The cryo-electron microscopy structure of Broad Bean Stain Virus suggests a common capsid assembly mechanism among comoviruses

François Lecorre, Josephine Lai-Kee-Him, Stéphane Blanc, Jean-Louis Zeddam, Stephano Trapani, Patrick Bron

The cryo-electron microscopy structure of Broad Bean Stain Virus suggests a common capsid assembly mechanism among comoviruses

François Lecorre[1], Joséphine Lai-Kee-Him[1], Stéphane Blanc[2], Jean-Louis Zeddam[3*], Stefano Trapani[1*], and Patrick Bron[1,*]

[1] Centre de Biochimie Structurale (CBS), INSERM, CNRS, Univ. Montpellier, 29 rue de Navacelles, 34090 Montpellier, France.

[2] INRA, Virus Insect Plant Laboratory, Joint Research Unit UMR 385 BGPI, Campus International de Baillarguet, Montpellier, France.

[3] IRD, Cirad, Montpellier University, Joint Research Unit UMR 186 IPME, Montpellier, France.

[*]Address correspondence to Patrick Bron: Fax; +33 491 266 720; e-mail: patrick.bron@cbs.cnrs.fr; or to Stefano Trapani: e-mail: Stefano.trapani@cbs.cnrs.fr; or to Jean-Louis Zeddam: jean-louis.zeddam@cirad.fr.

ABSTRACT

The Broad bean stain virus (BBSV) is a member of the genus Comovirus infecting Fabaceae. The virus is transmitted through seed and by plant weevils causing severe and widespread disease worldwide. BBSV has a bipartite, positive-sense, single-stranded RNA genome encapsidated in icosahedral particles. We present here the cryo-electron microscopy reconstruction of the BBSV and an atomic model of the capsid proteins refined at 3.22 Å resolution. We identified residues involved in RNA/capsid interactions revealing a unique RNA genome organization. Inspection of the small coat protein C-terminal domain highlights a maturation cleavage between Leu567 and Leu568 and interactions of the C-terminal stretch with neighbouring small coat proteins within the capsid pentameric turrets. These interactions previously proposed to play a key role in the assembly of the Cowpea mosaic virus suggest a common capsid assembly mechanism throughout all comovirus species.

Keywords: Cryo-EM, Three-dimensional structure, Single-particle analysis, BBSV, Comovirus, Virus coat protein, Single-stranded RNA genome

INTRODUCTION

Broad bean stain virus (BBSV) belongs to subfamily *Comoviridae* along with other plant-infecting members of the family *Secoviridae* within the order *Picornavirales*. This non-enveloped virus, which has a bipartite, single-stranded, positive-sense RNA genome, induces severe and widespread disease in Europe, North Africa, Asia, the Middle East and Australia (Al-Khalaf et al., 2002; Gibbs et al., 1968; Kumari and Makkouk, 1996; Makkouk and Azzam, 1986; Makkouk et al., 2012). The natural host range of BBSV is restricted to Fabaceae such as lentil (*Lens culinaris*), pea (*Pisum sativum*), faba bean (*Vicia faba*) and vetch (*Vicia palaestina*) (Makkouk and Azzam, 1986; Makkouk et al., 2012). BBSV is transmitted from plant to plant through two distinct strategies. It can be spread by weevil vectors (*Sitona* spp., *Apion* spp.) that ingest virus particles when feeding on infected plants and that retain the ability to transmit for days or weeks. Alternatively, it can be spread through host plant seeds, leading to an early infection of crops from the very first stage of plant development. Several strains of BBSV have been described that cause symptoms of varying severity ranging from mild to extremely severe, culminating in plant death (Kumari and Makkouk, 1996). Symptoms can vary cyclically during plant growth (Gibbs et al., 1968), and, as with other seed-borne viruses, they can remain undetectable for extended periods. The standard methods used to detect BBSV are serological tests based on ELISA and tissue blot immunoassays (Kumari and Makkouk, 1993). Seed transmission rates of BBSV are reported to be very high: up to 20% in faba bean (Mali et al., 2003), 50% in field pea (Fiedorow and Szlachetka-Wawrzyniak, 2002) and 27% in lentil (Al-Khalaf et al., 2002). Yield losses can also be high. For instance, pre-flowering infection in some lentils can lead to 77% yield loss (Mabrouk and Mansour, 1998). Because the maximum reported rate of seed

transmission is 50%, this sole mode of vertical transmission would reduce the virus incidence over host plant generations, underlining the importance of horizontal vector transmission.

The BBSV genome is composed of two positive-sense single-stranded RNA segments, termed RNA-1 and RNA-2, which are separately encapsidated, although both are required for virus infection (Figure 1A). Exploration of public databases indicates that only RNA-2 and the RNA-dependent RNA polymerase encoded by RNA-1 have been sequenced. RNA-2 encodes two capsid proteins (CPs), the large (L) and the small (S) coat protein, which are processed from a polyprotein precursor by the action of the RNA-1-encoded proteinase. Consequently, the presence of both genomic segments in an infected cell appears mandatory to ensure correct capsid assembly.

Our knowledge of the structural organization of comoviruses is well documented since atomic models based on crystallographic or cryo-electron microscopy (cryo-EM) have been reported for three comovirus: cowpea mosaic virus (CPMV; PDB entries 1NY7 (Lin et al., 1999)), 2BFU (Ochoa et al., 2006), 5A32, 5A33 (Hesketh et al., 2015) and 5FMO (Huynh et al., 2016)); bean pod mottle virus (BPMV; PDB entries 1PGL, 1PGW (Lin and Johnson, 2003) and 1BMV (Chen et al., 1989)); and red clover mottle virus (RCMV (VIPERdb entry 202, http://viperdb.scripps.edu) (Lin et al., 2000)). The comovirus capsid structures reveal a common structural organization (Lin et al., 2000; Lin and Johnson, 2003) in which the virions are composed of 60 copies of both the L and S coat proteins arranged into pseudo T = 3 icosahedral capsids with a diameter of ~30 nm. Each L subunit is composed of two jellyroll β-barrel domains, while each S subunit consists of a single jellyroll β-barrel. The S subunits are arranged around the icosahedral 5-fold symmetry axes and form protrusions (turrets) on the otherwise quasi-spherical

capsid surface. Each turret exhibits a central channel, which potentially connects the capsid interior to the exterior along the 5-fold axis, and its entry on the external side is constricted by a pentameric annulus of residues. The presence of a second more internal pentameric annulus along the turret channel, identified in CPMV and RCMV but not in BPMV, seems not to be a conserved feature among comoviruses (based on CP RNA-2 sequence comparison). This observation led to the classification of comovirus capsid structures into two subgroups, with RCMV and CPMV belonging to one subgroup, and BPMV and Cowpea severe mosaic virus (CPSMV) belonging to the other (Chen and Bruening, 1992; Lin et al., 2000).

However, relatively little is known about genome encapsidation and virus assembly. Some ordered ribonucleotides were observed within RNA-2-filled BPMV by X-ray diffraction (Chen et al., 1989), while single-particle cryo-EM studies on differently filled (either RNA-1 or RNA-2) and empty CPMV particles separated by gradient ultracentrifugation revealed several layers of RNA density and highlighted a few capsid residues interacting with genomic RNA (Hesketh et al., 2015; Hesketh et al., 2017). It was shown that the S subunit C-terminal region plays a critical role in virus assembly and genome encapsidation, although it is hydrolysed during maturation of virus particles (Lomonossoff and Johnson, 1991; Sainsbury et al., 2011; Taylor et al., 1999). Recent studies on CPMV (Hesketh et al., 2015; Huynh et al., 2016) revealed the presence of multiple cleavage sites and shed light on the role of some interactions of the S subunit C-terminus in favouring penton formation.

In the present work, we investigate for the first time the structural architecture of BBSV by cryo-EM single-particle analysis, and compute an atomic model of the icosahedral protein capsid refined at 3.22 Å resolution. The results provide a better understanding

of BBSV particle organisation and its internal RNA segment. Additionally, the extensive structural comparison of BBSV viral particle with other comoviruses strengthens our knowledge on structural organisation, genome encapsidation and assembly mechanisms within the comovirus family.

RESULTS

Cryo-EM images of frozen-hydrated BBSV virions purified from infected leaves revealed low concentrations of viral particles (Figure S1) with a diameter of ~30 nm that were mainly composed of RNA-filled capsids and, occasionally, empty particles. We then recorded a movie dataset of frozen-hydrated BBSV virions using a Titan Krios S-FEG electron microscope as described in Materials and Methods. A 3D density map (Figure 2A and B) was subsequently obtained from a homogeneous set of 17,233 RNA-filled particles. The resolution of the final map (refined by imposing icosahedral symmetry constraints) was estimated to be 3.22 Å by both the so-called gold-standard Fourier shell correlation (FSC) method using the 0.143 criterion (Scheres and Chen, 2012) and the high-resolution noise substitution method (Chen et al., 2013) (see Materials and Methods and Figure 3A). Local resolution estimations (Figure 4) indicate that many regions of the map reached even higher resolution than 3.2 Å in the inner part of the capsid shell, while the more disordered parts are near the 5-fold symmetry axis, consisting in regions located at the upper face of S subunits forming the turret (black arrows) and three regions located at the base of the turret (white arrows).

The resolution of the wild-type BBSV cryo-EM density map was high enough to clearly resolve many amino acid side chains (Figure 2B, C and D), and an atomic model of the BBSV capsid could be built and refined to approximately 3.22 Å resolution with no significant overfitting (Figure 3B). Table 1 summarises image processing and atomic model validation parameters. All CP residues (Figure 1B) could be built within the cryo-EM map except residues 371–377 (which constitute the C-terminal stretch of the L subunit) and residues 560–592 at the S subunit C-terminal end. A tentative model of residues 560–567 was built, although not included in the final refined structure (see

below). In Figure 1B, resolved residues of the L and S coat proteins are coloured in blue and red, respectively, while unresolved residues are black. It should be noted that the CP amino acid sequence of our BBSV strain, as determined by cDNA analysis, differs from the BBSV CP sequence available from UniProt (D0PSV7) at seven amino acid positions (Figure 1B and Figures S2 and S3). A ribbon representation of one CP protomer (one L subunit plus one S subunit forming one asymmetric unit of the icosahedral capsid) is presented in Figure 2E. The L subunit is composed of two jellyroll β-barrel domains, while the S subunit forms a single β-barrel domain. The whole BBSV capsid is formed from 60 copies of the protomer according to a pseudo T = 3 icosahedral triangulation. As expected, the folds of the BBSV coat proteins and the capsid assembly are similar to those of other known comovirus structures. A structure-based multiple sequence alignment of all published comovirus CP structures (including the results of this work) is presented in Figures S2 and S3. The aligned C^{α} position root mean square deviation (RMSD) values between our BBSV structure and those of homologous CPs, calculated by optimising the superposition of either individual chains or the whole icosahedral assembly, are listed in Table 2. The majority of the residues from all of the homologous structures are superimposable with good Q^{res} alignment scores (as defined by (Roberts et al., 2006)). The RMSD of the aligned protomer C^{α} positions varies from a minimum of 0.77 Å for RCMV (VIPERdb entry 202) to a maximum of 1.09 Å for CPMV (PDB entry 5a32). It is worth noting that PDB entry 2bfu is an exception, as it shows medium/low structural alignment scores along its entire sequence. Since the PDB validation scores published for this structure are extremely poor, we have not taken into account structural alignment results for PDB entry 2bfu.

The superposed C^{α} traces of the aligned protomers are shown in Figure 5. The few observed structural differences (white to red regions) are mostly located on the external

surface of the comovirus capsids. They comprise (i) three stretches of the L subunit (residues 51–54, 207–210 and 270–278 in BBSV) defining the shape of an elongated cleft centred at the 2-fold symmetry axis; (ii) the L subunit C-terminal stretch (with unresolved residues in most structures); (iii) three loops (residues 399–402, 419–421 and 456–457 in BBSV) at the top of the turrets formed by the S subunits around the 5-fold axes; and (iv) the S subunit C-terminal stretch, which is unresolved in BBSV (residues 560–592); the corresponding region in CPMV is involved in penton stabilisation during RNA encapsidation and subject to cleavage at different sites during CPMV maturation. The remaining structural differences are located on the internal surface of the capsid. They comprise (v) the L subunit N-terminal stretch, which has similar conformations in all structures except empty BPMV capsids (PDB and PDB-REDO entries 1pgw) in which it is mainly disordered; (vi) the end of the first β-barrel domain within the L subunit (residues 185–189 in BBSV), which is more tightly folded in CPMV due to the presence of a disulphide bond which is absent in the protein from other viruses; and (vii) the S subunit N-terminal stretch (residues 378–385 in BBSV), which points towards the top of the 5-fold turrets in all structures except BPMV, in which it is folded in an opposite direction; as a result, like RCMV and CPMV, BBSV has channels along the capsid turret axes that are distinctly narrower (Figure 6) than those observed in BPMV (Lin et al., 2000). It should be noted, however, that in BBSV the very first residue of the S subunit (Asn378) is not bent towards the channel axis as it is in RCMV and CPMV, but rather points towards the top of the turret.

When the whole BBSV CP icosahedral assembly was superposed with homologous structures, an increase in the aligned C^{α} RMSD values was observed and ranged from +1.50 to +2.92 Å (Table 2), indicating slight differences in the protomer assembly. The

calculated radial distribution of atoms (Figure S4) shows that the refined atomic model of the BBSV capsid is slightly larger than those of other comoviruses, although comparable in thickness. It is not clear if these differences are real features of the protomer assemblies (cumulative effect of small, less than 0.5 Å, displacements of the protomers) or an artefact related to the cryoEM magnification accuracy.

The BBSV CP refined atomic model presented above does not account for all of the features observed within the capsid region in the reconstructed cryo-EM density map. Notably, additional density is visible on the external surface of the turrets located at the icosahedral 5-fold vertices (Figure 7). Indeed, the unsharpened cryo-EM map reveals a continuous two-lobed density forming a bridge between adjacent S subunits (Figure 7A, green dotted line, 1 and 2). In the sharpened cryo-EM map, well-resolved density can be observed only in the lobe 2 region (Figure 7B and C), which adopts an omega-like shape. The density of lobe 1 is in continuity with the last refined residue of the S subunit model (Asp559), while the omega-like density is attached to the β-strands of the adjacent S subunit. According to cDNA sequencing, the full-length S subunit contains 33 additional amino acids beyond Asp559, which could explain the presence of additional density. A tentative model corresponding to residues 560 to 567 (green sequence in Figure 1C) could be built into this region of the map where the density corresponding to Arg560 and Val561 is barely visible, that for Ser562 and Val563 are not visible at all and the subsequent four residues (564- Ala-Arg-Pro-Leu567) fit into the density contacting the neighbouring S subunit (Figure 4C). Numerous interactions with neighbouring S subunit are observed through Asp418, Glu523, Ser405, Arg471, Phe469 and Leu525. Due to the lack of continuity, we preferred to not include this part of BBSV in the final refined structure.

Figure 8 shows a cross-sectional view of an unsharpened BBSV map at 4.0 Å. Significant extra density is visible inside the capsid that can be ascribed to the packaged RNA genome. Wild-type BBSV particles contain either the full-length single-stranded RNA-1 or RNA-2; thus the density present inside the capsid is an averaged image of both RNA genomes. RNA appears as a concentric shell of density just beneath the capsid, protruding along the 5-fold symmetry axis. The shell of density is ~20 Å thick. Three different areas of interaction between the capsid and the averaged RNA shell were observed. The most extended interactions occur near the CP 5-fold axes through Arg382 (S subunits), while other interactions take place through the L subunits (residues Asn175 and Gln190).

DISCUSSION

The cryo-EM single particle analysis of BBSV virions allowed us to compute a final map of the RNA-filled BBSV virion at an average resolution of 3.22 Å. Local resolution estimation showed that many regions extend to higher resolution in the inner part of the capsid shell, while more disordered regions of the structure are associated with the penton turrets (Figure 4). We sequenced the cDNA of the BBSV strain corresponding to the S and L coat proteins and compared the CP sequences with those of CPMV, BPMV and RCMV (Figures S2 and S3). The sequence identity shared between the BBSV and CPMV, BPMV and RCMV sequences is 53%, 53% and 76% respectively. As expected, the overall structure of BBSV coat proteins is very similar to those of other known comovirus structures, with the icosahedral asymmetric unit formed from one S and one L subunit. Most structural differences are located on the external surface of the capsid, mostly near the 2-fold and 5-fold axes. Unlike CPMV and RCMV, BBSV does not exhibit a second, more internal, annular constriction of the channels running through the capsid along the 5-fold axes. This is not due to a completely different orientation of the S subunit N-terminal stretch (as in BPMV), but rather to a different conformation of just one (the first) amino acid, which renders the inner opening of the 5-fold channel slightly larger and more regularly shaped in BBSV than in CPMV and RCMV. Consequently, in the structural subgrouping organization of comovirus capsid structure proposed by Lin et al. (Lin et al., 2000), BBSV represents an intermediate state between CPMV and BPMV, defining then a new subgroup.

Our knowledge about the organisation of encapsidated genomic RNA within comovirus particles is limited. In RNA-2-filled BPMV particles, ordered ribonucleotides were observed by X-ray diffraction to adopt a trefoil shape near the 3-fold axes (Chen et al.,

1989). Differently filled (either RNA-1 or RNA-2) and empty CPMV particles could be separated by gradient ultracentrifugation and analysed by cryo-EM single-particle analysis (Hesketh et al., 2015; Hesketh et al., 2017), revealing several circular layers of RNA density. The discrete bridges of density observed in CPMV between the protein capsid and the RNA genome indicate RNA-1 interactions with Arg17 and Trp190, and RNA-2 interactions with Arg17 and Asn174, from the L subunit (corresponding to Lys17, Thr191 and Asn175 in BBSV). In our BBSV map, RNA density which results from averaged contributions of single-stranded RNA-1 and RNA-2 appears as a single layer of density protruding towards the 5-fold axes. We identified three main regions of contact with the BBSV capsid, corresponding to Asn175 and Gln190 of the L subunit and Arg382 of the S subunit. The BBSV L subunit residues involved in RNA interactions are at the same position (Asn175) or in close proximity (Gln190) to equivalent residues in CPMV, indicating that they are key positions for capsid-genome interactions in all comoviruses. Importantly, this is the first demonstration of an RNA genome interacting directly with the S subunits within the 5-fold channel axes of a comovirus. In this region, BBSV and RCMV have a highly similar S subunit N-terminal stretch in which the key Arg382 (Arg381 in RCMV) is conserved. It would be interesting to further investigate by cryo-EM the details of RNA-filled RCMV particles in order to establish if the presence of RNA in the 5-fold axis channels of comoviruses is related to the presence of several arginines.

As far as the BBSV capsid structure is concerned, lowest resolution regions in the reconstructed density were identified at vicinity of five fold symmetry axes. Two regions are observed at the upper surface of the capsid turrets. They correspond to two loops that have been previously shown as structurally variable between comoviruses, containing residues 399–402 and 419–421. Therefore, it seems that these loops are very flexible among comoviruses. An interesting lowest resolution region having an

elongated shape in the unsharpened map is located at the base of the pentameric unit forming the turret. The BBSV cryo-EM map permitted the unambiguous identification of the C-terminal part of the BBSV S chain up to residue Asp559. As the elongated density observed in the unsharpened map is in continuity with the last residue Asp559, a tentative model for eight further residues (560 to 567), forming a bridge between neighbouring S subunits within a penton, was then generated. Although there is no available information about the location of cleavage sites within the C-terminus of the BBSV S subunit, the presence of residual density in which residues up to Leu567 can be modelled and the sequence proximity of the corresponding main cleavage site on the CPMV S sequence (Leu563–Leu564) indicate that BBSV is also cleaved during maturation, here between Leu567–Leu568. The proposed conformation of BBSV residues 560–567 differs significantly from that observed in some CPMV structures, and includes various interactions with the neighbouring S subunit therefore stabilizing the pentameric unit forming the turret. It is very interesting because the interaction of the S subunit C-terminal region with neighbouring S subunit was shown to play a critical role in virus assembly and genome encapsidation (Hesketh et al., 2015; Lomonossoff and Johnson, 1991; Sainsbury et al., 2011; Taylor et al., 1999). Indeed, in CPMV, the interaction zone between a small helix formed by residues 564 to 576 of the S subunit with the neighbouring S subunits in unmatured CPMV; was proposed as key in stabilizing the formation of the penton and the subsequent formation of the CPMV particle. In BBSV, interactions of the residues 560–567 of S subunit with the neighbouring S subunits, strongly suggests a capsid assembly mechanism similar to that described for CPMV. According to sequence identity of capsid proteins among comoviruses, it also suggests a common capsid assembly mechanism throughout most of comovirus species.

CONCLUSIONS

The cryo-EM single particle analysis of BBSV allowed us to build a density map at 3.22 Å resolution and refine an atomic model of coat proteins forming the capsid of viral particles. This work provides a better understanding of BBSV particle organization and its internal RNA segment, allowing us notably to observe a direct interaction of the RNA genome with the S subunits within the 5-fold channel axes never described previously. The structural characterization of the small coat protein C-terminal domain allowed us to define a maturation cleavage between Leu567 and Leu568 and visualized interactions of the C-terminal stretch with neighbouring small coat proteins within the capsid pentameric turrets. These interactions, observed for the second time for a comovirus, strongly suggest a common capsid assembly mechanism throughout most of comovirus species. Therefore the structural study of BBSV strengthens our knowledge on structural organisation, genome encapsidation and assembly mechanisms within the comovirus family.

MATERIALS AND METHODS

Purification of the BBSV virus particles

One hundred grams of fresh Vicia faba FBNSV-infected leaves (with petioles) exhibiting symptoms were collected and homogenized in 20 mM citrate buffer containing 20 mM Na2SO3 (pH 6). Macerozyme at 0.1% final was added to the mixture, which was then shaken overnight at 4°C. The extract was filtered through cheesecloth (Miracloth), 25% of cold chloroform/iso-butanol (v/v) was progressively added to the filtrate before agitation at 4°C for 30 min. The mixture was then centrifuged at 15,200 g for 20 min (4°C) to separate the aqueous butanol (upper phase) phase containing virus particles from the chloroform phase (bottom phase) containing sedimented debris and solubilized lipids and pigments. The virus particles were precipitated by centrifugation at 302,000 g for 3 h (4°C) and the resulting pellet resuspended in 20 mM citrate buffer (pH 6). The virus particles suspension was finally deposited on a discontinuous cesium chloride gradient (10-15-20-30-40%) and centrifuged at 201,000 g for 4 h (4°C). The bands corresponding to viral material were collected and centrifuged again (302,000 g for 3h at 4°C). The final pellet was resuspended in 20 mM phosphate buffer (pH 7.4).

CP Sequencing

Amino-acid sequences of the two coat proteins of our BBSV strain were assessed by analysis of the nucleotide sequences of cDNA clones.

Cryo-EM sample preparation and imaging

Three microliters of purified BBSV viral particles at ~2 mg/ml were applied to glow-discharged Quantifoil R 2/2 grids (Quantifoil Micro tools GmbH, Germany), blotted for 1s and then flash frozen in liquid ethane using the semi-automated plunge freezing device CP3 (Gatan inc.) at 95% relative humidity. Preliminary images were recorded using a JEOL 2200 FS electron microscope operating at 200 kV in zero-energy-loss mode with a slit width of 20 eV and equipped with a 4k x 4k slow-scan CDD camera (Gatan inc.). High resolution data were collected using a spherical aberration (Cs) corrected Titan Krios S-FEG transmission electron microscope (Thermo Fisher Scientific, Eindhoven, Netherlands) (IGBMC-CERBM, Strasbourg, France) operating at 300 kV at a nominal magnification of X 59,000 under low-dose conditions and with defocus ranging from -0.8 to -2.5 µm. The pixel size estimated at the specimen level was 1.13 Å (this estimation could be further refined to 1.11 Å during atomic model building, see below). A total of 2,899 movies were recorded using the EPU automated acquisition software on a Falcon II camera with dose fractionation. The total exposure time was 2 s corresponding to a dose of 45 $e^-/Å^2$. Seven individual frames were collected with an electron dose of 3.5 $e^-/Å^2$ per frame, therefore giving to a total dose of 24.5 $e^-/Å^2$. The first frame was collected after an exposition corresponding to an electron dose of 1.25 $e^-/Å^2$.

Processing of Titan Krios S-FEG movies and high-resolution reconstruction

The seven frames of each movie were computationally corrected for drift and beam-induced movement in RELION-3 (Zivanov et al., 2018). The contrast transfer function of

each micrograph was determined using Gctf program (Zhang, 2016). On 2,899 micrographs, only 1,779 micrographs were used for further processing and 37,326 particles were automatically selected using the Gautomatch program. A first round of reference-free two-dimensional (2D) classification allowed us to select 27,648 particles that were used to compute a *de novo* 3D model. We then used 3D classification into four classes to select a subset of 17,233 particles contributing to the best class. Three-dimensional refinement yielded 3.95 Å resolution. We then proceeded with per-particle defocus refinement and Bayesian polishing, improving the resolution to 3.3 Å. The final resolution after post-processing was then estimated at 3.22 Å by the so-called gold-standard Fourier shell correlation (FSC) using the 0.143 criterion (Scheres and Chen, 2012). The final resolution value was further validated by the high resolution noise substitution method (which has proven to be a robust way to assess the amount of noise in high spatial frequencies (Chen et al., 2013)), whereby the FSC of a data set with randomized phases beyond 4.4 Å was subtracted from the FSC of the original data set. The refined maps were post-processed through model masking and *B*-factor sharpening (using an empirically derived *B* value of −101 $Å^2$). Local resolution was estimated using Blocres from Bsoft package (Heymann, 2001). Figures were prepared using Chimera (Pettersen et al., 2004) and Pymol (Schrodinger, 2015).

Model Building and refinement

A preliminary atomic model of the BBSV capsid was set up by rigid-body fitting the structure of a CPMV CP protomer (PDB-REDO entry 1ny7) into the BBSV cryo-EM reconstructed density map. At the same time, the value of the map voxel size (initially estimated at 1.13 Å, see *§ Cryo EM sample preparation and imaging*) was refined by

maximizing the correlation between the model and the density map (final voxel size: 1.11 Å). The amino acid side-chains were then manually replaced and fit into the density based on the BBSV CP sequence. This preliminary model was iteratively refined against the experimental density map – following the general guidelines described by (Brown et al., 2015) and (Afonine et al., 2018a) by alternated steps of automatic refinement under stereochemical restraints with REFMAC5 (Murshudov et al., 2011) and PHENIX.real_space_refine (Afonine et al., 2018b), manual model inspection and correction with Coot (Emsley and Cowtan, 2004) and model geometry validation with MolProbity (Chen et al., 2010). In order to save computation time and memory and, at the same time, correctly avoid steric clashes between symmetry related neighboring chains, only a part of the whole icosahedral capsid density was taken into account, which consisted of six L chains and five S chains: one central protomer (one L and one S chain), two neighboring L chains and four neighboring S chains grouped around a 5-fold symmetry axis, two neighboring L chains related to the central protomer by a 3-fold symmetry axis, and one L chain related to the central protomer by a 2-fold symmetry axis. During refinement, tight non-crystallographic symmetry (NCS) restraints were applied in order to maintain equivalent chains in the same conformation. Standard stereochemical restraints with optimized weights were applied during all refinement steps, while secondary structure and homology restraints generated by ProSMART (Brown et al., 2015) and derived from two CPMV CP structures (PDB entry 5a33 and PDB-redo entry 1ny7) were applied at the beginning of the refinement only and then gradually released in the final steps. Atomic isotropic displacement parameters (one parameter per residue) were applied and refined in the final steps. The degree of model overfitting into noise was assessed according to the procedure described by (Brown et al., 2015): the model atomic coordinates were randomly modified (0.5 Å maximum

shift), the perturbed model was re-refined against a half-data reconstructed map (work map), and the FSC's between the re-refined model (extended by applying full icosahedral symmetry) and either the work map or the remaining half-data map (free map) (Figure 3B).

CP structure based alignments and analysis

The sequence and structure of the BBSV capsid protein has been compared to all the RCMV, CPMV and BBSV CPs structures available from public databases (PDB, http://www.rcsb.org; PDB-REDO, https://pdb-redo.eu/; and VIPERdb, http://viperdb.scripps.edu/) (13 structures, see Table 2). Structure-based alignments were carried out automatically using STAMP (Russell and Barton, 1992) and Multiseq (Roberts et al., 2006) and manually corrected in order to include the structurally unresolved residues listed in the "REMARK 465" cards from the PDB files or described in the literature as being present though unresolved. Structures were compared both at the protomer level (in order to compare chain folds) by optimizing the superposition of individual chains and at the whole assembly level (in order to compare capsid sizes) by aligning the set of icosahedral symmetry axes (thus without distorting the icosahedral assemblies).

The spherically averaged capsid atom density (number of non-H atoms per unit volume) was calculated as a function of the distance (radius) from the capsid center (Figure 6B) by counting (MOLEMAN2 (Kleywegt, 1997)) all the atoms contained within spherical shells with varying radii (radius step 0.667 Å, shell thickness: 2 Å). The cylindrically averaged atom density around the icosahedral symmetry axes (Figure 6A) was

calculated (MOLEMAN2) by counting the number of atoms contained within thin cylindrical shells coaxial with the symmetry axis and having varying radii (r) and positions (z) along the symmetry axis (z and r step: 1 Å; shell z- and r-thickness: 3.76 Å).

SUPPLEMENTARY INFORMATION

Supplemental information includes 3 figures and can be found with this article online at http://dx.doi.org/……..

ACCESSION NUMBERS

The refined 3.22 Å resolution map of the BBSV capsid has been deposited with the Electron Microscopy Data Bank (entry EMD-4504). The refined atomic model of the BBSV capsid has been deposited with the Protein Data Bank (entry 6QCC).

ACKNOWLEDGMENTS

F.L. acknowledges funding from Gouvernement de Nouvelle Calédonie. This work was supported by the Centre National de la Recherche Scientifique (CNRS), the Institut National de la Santé et de la Recherche Médicale (INSERM), the University of Montpellier and from the French Infrastructure for Integrated Structural Biology, a national infrastructure supported by the French National Research Agency (ANR-10-INBS-05).

REFERENCES

Afonine, P.V., Klaholz, B.P., Moriarty, N.W., Poon, B.K., Sobolev, O.V., Terwilliger, T.C., Adams, P.D., Urzhumtsev, A., 2018a. New tools for the analysis and validation of cryo-EM maps and atomic models. Acta crystallographica. Section D, Structural biology 74, 814-840.

Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., Adams, P.D., 2018b. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta crystallographica. Section D, Structural biology 74, 531-544.

Al-Khalaf, M., Makkouk, K., Kasem, A.H., 2002. Seed transmission of broad bean stain virus in lentil with respect to genotype variability and seed size. Arab Journal of Plant Protection 20, 106-110.

Brown, A., Long, F., Nicholls, R.A., Toots, J., Emsley, P., Murshudov, G., 2015. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. Acta crystallographica. Section D, Biological crystallography 71, 136-153.

Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H., Henderson, R., 2013. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy 135, 24-35.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta crystallographica. Section D, Biological crystallography 66, 12-21.

Chen, X., Bruening, G., 1992. Cloned DNA copies of cowpea severe mosaic virus genomic RNAs: infectious transcripts and complete nucleotide sequence of RNA 1. Virology 191, 607-618.

Chen, Z.G., Stauffacher, C., Li, Y., Schmidt, T., Bomu, W., Kamer, G., Shanks, M., Lomonossoff, G., Johnson, J.E., 1989. Protein-RNA interactions in an icosahedral virus at 3.0 A resolution. Science 245, 154-159.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta crystallographica. Section D, Biological crystallography 60, 2126-2132.

Fiedorow, Z., Szlachetka-Wawrzyniak, E., 2002. Transmission of broad bean stain virus (BBSV) by seeds of pea (Pisum sativum L.). Plant Breed Seed Sci 46, 81–88.

Gibbs, A., Giussani-Belli, G., Smith, H.G., 1968. Broad-bean stain and true broad-bean mosaic viruses.

Hesketh, E.L., Meshcheriakova, Y., Dent, K.C., Saxena, P., Thompson, R.F., Cockburn, J.J., Lomonossoff, G.P., Ranson, N.A., 2015. Mechanisms of assembly and genome packaging in an RNA virus revealed by high-resolution cryo-EM. Nature communications 6, 10113.

Hesketh, E.L., Meshcheriakova, Y., Thompson, R.F., Lomonossoff, G.P., Ranson, N.A., 2017. The structures of a naturally empty cowpea mosaic virus particle and its genome-containing counterpart by cryo-electron microscopy. Scientific reports 7, 539.

Heymann, J.B., 2001. Bsoft: image and molecular processing in electron microscopy. Journal of structural biology 133, 156-169.

Huynh, N.T., Hesketh, E.L., Saxena, P., Meshcheriakova, Y., Ku, Y.C., Hoang, L.T., Johnson, J.E., Ranson, N.A., Lomonossoff, G.P., Reddy, V.S., 2016. Crystal Structure and Proteomics Analysis of Empty Virus-like Particles of Cowpea Mosaic Virus. Structure 24, 567-575.

Kleywegt, G.J., 1997. Validation of protein models from Calpha coordinates alone. Journal of molecular biology 273, 371-376.

Kumari, S.G., Makkouk, K., 1993. Evaluation of different ELISA procedures for the detection of pea seed-borne mosaic potyvirus and broad bean stain comovirus in lentil leaf extracts. Arab Journal of Plant Protection 11, 86-91.

Kumari, S.G., Makkouk, K., 1996. Transmission of broad bean stain comovirus and broad bean mottle by weevils in Syria. Phytopathologia Meditterranea 35, 124-126.

Lin, T., Chen, Z., Usha, R., Stauffacher, C.V., Dai, J.B., Schmidt, T., Johnson, J.E., 1999. The refined crystal structure of cowpea mosaic virus at 2.8 A resolution. Virology 265, 20-34.

Lin, T., Clark, A.J., Chen, Z., Shanks, M., Dai, J.B., Li, Y., Schmidt, T., Oxelfelt, P., Lomonossoff, G.P., Johnson, J.E., 2000. Structural fingerprinting: subgrouping of comoviruses by structural studies of red clover mottle virus to 2.4-A resolution and comparisons with other comoviruses. Journal of virology 74, 493-504.

Lin, T., Johnson, J.E., 2003. Structures of picorna-like plant viruses: implications and applications. Advances in virus research 62, 167-239.

Lomonossoff, G.P., Johnson, J.E., 1991. The synthesis and structure of comovirus capsids. Progress in biophysics and molecular biology 55, 107-137.

Mabrouk, O., Mansour, A.N., 1998. Effect of pea seedborne mosaic and broad bean stain viruses on lentil growth and yield in Jordan. Scientia Horticulturae 73, 175-178.

Makkouk, K., Azzam, O.I., 1986. Detection of broad bean stain virus in lentil seed groups. LENS Newsletter 13, 37-38.

Makkouk, K., Pappu, H., Kumari, S.G., 2012. Virus diseases of peas, beans, and faba bean in the Mediterranean region. Advances in virus research 84, 367-402.

Mali, V.R., Subr, Z., Kudela, O., 2003. Seed transmission of como and potyviruses in fababean and vetch cultivars introduced into Slovakia. Acta Phytopathologica and Entomologica Hungarica 38, 87-97.

Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta crystallographica. Section D, Biological crystallography 67, 355-367.

Ochoa, W.F., Chatterji, A., Lin, T., Johnson, J.E., 2006. Generation and structural analysis of reactive empty particles derived from an icosahedral virus. Chemistry & biology 13, 771-778.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry 25, 1605-1612.

Roberts, E., Eargle, J., Wright, D., Luthey-Schulten, Z., 2006. MultiSeq: unifying sequence and structure data for evolutionary analysis. BMC bioinformatics 7, 382.

Russell, R.B., Barton, G.J., 1992. Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. Proteins 14, 309-323.

Sainsbury, F., Saunders, K., Aljabali, A.A., Evans, D.J., Lomonossoff, G.P., 2011. Peptide-controlled access to the interior surface of empty virus nanoparticles. Chembiochem : a European journal of chemical biology 12, 2435-2440.

Scheres, S.H., Chen, S., 2012. Prevention of overfitting in cryo-EM structure determination. Nature methods 9, 853-854.

Schrodinger, LLC, 2015. The PyMOL Molecular Graphics System, Version 1.8.

Taylor, K.M., Spall, V.E., Butler, P.J., Lomonossoff, G.P., 1999. The cleavable carboxyl-terminus of the small coat protein of cowpea mosaic virus is involved in RNA encapsidation. Virology 255, 129-137.

Zhang, K., 2016. Gctf: Real-time CTF determination and correction. Journal of structural biology 193, 1-12.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., Scheres, S.H., 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife 7.

FIGURE LEGENDS

Figure 1. The BBSV genome and protein organization. (A) Organization of the two RNA segments composing the genome of BBSV. RNA-1 encoded poly-protein is here proposed to encompass non-structural products according to homologies with the related genome of Cowpea mosaic virus. RNA-2 encoded poly-protein contains a movement protein and the coat proteins (large and small subunits). The parts in color correspond to known sequenced regions. **(B)** Amino acid sequence of BBSV coat proteins derived from cDNA sequencing. Residue substitutions with respect to UniProt sequence D0PSV7 are highlighted. Residues for which atomic coordinates have been computed and refined based on the cryo EM map are colored in blue (large CP) and red (small CP). Residues initially unresolved are black. **(C)** Protein organization of the C-terminal region of the small coat protein. The attribution of residues 560 to 567 (green) results in the further inspection of the BBSV cryo-EM map (green domains in Figure 7).

Figure 2. The cryo-EM structure of the BBSV particle. (A) Isosurface representation of the icosahedral capsid density reconstructed at 3.22 Å resolution. The surface is colored according to the distance from the capsid center as indicated by the scale bar. **(B)** Close-up view on the turret around the top 5-fold axis. **(C and D)** Two details of the density within the L and S subunits respectively; the corresponding refined atomic models are shown. **(E)** Ribbon representation of the BBSV capsid protomer.

Figure 3. Resolution estimation and validation. (A) Curves of the 3D Fourier shell correlation (FSC) between two independent data set halves. FSC_{data}: gold standard FSC

calculated using the original data; FSC_{random}: FSC calculated after randomization of phases at frequencies higher than 4.4 Å; FSC_{true}: FSC of the original data after subtraction of FSC_{random}; FSC_{true} is nearly identical to FSC_{random} up to 3.22 Å resolution. **(B)** Atomic model overfitting test. The FSC curves of the BBSV CP atomic model re-refined against one half dataset map only (after random perturbation of the atomic coordinates) are shown: FSC_{work} is calculated between the re-refined model and the half dataset map against which the model was re-refined, while FSC_{free} is calculated between the re-refined model and the remaining half dataset map. The two curves are extremely similar, indicating that the degree of model overfitting into noise is very small up to the maximum resolution used during model refinement.

Figure 4. Local resolution map. (A) Surface view of the unsharpened cryo-EM map colored by local resolution as evaluated by Blocres. **(B)** Cross-sectional view of the unsharpened cryo-EM map colored by local resolution. **(C)** Close view on a capsid turret around a 5-fold axis. Disordered regions located at the upper surface of a S subunit forming the capsid turret and those located at the base of the pentameric units are indicated by black and white arrows respectively.

Figure 5. Superposed C-alpha traces of known comovirus capsid protomer structures. The atomic structures used here are listed in Table 2; except for PDB entry 1bfu, which is not included in the figure. Colors according to Q^{res} structural similarity score (see Materials and Methods). The small subunit occupies the upper half of the picture.

Figure 6. Comparision of the averaged atom density distributions of the BBSV capsid with those of other known comovirus capsid structures. (A) Circularly averaged atom density (number of non-H atoms per unit volume) around a 5-fold symmetry axis as a function of the distances *z* (along the axis) and *r* (from the axis); density values at negative *r* are just a copy of those at positive *r*. **(B)** Spherically averaged atom density volume as a function of the distance from the capsid center.

Figure 7. The structure of the C-terminal extension of the small subunit. (A) View of the turret located at the icosahedral 5-fold symmetry axis in the unsharpened BBSV map. An unattributed continuous two-lobed density is surrounded by green dotted lines and labeled 1 and 2. It forms a bridge between adjacent subunits of the turret. **(B)** Close-up view of the area 1 and 2 in the sharpened BBSV map. The extra densities are colored in green. **(C)** Proposed atomic model for the C-terminal extension of the S subunit.

Figure 8. Interactions between the capsid protein and RNA genome. (A) Cross-sectional view of the unsharpened BBSV map. The cutting plane contains 5-, 3- and 2-fold symmetry axes. **(B)** Close-up view of the three interaction sites between the capsid and the RNA genome. **(C)** Close-up view of the two interactions between the RNA genome and residues Asn175 and Gln190 from the large subunit of the capsid protein. **(D)** Close-up view of the interaction between the RNA genome and residue Arg382 of the small subunit of the capsid protein.

A

RNA-1

RNA-2

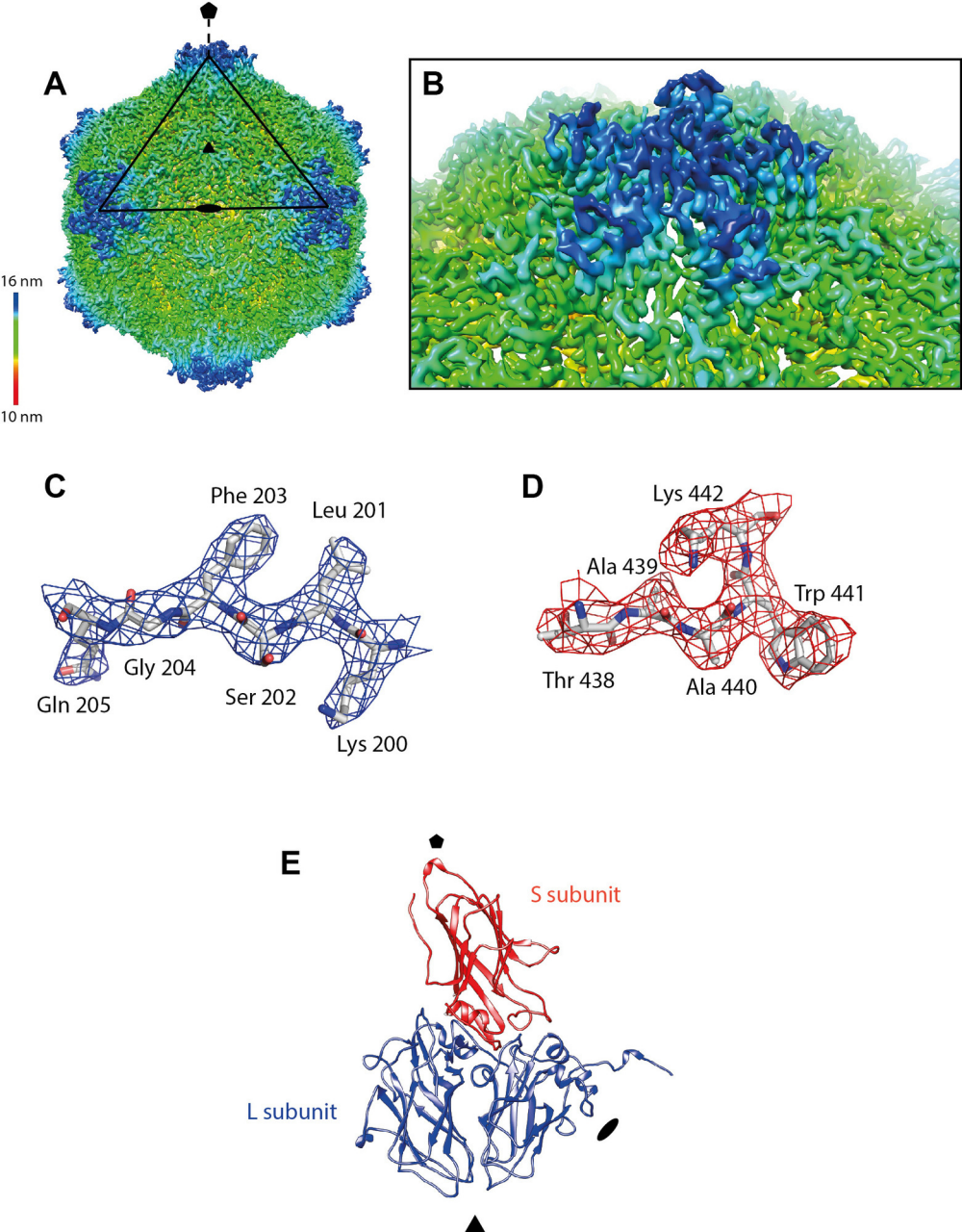
B

```
   1 MDVDLFKLSL  DDTSSVKGSL  LDTRFAQVRV  VIPKAMAGGN  ELLNSNLYDI   50
                                                     (N)
  51 LVVDNNFRAA  AALAHTHIIE  GQIKCVCTIN  LPENTGCCLA  LCVNSSNRGQ  100
       (G)
 101 FSTDIYTIGS  QDRMLWNPAC  SKNSTFTFNP  NPCGTGWSLE  FLRRTKFHIS  150
                                          (S)
 151 VVCVSGWSAQ  PQTDLVMTMD  FFVANVPCVP  RIYNLGSPGQ  TLWLNRWMGK  200

 201 LSFGQGVSND  IKSMPLAIGG  GAGAKDSILM  NMTNAYLSLW  RYFHGDLVFE  250

 251 VNKMSSPYIK  STVTFFIGFG  GVSFQPELED  FPNKLVQFSE  VQEKIELKFT  300
                             (E)
 301 RAEFLTAWST  QVDPAAQLAN  DGCPYLYAMV  HDSTASTIVG  DFNLGVTLTR  350

 351 IENFAGIGCN  PGIQGARLLG  SAIATPQNAV  VRSSPGIYSN  CFSLRAPLKP  400
                                          (I)
 401 DGPKSFTCDL  MGGGVVTDGD  TGWQVTVRNT  PVSNLLRTAA  WKRGTVHVQV  450
       (L)
 451 VLAGASVKRS  DWDSTVQIFL  RQSMATSSYD  AKIWDICQPG  AAMLEFSFDV  500

 501 VGPNSGFEMW  DSNWASQTSW  FLEFLISNPA  QNTLFEVNLR  LDENFSVAGT  550

 551 TLMPPFVLDR  VSVARPLLGK  QTKTVARSAR  VVRETKEASE  SP         592
                             (I)
```

C

```
551 559 560   567 568                      592
|     | |       |▲| +  +   +  +  +- +-  -  |
MPPFVLD RVSVARPL LGKQTKTVARSARVVRETKEASESP
                ▼
           cleavage
```

Figure 1


A
16 nm
10 nm
B
C
Phe 203
Leu 201
Gly 204
Gln 205
Ser 202
Lys 200
D
Lys 442
Ala 439
Trp 441
Thr 438
Ala 440
E
S subunit
L subunit

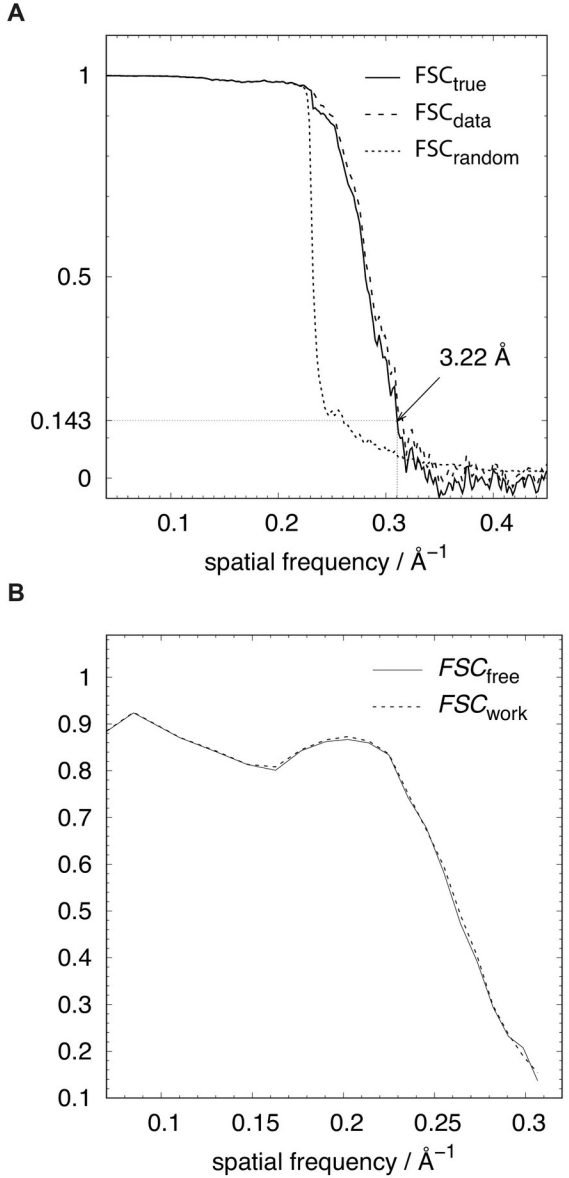

A
1
0.5
0.143
0
FSC_true
FSC_data
FSC_random
3.22 Å
0.1
0.2
0.3
0.4
spatial frequency / Å^-1
B
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
FSC_free
FSC_work
0.1
0.15
0.2
0.25
0.3
spatial frequency / Å^-1

Figure 3

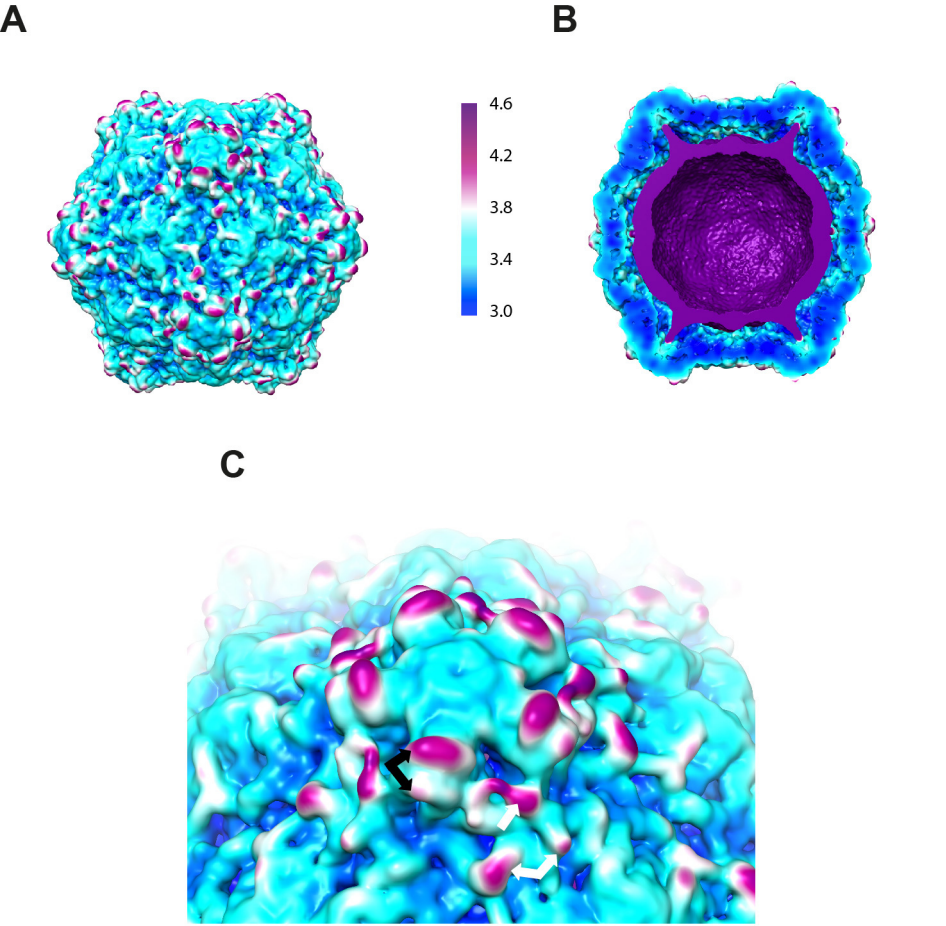


Figure 4

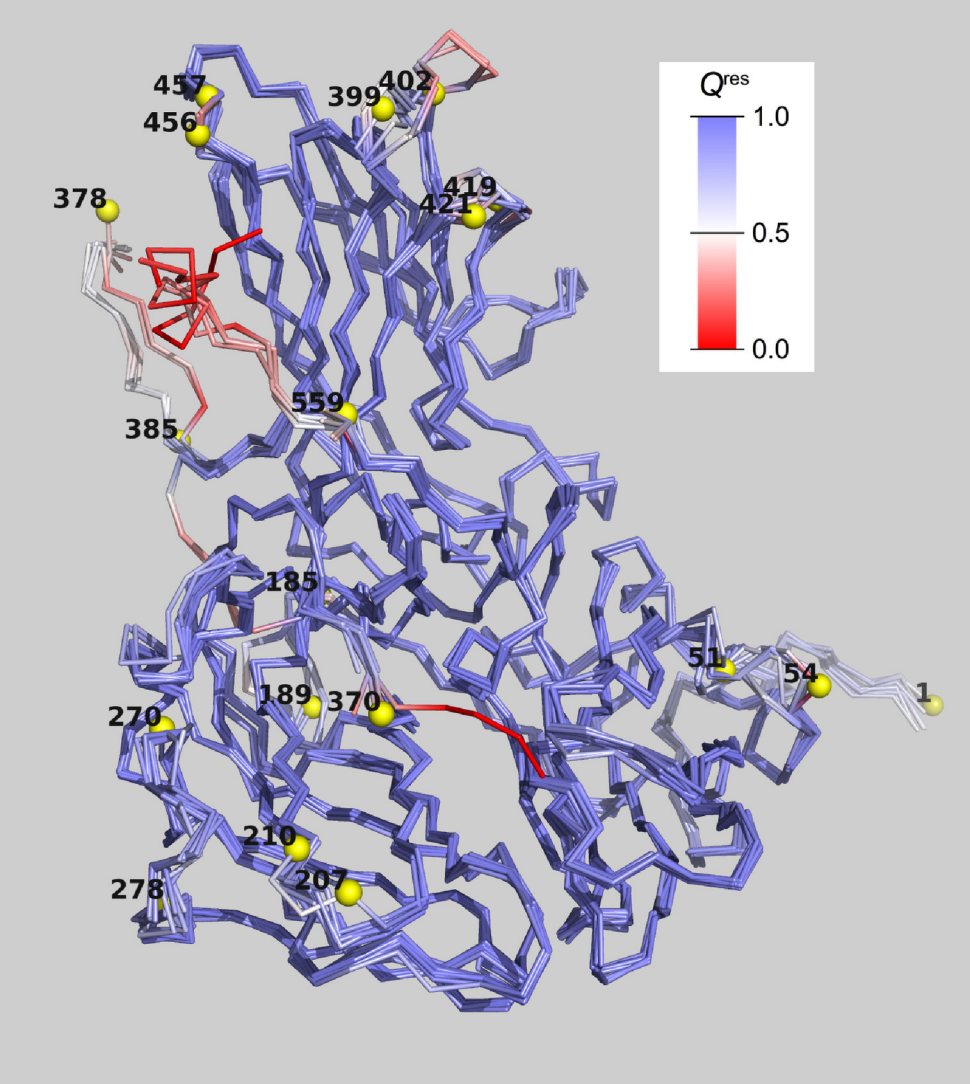

Qres
1.0
0.5
0.0
457
456
399
402
378
419
421
559
385
185
51
54
1
189
370
270
210
207
278

Figure 5

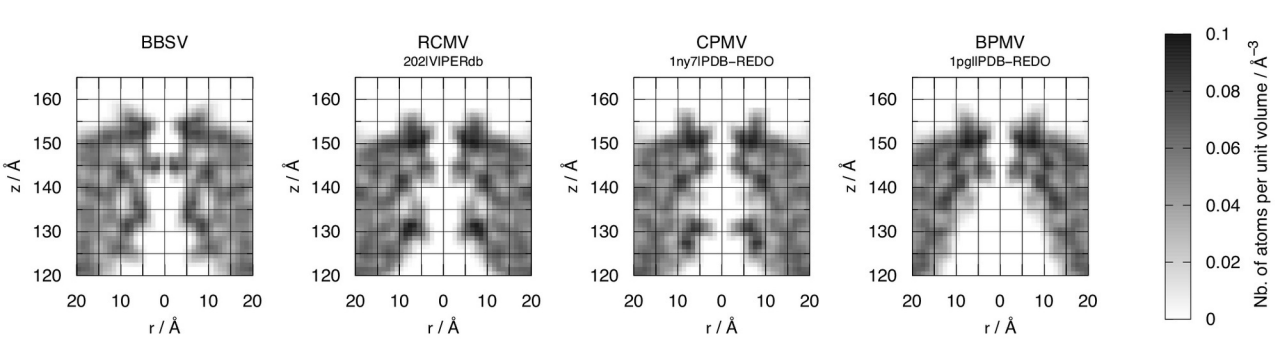

BBSV
RCMV
202lVIPERdb
CPMV
1ny7lPDB-REDO
BPMV
1pgllPDB-REDO
z / Å
r / Å
160
150
140
130
120
20
10
0
10
20
0.1
0.08
0.06
0.04
0.02
0
Nb. of atoms per unit volume / Å^{-3}

Figure 6

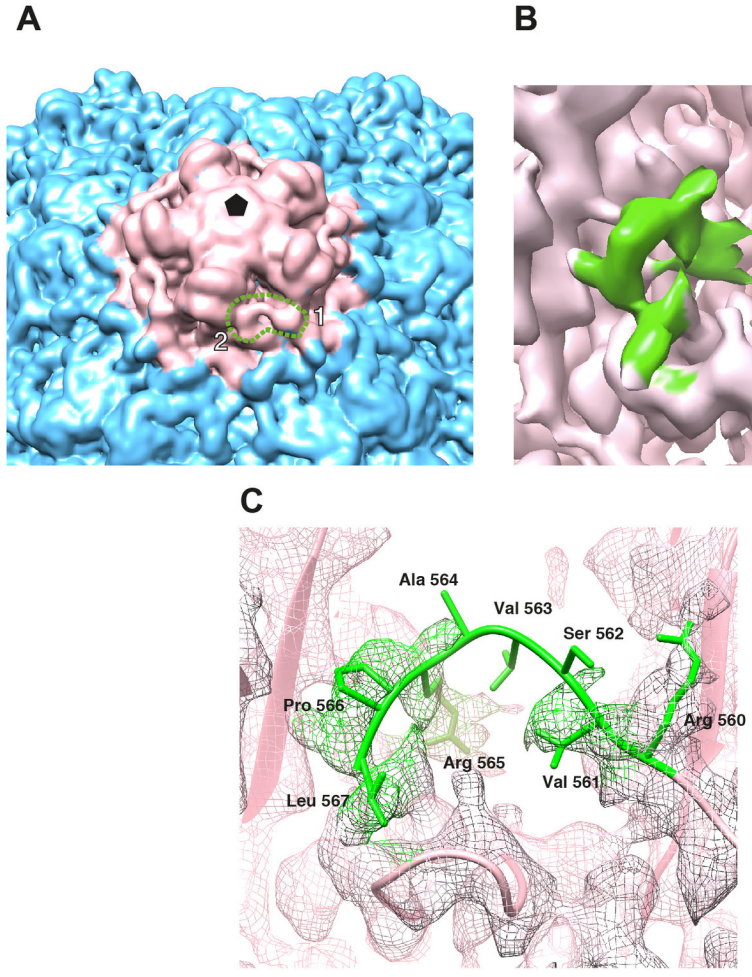

A
B
C
1
2
Ala 564
Val 563
Ser 562
Pro 566
Arg 560
Arg 565
Val 561
Leu 567

Figure 7

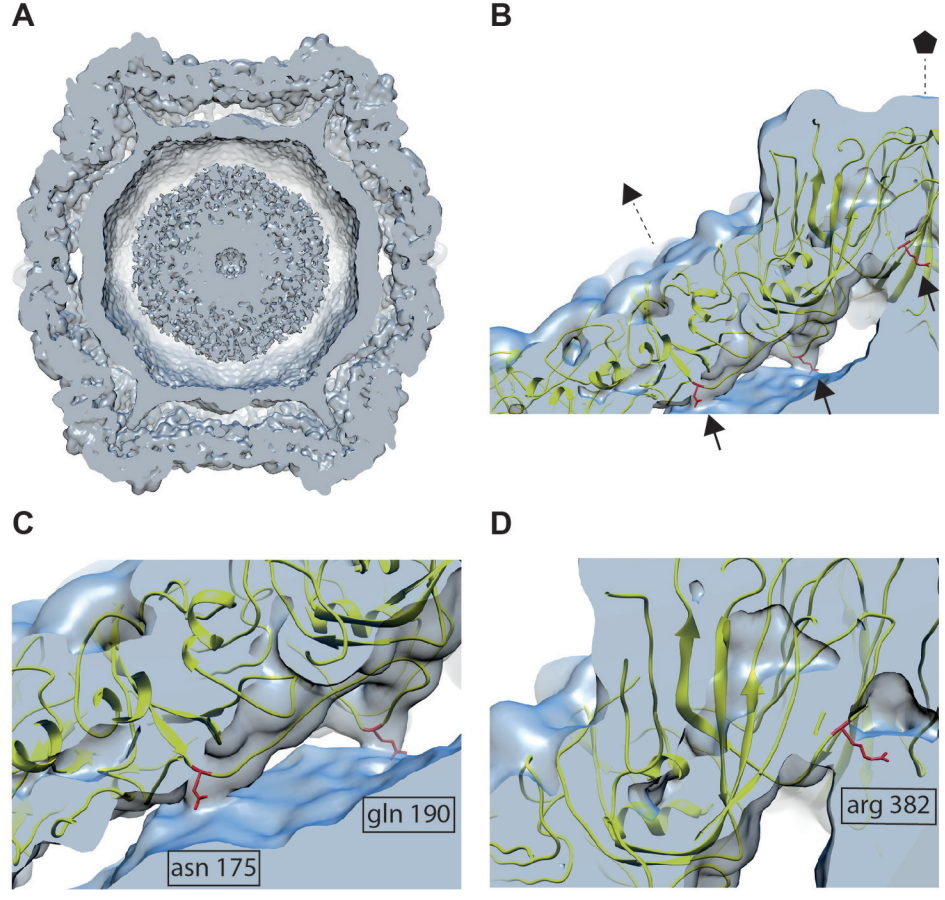


Figure 8

Single particle 3D reconstruction	
Data collection	
Electron source voltage	300 kV
Detector	Falcon II
Pixel size (unrefined value)	1.13 Å
Total dose	24.5 e^{-}/Å
Image processing	
Nb. of micrographs	1,779
Nb. of particles	17,233
Refined map pixel size	1.11 Å
Final resolution (@1.11 Å/pix)	3.22 Å
Sharpening *B*-factor	–101 $Å^2$
Atomic model refinement	
PHENIX real-space refinement statistic	
Model-to-map fit (CC_mask)	0.8591
Nb. Of refined atoms per asymmetric unit	4253
Min., mean and max. grouped atomic *B*-factor	30/49/84 $Å^2$
Molprobity validation statistics	
Clashscore	4.48
Poor / favored rotamers	5.96% / 84.15%
Ramachandran outliers / favored	0% / 96.5%
MolProbity score	2.04
C^{β} deviations >0.25 Å	0%
Bad bonds / angles	0% / 0%
Cis Prolines	0%
C^{α} Geometry / BLAM outliers	0.55% / 1.84%

Table 1 : Summary of single particle analysis parameters and atomic model refinement statistics.

virus	database entry		optimal rmsd / Å	
			single protomer chains	icosahedral assembly
RCMV	202	VIPERdb	0.77	1.67
CPMV	1ny7	PDB	1.03	1.96
CPMV	1ny7	PDB-REDO	1.00	1.95
CPMV	5a32	PDB	1.09	2.92
CPMV	5a33	PDB	1.02	2.05
CPMV	5fmo	PDB	1.03	2.09
CPMV	5fmo	PDB-REDO	1.02	2.28
CPMV	2bfu	PDB	1.90	2.54
BPMV	1bmv	PDB	1.00	1.50
BPMV	1pgl	PDB	0.98	1.80
BPMV	1pgl	PDB-REDO	0.94	1.77
BPMV	1pgw	PDB	0.98	1.72
BPMV	1pgw	PDB-REDO	0.95	1.69

Table 2 : Global structural similarity between BBSV and known comovirus CP structures. The root mean square deviations (RMSD) of the aligned (see Figures S2 and S3) CP Cα positions after optimal superposition of either individual chains or whole icosahedral assemblies are reported.