze their hydrodynamic properties.

2.6. RNA-binding activity

The 5′ UTR and 3′ UTR genes were PCR amplified from cDNA of BCMV using gene-specific primers containing T7 promoter sequence

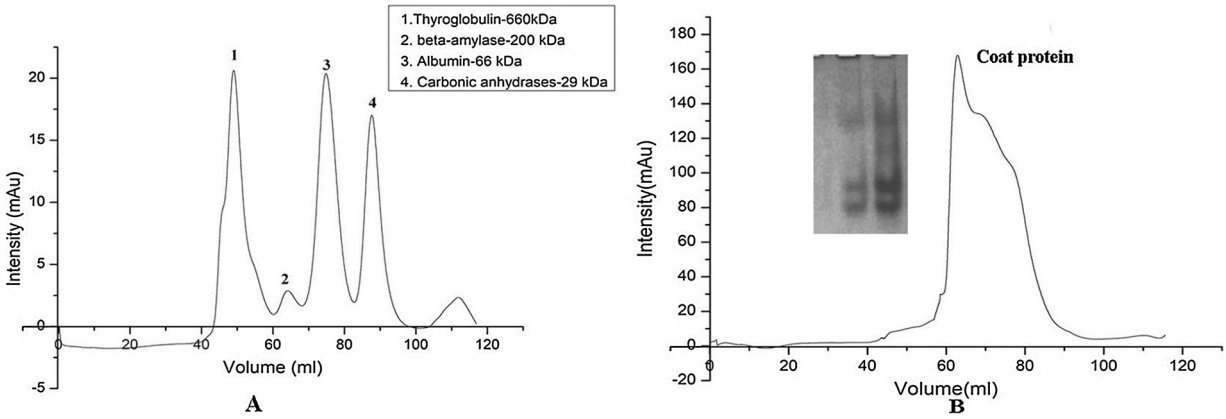

Fig. 1. Wild Type coat protein was analyzed on size-exclusion chromatography (SEC): (A) The SEC analysis of protein standards. Molecular mass markers (Thyroglobulin: 669 kDa; β-Amylase: 223 kDa; Albumin: 66 kDa; Carbonic anhydrase: 29 kDa). (B) Wild Type coat protein showed retention volume between 60 ml–90 ml. The Native-PAGE analysis at 0.2 and 0.5 mg ml^{-1} further reveal the oligomeric nature of CP.

(Table S2). The purified PCR products of 5′ UTR and 3′ UTR genes were used as template for *in-vitro* transcription. The 5′ UTR and 3′ UTR RNAs were transcribed using a HiScribe™ T7 In-Vitro Transcription Kit (New England Biolabs) according to the manufacturer's protocol. The EMSA experiment was carried out to study the RNA binding activity of mutant ΔNΔC37CP and ΔN37ΔC46CP proteins. The mutant proteins were mixed in a binding buffer (20 mM HEPES pH 7.5, 85 mM KCl, 1 mM $MgCl_2$, 1 mM DTT) with 1:1 M ratio of UTR RNAs for 1 h at 4 °C. The mutant protein-RNA complexes were analyzed on a pre-run 6% native polyacrylamide gel as mentioned above. After electrophoresis at 100 V for 60 min, the gel was stained in a solution of ethidium bromide and documented with Gel Imaging System (Bio-Rad).

2.7. Homology modeling and RNA-protein docking

The three dimensional (3D) structural model of mutant ΔNΔC37CP was elucidated using RaptorX (Wang et al., 2016) and I-TASSER (Roy et al., 2010) protein tertiary structure prediction server. The structure prediction was performed on the basis of sequence similarity by both the server which used coat protein structure of watermelon mosaic virus potyvirus (PDB ID: 5ODV) (Zamora et al., 2017) as a template. Stereo-chemical properties of the generated 3D model was examined through PROCHECK software (http://swissmodel.expasy.org/) (Laskowski et al., 1993) to assess geometry of the residues in protein structure. The RNA secondary structure as well as RNAfold of 3′UTR was observed through ViennaRNA webserver (Gruber et al., 2008). The 3D structure of 3′UTR RNA was simulated using SimRNA server (Boniecki et al., 2015). The predict RNA model was elucidated on the basis of Monte Carlo method which explained conformational space, potential energy, secondary structure with high accuracy and prediction of pseudoknots. The 3D model of 3′ UTR RNA was simulated with 3D model of mutant ΔNΔC37CP using NPdock (http://genesilico.pl/NPDock) (Tuszynska et al., 2015) and HDOCK (http://hdock.phys.hust.edu.cn/)(Yan et al., 2017) web servers for protein-RNA interaction prediction study. The best RNA-protein model was predicted based on 1,000 simulation steps with initial temperature of 15,000 K, and 295 K for the last step of simulation, RMSD cutoff for clustering, docking score and structure refinement.

3. Results and discussion

3.1. Purification of the recombinant coat protein

Phaseolus vulgaris plants are the susceptible host for bean common mosaic virus (BCMV) infection which causes severe yield loss in the crop. BCMV is a seed-borne virus (Salem et al., 2010). The BCMV infection was observed at the 3-leaf stage seedlings of *P. vulgaris* seeds, in which 14.0% disease incidence was found. The infected leaves showed leaf rolling, green-yellow mosaic, chlorosis and mottling upon BCMV infection. The amplified PCR product of coat protein was observed at ~717 bp in agarose gel (Fig. S1A). Later, the CP gene was cloned into pET28a expression vector. Insertion of CP gene was confirmed by sequencing.

The expression vector was transformed into expression *E. coli* strain and the protein expression was carried out. Induced recombinant coat protein was purified through Ni-NTA column. Further, the eluted fractions showed protein bands corresponding to 30 kDa on SDS-PAGE which was equal to the theoretical molecular mass of His-tag containing coat protein (Fig. S1B). The 30 kDa protein was subjected to matrix-assisted laser desorption/ionization - mass spectrometry (MALDI-MS), and a mass spectra profile of peptides that made up the protein was built using MASCOT. These peptides encompassed the amino acid stretches of the coat protein (Fig. S1C). After the removal of N-terminal His-tag, purification of CP was achieved through Superdex-200 Gel-filtration column. The GFC profile of CP revealed that the retention volume of CP was between 60 ml–90 ml, corresponding to the similar retention volume of 200 kDa to 29 kDa proteins (Fig. 1B). The Native-PAGE analysis showed that CP existed in different oligomeric forms, the extent of which was concentration dependent (Fig. 1B). The number of oligomeric forms was more at 0.5 mg ml^{-1}, when compared with 0.2 mg ml^{-1} sample. This result indicates that CP tends to form high molecular weight oligomers even in the presence of 1 mM DTT. The CP monomer contains one cysteine, and it might be possible that the CP underwent oligomerization due to the formation of inter-molecular disulfide bond formation. To assess this, the cysteine amino acid was mutated to alanine in the wild-type coat protein. In the absence of cysteine, the elution profile of coat protein remained unchanged. Previous *in vitro* studies have shown that the CP monomers assembled to form ring-like structure in the presence or absence of RNA (Jagadish et al., 1991). The coat protein of pepper vein banding virus (PVBV) assembled to form flexuous rod-shaped structure where the oligomerization could be disturbed and reformed under the control of ionic strength and pH conditions (Anindya and Savithri, 2003).

3.2. Equilibrium unfolding studies of coat protein

3.2.1. Tertiary structure assessment through intrinsic and extrinsic fluorescence

To understand the structural properties of CP and its behavior in stress conditions, spectroscopic assays were performed. Intrinsic tryptophan (Trp) fluorescence provides important information regarding the tertiary structure integrity of a protein. Tryptophan residue is

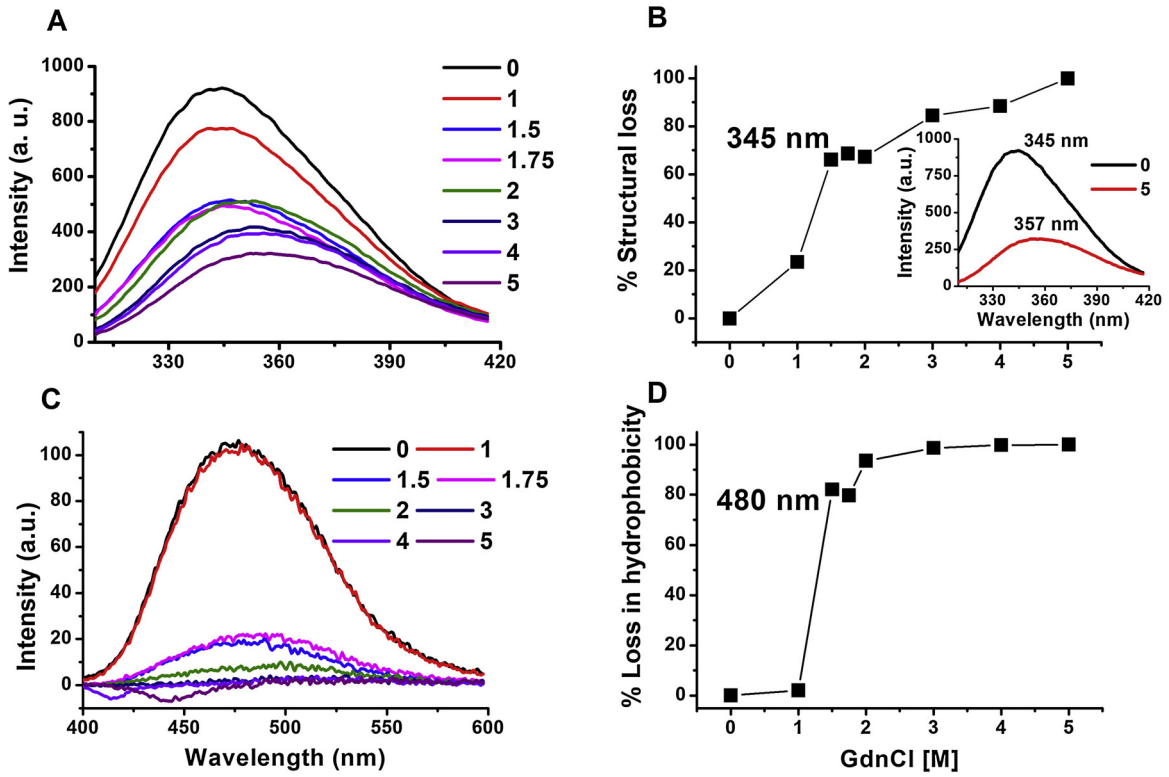

Fig. 2. Tertiary structure analysis of CP: (A) Intrinsic Trp fluorescence spectra of CP incubated at varying concentration of GdnCl, (B) The corresponsing loss in the signal intensity at 345 nm is plotted. In the inset, fluorescence emission maximum is shown for native CP and CP denatured in 5 M GdnCl. (C) & (D) ANS fluorescence spectra and corresponding loss in surface hydrophobicity, respectively for CP denatured at varying concentration of GdnCl.

hydrophobic in nature, which tends to bury inside the protein core. However, structure of some proteins is not globular by account of their intrinsic disorder or due to structural perturbation. The Trp of those proteins come in contact with water and display characteristic spectrum. Intrinsic tryptophan fluorescence of CP was monitored at 295 nm excitation. The tryptophan fluorescence emission maximum (λ_{max}) of native protein was corresponding to 345 nm, which pointed towards a slight structural disorder in CP (Fig. 2A & B). In general, a properly folded protein (globular proteins) exhibits λ_{max} corresponding to 330 nm in native condition. To understand the unfolding mechanism of coat protein, it was subjected to denaturation using different concentration of GdnCl. With increasing GdnCl concentration, the fluorescence intensity of CP decreased with a concomitant red shift in its λ_{max} (Fig. 2A). It was observed that there was no further change in the fluorescence intensity with λ_{max} corresponding to 357 nm at 5 M GdnCl, which is a sign of a completely open protein structure, wherein all the Trp residues have been exposed to the surrounding solvent. Further, we monitored the percent loss in fluorescence intensity at 345 nm and observed that the unfolding profile was non-sigmoidal in nature which lacked a well defined pre- and post-transition baseline. In other words, the transition was gradual and non-cooperative. This kind of gradual transition has been observed in molten globular stage of many proteins, which is characterized by a loose tertiary structure.

To further understand the tertiary structural perturbation, the corresponding ANS fluorescence was monitored at varying denaturant concentration. In general, the emission intensity of free ANS is very less, but enhances significantly after binding to the exposed hydrophobic pockets of the protein. The ANS fluorescence intensity of CP displayed a λ_{max} of 480 nm in the native state, indicating that its hydrophobic pockets were significantly exposed (Fig. 2C). This is quite different from the globular proteins which show negligible ANS fluorescence in their native state. The result was in accordance with the intrinsic Trp fluorescence data. The ANS intensity decreased dramatically at 1.5 M GdnCl, which indicates a sudden loss in the structure.

3.2.2. *Secondary structure assessment through far UV CD*

Far UV circular dichroism (CD) gives good approximation of overall secondary structural content of a protein. The GdnCl mediated denaturation indicated that the loss in secondary structure was gradual beyond 1 M GdnCl (Fig. 3). Hence, it can be concluded from the intrinsic Trp fluorescence and far UV CD spectra results that the CP displayed spectral properties of a polypeptide with major unstructured region. Many nucleic acid binding proteins are known to have disordered regions and such regions tend to gain structure when bound to their nucleic acid partners. Thus, we can conclude that CP subunit has loosely packed tertiary structure but gained a compact and stable structure when it makes contact with another subunit or viral RNA. These outcomes are supported by the previous reports that the structure of potato virus A (PVA) CP includes a significant disordered segment

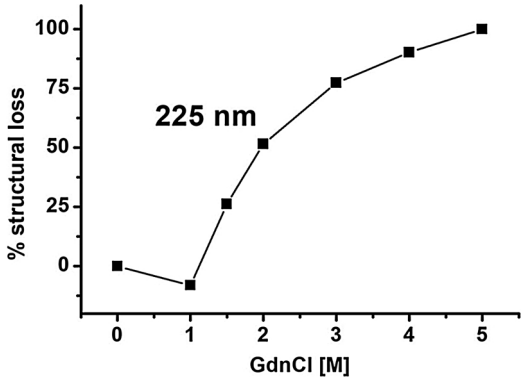

Fig. 3. Secondary structure analysis of CP at varying GdnCl concentration: A corresponding percent loss in signal intensity at 222 nm is shown.

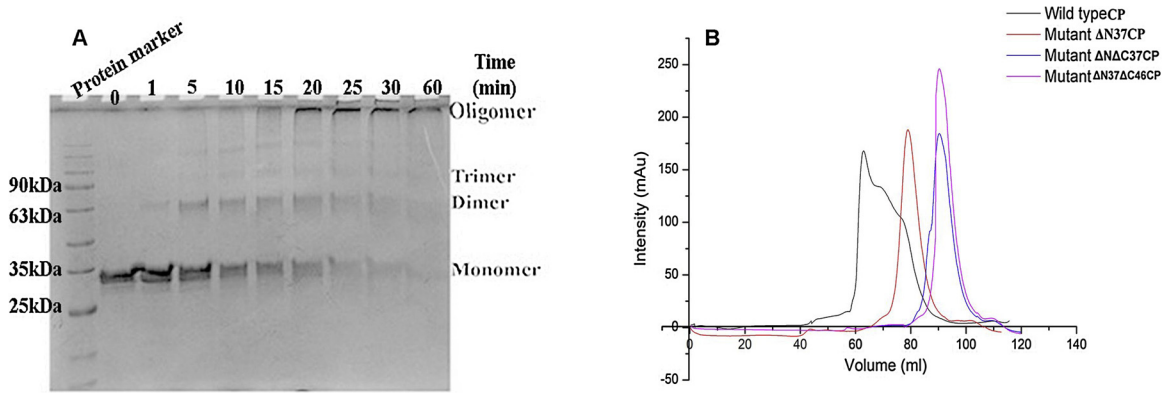

Fig. 4. Analysis of oligomerization property of coat protein: (A) Glutaraldehyde cross-linking of coat protein was analyzed by 12% SDS–PAGE. (B) The SEC elution profile of wild type CP, ΔN37CP, ΔNΔC37CP, and ΔN37ΔC46CP were shown.

and sustained stability by interactions between them (Ksenofontov et al., 2013).

The importance of this study was to investigate the viral CP structural behavior in the absence of viral RNA. The viral CP despite of having an unstructured region plays a major role in the virus packaging through oligomerization. The next step of the study was to investigate how the unstructured CP oligomerize.

3.3. Oligomerization of coat protein

The cross-linking experiment was performed using glutaraldehyde to investigate the oligomerization property of coat protein. Time-course cross-linking results suggested that CP exist predominantly in a dimer form. After 1 h of cross-linking reaction, all units of CP were shifted to high oligomers, possibly due to non-specific crosslinking (Fig. 4A). This result indicated that the dimerization could be a basic pre-requisite for CP assembly or the dimers might be the building blocks of CP assembly. Dimerization/oligomerization of protein has to happen through interaction between domains or terminal residues. Earlier reports observed that the oligomerization of CP subunits was carried out by amino- and carboxy-terminal residues in potexviruses (DiMaio et al., 2015). It was shown that the N- and C- terminal exposed regions are essential for initiation of assembly using the N- and C- terminal deletion mutants in the case of pepper vein banding virus (Anindya and Savithri, 2003). Hence, oligomeric property of CP depend upon the composition of residues present at the terminal ends. In the present study, the unstructured domain (37 amino-acids of N-terminal) was deleted from the wild-type coat protein, referred to as ΔN37CP to investigate the role of the N-terminal residues. The purified mutant ΔN37CP showed a single peak of ~23 kDa protein profile on GFC, which corresponds to the monomer of mutant CP (Fig. 4B). These results suggested that the N-terminal residues are crucial in the polymerization of coat protein in BCMV virion formation. Thus, the role of N-terminal residues varies from virus to virus. For example, the N-terminal residues were responsible for the interaction between CP subunits in pepino mosaic virus (PepMV) and watermelon mosaic virus (WMV) (Agirrezabala et al., 2015; DiMaio et al., 2015; Zamora et al., 2017), while N-terminus interacted with 3′ untranslated region of viral RNAs to facilitate alfalfa mosaic virus (AMV) infection (Yusibov et al., 1996). Earlier reports states that the CP oligomerized in the presence of RNA/DNA to form different types of structure like spherical/ovoid or tubular, which depends upon the length of the nucleic acid (Hull, 2013).

3.4. Interaction with the untranslated region of the viral genome

It is well known that the coat protein of RNA viruses displays high affinity towards the viral RNA to maintain the RNA stability and thus facilitate the cell-to-cell spread of the virus particle (Hyodo et al., 2014). The untranslated regions (UTRs) are present at both ends of the genome with polyA tail at the 3′ end, and a covalently linked virus-encoded protein (VPg) at 5′ end (Oruetxebarria et al., 2001; Puustinen et al., 2002). Mutational analysis of coat protein was performed to find out the nucleic acid binding site. The C-terminal mutant was generated by deleting 37 residues from the mutant ΔN37CP, which was referred to as ΔNΔC37CP. The N-terminal ΔN37CP and C-terminal ΔNΔC37CP mutants were analyzed for 3′ UTR-RNA binding assay, in which both the mutants showed binding with 3′ UTR-RNA. Further, ΔN37ΔC46CP mutant was constructed by removing additional nine amino acids from the C-terminal of ΔNΔC37CP mutant. The 3′ UTR-RNA binding assay was performed with both ΔNΔC37CP and ΔN37ΔC46CP mutant proteins. The mutant ΔNΔC37CP showed binding with 3′ UTR-RNA as observed through retarded migration of 3′ UTR-RNA on EMSA gel, whereas, there was no shift observed with mutant ΔN37ΔC46CP when compared with the control 254 nt long 3′ UTR-RNA (Fig. 5B). These results indicated that the wild type CP, ΔN37CP and ΔNΔC37CP mutants showed 3′ UTR-RNA binding activity due to the presence of nine amino acids, which were not present in ΔN37ΔC46CP mutant.

The mutant ΔNΔC37CP contains AALSNVSSK extra residues at the C-terminal as compared to ΔN37ΔC46CP, which could be crucial for RNA binding. This RNA binding region is conserved among potyviruses as observed through multiple sequence alignment (Fig. S2). This extra RNA binding amino acid stretch contain hydrophobic and polar amino acids, which might play a crucial role in the interaction with nucleic acid. Hydrogen and hydrophobic interaction are the key factors responsible for protein-DNA/RNA complex formation (Jones et al., 2001). This RNA binding amino acid stretch contains 27% serine, which could play important role in CP-RNA interaction. Previously, it was observed that the interaction of coat protein with viral RNA is regulated by the phosphorylation of serine, which was under the control of protein kinase CK2 (Ivanov et al., 2003). The RNA binding site was observed in watermelon mosaic virus at the amino acid position S140-R172, which was a long stretch of residues (Zamora et al., 2017). The cryo-EM structure of pepino mosaic virus (PepMV) revealed a RNA binding pocket near the residue position D163 in the coat protein (Agirrezabala et al., 2015). There are no reports on the crystal structure of coat protein which can explain the correct position of amino acids or domains which interacts with the nucleic acid.

The 144 nt long 5′ UTR-RNA was also tested for its binding with mutant ΔNΔC37CP. The mutant ΔNΔC37CP showed less binding with 5′ UTR-RNA, whereas the mutant ΔN37ΔC46CP did not bind to the 5′ UTR RNA as observed through EMSA (Fig. 5C). Hence, the mutant ΔNΔC37CP showed a higher RNA binding with the 3′ UTR as compared to the 5′ UTR. Both the UTR RNA secondary structures were analyzed online using RNAfold web server which showed more hairpin like secondary structures or pseudo-knots in the 3′ UTR than in the 5′ UTR (Fig. 5A&D). A previous study describes that the 5′ UTR is much less

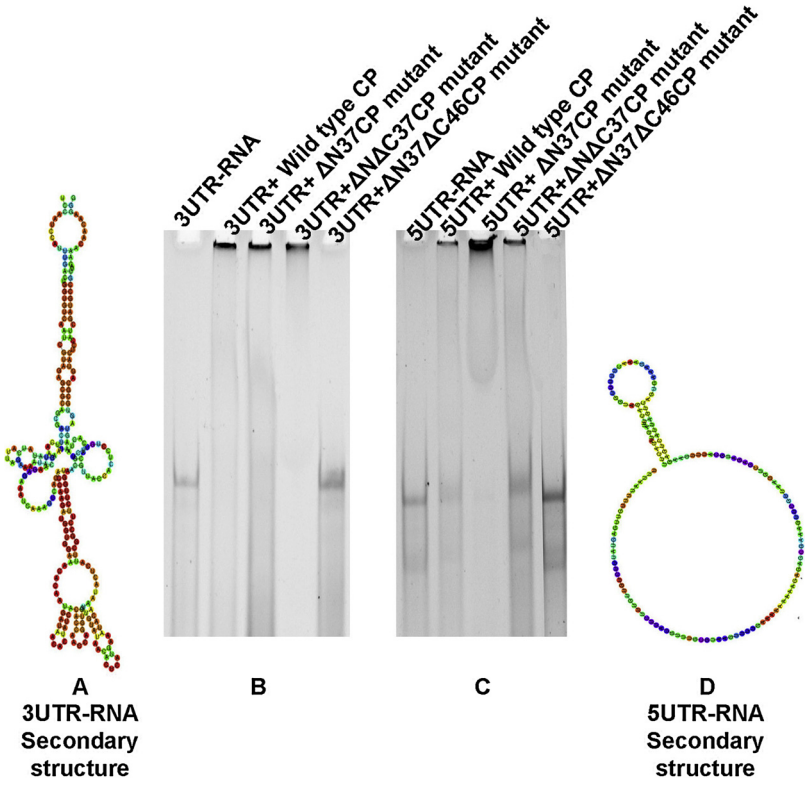

Fig. 5. The RNA binding activity of coat protein mutants. (A) The 3′ UTR RNA secondary structure; (B) EMSA was performed for analysis of 3′ UTR-RNA binding with CP mutants. Lane 1 contain 3′ UTR-RNA as control, lane 2 contain wild type CP and 3′ UTR-RNA complex, lane 3 contain ΔN37CP mutant protein and 3′ UTR-RNA, lane 4 contain complex of 3′ UTR-RNA and ΔNΔC37CP mutant, lane 5 contain complex of 3′ UTR-RNA and ΔN37ΔC46CP. (C) The mutant proteins ΔNΔC37CP and ΔN37ΔC46CP interaction with 5′UTR RNA was analyzed on EMSA gel. Lane 1 contain 5′ UTR-RNA as control, lane 2 contain complex wild type CP and 5′ UTR-RNA, lane 3 contain ΔN37CP mutant protein and 5′ UTR-RNA, lane 4 contain complex of 5′ UTR-RNA and ΔNΔC37CP mutant, lane 5 contain complex of 5′ UTR-RNA and ΔN37ΔC46CP. (D) The secondary structure of circularized 5′ UTR.

structured in *Potyviridae* (Kneller et al., 2006). In positive-strand RNA viruses, the 5′ UTR of viral RNAs interact covalently with the 3′ UTR to circularize the RNA for efficient translation (Iwakawa et al., 2012). The 3′ UTR had a tendency to adopt conformational changes for binding to coat protein to inhibit negative strand promoter activity and enhanced viral RNA translation (Neeleman et al., 2004). The CP of alfalfa mosaic virus (AMV) binds to the 3′ UTR which leads to 50–100 fold enhanced translation of viral RNA in tobacco protoplasts (Neeleman et al., 2001). The binding of the CP with the 3′ UTR, resulted in the conformational change in its pseudo-knot to play as a controller to regulate the translation and replication of the viral genome (Miras et al., 2017). In brome mosaic virus (BMV), the 3′ UTR act as a cis- and trans- element for viral genome encapsidation through interaction with the coat protein (Yi et al., 2009).

3.5. In-silico study of a nucleoprotein complex

The three dimensional (3D) homology model of mutant ΔNΔC37CP was generated using two different software; RaptorX and I-TASSER which provided 3 top predicted models based on their scoring algorithms (Fig. 6A). The 3D model of CP mutant was generated using PDB 5odvA as the template which showed 83% identity with the template. The 3D model by RaptorX contained 35% helix and 49% coiled structure. Table S3 illustrated the confidence score (C-value), TM-score, root mean square deviation (RMSD) and Cluster density of best-predicted 3D model of CP by I-TASSER. The superimposition of predict 3D model of CP mutant with template 5odvA showed 1.8 ± 1.5 Å RMSD. Additionally, the modeled mutant ΔNΔC37CP was analyzed by PROCHECK for Ramachandran plot analysis which revealed that 85.4% residues were present in the favored region, 13.2% in the allowed region and 1.3% in the disallowed regions (Fig. S3). The earlier report predicted the 3D model of potato virus A (PVA) coat protein which contained the high content of random coil structure (40–66%) and 25–45% α-helix structure (Ksenofontov et al., 2013).

The 3D structure model of 3′ UTR RNA was generated by simulation using SimRNA server. The simulated 3D structure of RNA was generated with high accuracy pseudoknots and folds (Fig. 6B). The predicted model of 3′ UTR was docked for interaction with the 3D model of mutant ΔNΔC37CP using NPdock and HDOCK server. The Lowest energy model nucleoprotein complex showed docking score −191.90 with RMSD 44.87 °A in HDOCK server. The NPdock server also simulated the best nucleoprotein complex on the basis of lower energy structure. The binding of C-terminal of mutant ΔNΔC37CP with 3′ UTR was observed in predict 3D model of RNA-protein complex through both servers (Fig. 6C). The 3′ UTR played a switch controller between viral genome undergoing the replication and translation process in many plant viruses to facilitating infection (Simon and Miller, 2013). The CP binds to 3′ UTR for stabilizing the translation machinery (Bol, 2005; Neeleman et al., 2004).

In conclusion, this study explains the assembly and RNA binding residues of BCMV coat protein through mutations. This study describes that only first 37 N-terminal residues of CP are responsible for BCMV virion assembly. The assembly of some potyviruses occurs due to both terminal residues while some virion assembly happened without N- and C-terminal residues. Thus, assembly of CP differs within the potyviruses due to the amino acids variation in the terminal regions. The coat protein-RNA interaction study also suggested that RNA binding domain corresponding to 194–202 amino acid stretch is conserved among the potyviruses despite variation in assembly process. From therapeutics point of view, mutation in the RNA binding domain could be suppressing the cell to cell movement of viral genome. The conserved RNA binding site in CPs might be used as a target for antiviral compounds that block genome packaging of potyviruses.

Author contributions

SK, RK, DKG and IG performed the experiments and wrote the manuscript. AKP designed the experiment and corrected the manuscript.

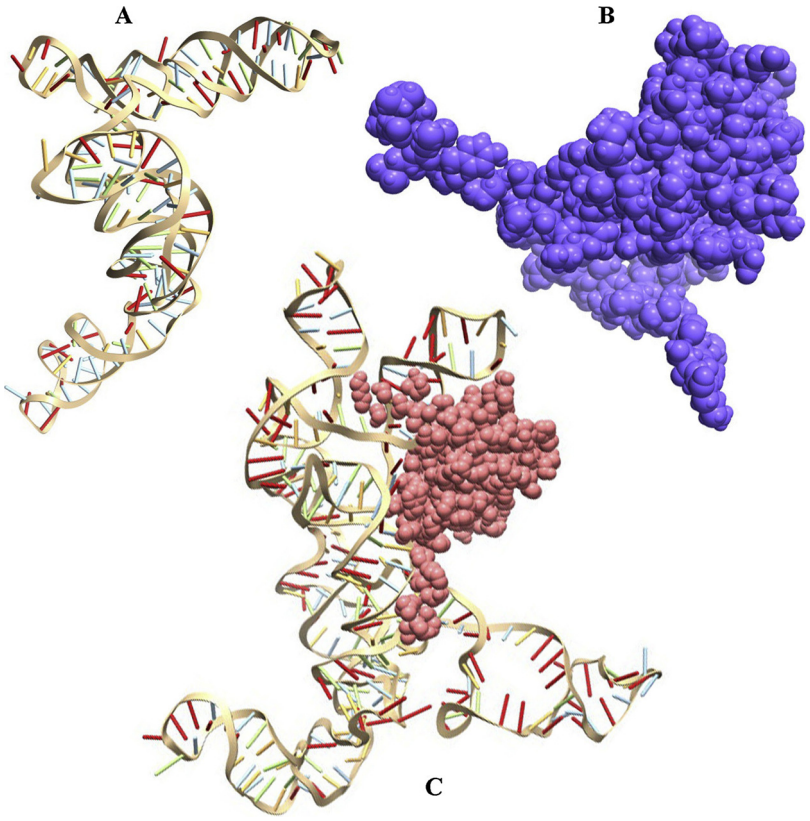

Fig. 6. (A) Three-dimensional (3D) structural model of 3′ UTR RNA. (B) The 3D structure of simulated mutant protein ΔNΔC37CP. (C) The predicted 3D structure of 3′ UTR RNA- mutant protein ΔNΔC37CP complex.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgement

Authors acknowledge the infrastructural support of IIT Delhi. SK acknowledges the young-scientist research grant (ID: YSS/2014/000819) from SERB, Department of Science & Technology, Govt. of India. Authors thank to Dr. Hariprasad P. for providing screen-house facility.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.virusres.2019.197755.

References

Agirrezabala, X., Méndez-López, E., Lasso, G., Sánchez-Pina, M.A., Aranda, M., Valle, M., 2015. The near-atomic cryoEM structure of a flexible filamentous plant virus shows homology of its coat protein with nucleoproteins of animal viruses. Elife 4, e11795.
Anindya, R., Savithri, H., 2003. Surface-exposed amino-and carboxy-terminal residues are crucial for the initiation of assembly in Pepper vein banding virus: a flexuous rod-shaped virus. Virology 316 (2), 325–336.
Besong-Ndika, J., Ivanov, K.I., Hafrèn, A., Michon, T., Mäkinen, K., 2015. Co-translational CP-mediated inhibition of potyviral RNA translation. J. Virol., JVI 02915–02914.
Bol, J.F., 2005. Replication of alfamo-and ilarviruses: role of the coat protein. Annu. Rev. Phytopathol. 43, 39–62.
Boniecki, M.J., Lach, G., Dawson, W.K., Tomala, K., Lukasz, P., Soltysinski, T., Rother, K.M., Bujnicki, J.M., 2015. SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. Nucleic Acids Res. 44 (7) e63–e63.
DiMaio, F., Chen, C.-C., Yu, X., Frenz, B., Hsu, Y.-H., Lin, N.-S., Egelman, E.H., 2015. The molecular basis for flexibility in the flexible filamentous plant viruses. Nat. Struct. Mol. Biol. 22 (8), 642.
Dolja, V., Haldeman, R., Robertson, N., Dougherty, W., Carrington, J., 1994. Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants. EMBO J. 13 (6), 1482–1491.
Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R., Hofacker, I.L., 2008. The vienna RNA websuite. Nucleic Acids Res. 36 (suppl_2), W70–W74.
Hull, R., 2013. Plant Virology. . Academic press.
Hyodo, K., Kaido, M., Okuno, T., 2014. Host and viral RNA-binding proteins involved in membrane targeting, replication and intercellular movement of plant RNA virus genomes. Front. Plant Sci. 5, 321.
Ivanov, K.I., Mäkinen, K., 2012. Coat proteins, host factors and plant viral replication. Curr. Opin. Virol. 2 (6), 712–718.
Ivanov, K.I., Puustinen, P., Gabrenaite, R., Vihinen, H., Rönnstrand, L., Valmu, L., Kalkkinen, N., Mäkinen, K., 2003. Phosphorylation of the potyvirus capsid protein by protein kinase CK2 and its relevance for virus infection. Plant Cell 15 (9), 2124–2139.
Iwakawa, H.-o., Tajima, Y., Taniguchi, T., Kaido, M., Mise, K., Tomari, Y., Taniguchi, H., Okuno, T., 2012. Poly (A)-binding protein facilitates translation of an uncapped/nonpolyadenylated viral RNA by binding to the 3′untranslated region. J. Virol., JVI 00538–00512.
Jagadish, M.N., Ward, C.W., Gough, K.H., Tulloch, P.A., Whittaker, L.A., Shukla, D.D., 1991. Expression of potyvirus coat protein in Escherichia coli and yeast and its assembly into virus-like particles. J. Gen. Virol. 72 (7), 1543–1550.
Jones, S., Daley, D.T., Luscombe, N.M., Berman, H.M., Thornton, J.M., 2001. Protein–RNA interactions: a structural analysis. Nucleic Acids Res. 29 (4), 943–954.
Kendall, A., McDonald, M., Bian, W., Bowles, T., Baumgarten, S.C., Shi, J., Stewart, P.L., Bullitt, E., Gore, D., Irving, T.C., 2008. Structure of flexible filamentous plant viruses. J. Virol. 82 (19), 9546–9554.
Kneller, E.L.P., Rakotondrafara, A.M., Miller, W.A., 2006. Cap-independent translation of plant viral RNAs. Virus Res. 119 (1), 63–75.
Ksenofontov, A.L., Paalme, V., Arutyunyan, A.M., Semenyuk, P.I., Fedorova, N.V., Rumvolt, R., Baratova, L.A., Järvekülg, L., Dobrov, E.N., 2013. Partially disordered structure in intravirus coat protein of potyvirus potato virus A. PLoS One 8 (7), e67830.
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 (2), 283–291.
Miras, M., Miller, W.A., Truniger, V., Aranda, M.A., 2017. Non-canonical translation in plant RNA viruses. Front. Plant Sci. 8, 494.
Neeleman, L., Linthorst, H.J., Bol, J.F., 2004. Efficient translation of alfamovirus RNAs requires the binding of coat protein dimers to the 3′ termini of the viral RNAs. J. Gen. Virol. 85 (1), 231–240.
Neeleman, L., Olsthoorn, R.C., Linthorst, H.J., Bol, J.F., 2001. Translation of a non-polyadenylated viral RNA is enhanced by binding of viral coat protein or poly-adenylation of the RNA. Proc. Natl. Acad. Sci. 98 (25), 14286–14291.
Oruetxebarria, I., Guo, D., Merits, A., Mäkinen, K., Saarma, M., Valkonen, J.P., 2001. Identification of the genome-linked protein in virions of Potato virus A, with

comparison to other members in genus Potyvirus. Virus Res. 73 (2), 103–112.
Puustinen, P., Rajamäki, M.-L., Ivanov, K.I., Valkonen, J.P., Mäkinen, K., 2002. Detection of the potyviral genome-linked protein VPg in virions and its phosphorylation by host kinases. J. Virol. 76 (24), 12703–12711.
Revers, F., García, J.A., 2015. Molecular Biology of Potyviruses, Advances in Virus Research Vol. 92. Elsevier, pp. 101–199.
Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5 (4), 725.
Salem, N., Ehlers, J., Roberts, P., Ng, J., 2010. Biological and molecular diagnosis of seedborne viruses in cowpea germplasm of geographically diverse sub-Saharan origins. Plant Pathol. 59 (4), 773–784.
Simon, A.E., Miller, W.A., et al., 2013. 3′ cap-independent translation enhancers of plant viruses. Annu. Rev. Microbiol. 67, 21–42.
Studier, F.W., 2014. Stable Expression Clones and Auto-induction for Protein Production in E. coli, Structural Genomics. Springer, pp. 17–32.
Tuszynska, I., Magnus, M., Jonak, K., Dawson, W., Bujnicki, J.M., 2015. NPDock: a web server for protein–nucleic acid docking. Nucleic Acids Res. 43 (W1), W425–W430.
Urcuqui-Inchima, S., 2001. Potyvirus proteins: a wealth of functions. Virus Res. 74, 157–175.
Wang, S., Li, W., Liu, S., Xu, J., 2016. RaptorX-Property: a web server for protein structure property prediction. Nucleic Acids Res. 44 (W1), W430–W435.
Worrall, E.A., Wamonje, F.O., Mukeshimana, G., Harvey, J.J., Carr, J.P., Mitter, N., 2015. Bean Common Mosaic Virus and Bean Common Mosaic Necrosis Virus: Relationships, Biology, and Prospects for Control, Advances in Virus Research Vol. 93. Elsevier, pp. 1–46.
Yan, Y., Zhang, D., Zhou, P., Li, B., Huang, S.-Y., 2017. HDOCK: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. Nucleic Acids Res. 45 (W1), W365–W373.
Yi, G., Letteney, E., Kim, C.-H., Kao, C.C., 2009. Brome mosaic virus capsid protein regulates accumulation of viral replication proteins by binding to the replicase assembly RNA element. Rna.
Yusibov, V., Kumar, A., North, A., Johnson, J.E., Loesch-Fries, L.S., 1996. Purification, characterization, assembly and crystallization of assembled alfalfa mosaic virus coat protein expressed in Escherichia coli. J. Gen. Virol. 77 (4), 567–573.
Zamora, M., Méndez-López, E., Agirrezabala, X., Cuesta, R., Lavín, J.L., Sánchez-Pina, M.A., Aranda, M.A., Valle, M., 2017. Potyvirus virion structure shows conserved protein fold and RNA binding site in ssRNA viruses. Sci. Adv. 3 (9) eaao2182.