



An Examination of the Electrostatic Interactions between the N-Terminal Tail of the Brome Mosaic Virus Coat Protein and Encapsidated RNAs

Peng Ni¹, Zhao Wang², Xiang Ma³, Nayaran Chandra Das⁴, Paul Sokol^{4,5}, Wah Chiu², Bogdan Dragnea³, Michael Hagan^{6*} and C. Cheng Kao^{1*}

¹Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, IN 47405, USA

²National Center for Molecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA

³Chemistry Department, Indiana University, Bloomington, IN 47405, USA

⁴Center for Exploration of Energy and Matter, Indiana University, Bloomington, IN 47408, USA

⁵Physics Department, Indiana University, Bloomington, IN 47405, USA

⁶Physics Department, Brandeis University, Waltham, MA 02454, USA

Received 1 November 2011;
received in revised form
17 March 2012;
accepted 26 March 2012
Available online
1 April 2012

Edited by D. E. Draper

Keywords:

Brome Mosaic Virus;
capsid;
arginine-rich tail;
RNA encapsidation;
electrostatic interaction

The coat protein of positive-stranded RNA viruses often contains a positively charged tail that extends toward the center of the capsid and interacts with the viral genome. Electrostatic interaction between the tail and the RNA has been postulated as a major force in virus assembly and stabilization. The goal of this work is to examine the correlation between electrostatic interaction and amount of RNA packaged in the tripartite Brome Mosaic Virus (BMV). Nanoindentation experiment using atomic force microscopy showed that the stiffness of BMV virions with different RNAs varied by a range that is 10-fold higher than that would be predicted by electrostatics. BMV mutants with decreased positive charges encapsidated lower amounts of RNA while mutants with increased positive charges packaged additional RNAs up to ~900 nt. However, the extra RNAs included truncated BMV RNAs, an additional copy of RNA4, potential cellular RNAs, or a combination of the three, indicating that change in the charge of the capsid could result in several different outcomes in RNA encapsidation. In addition, mutant with specific arginines changed to lysines in the capsid also exhibited defects in the specific encapsidation of BMV RNA4. The experimental results indicate that electrostatics is a major component in RNA encapsidation but was unable to account for all of the observed effects on RNA encapsidation. Thermodynamic modeling incorporating the electrostatics was able to predict the approximate length of the RNA to be encapsidated for the majority of mutant virions, but not for a mutant with extreme clustered positive charges. Cryo-electron microscopy of virions that encapsidated an additional copy of RNA4 revealed that, despite the increase in RNA encapsidated, the capsid structure was

*Corresponding authors. E-mail addresses: hagan@brandeis.edu; ckao@indiana.edu.

Abbreviations used: BMV, Brome Mosaic Virus; CP, coat protein; CsCl, cesium chloride; WT, wild type; DSF, differential scanning fluorimetry; UTR, untranslated region; EM, electron microscopy; AFM, atomic force microscopy; NIH, National Institutes of Health.

minimally changed. These results experimentally demonstrated the impact of electrostatics and additional restraints in the encapsidation of BMV RNAs, which could be applicable to other viruses.

© 2012 Elsevier Ltd. All rights reserved.

Introduction

Viruses have evolved distinct mechanisms to ensure the proper encapsidation of their highly negatively charged genomes. Among DNA viruses, SV40 forms minichromosomes using cellular histones while adenovirus complexes its DNA with an adenovirus-encoded protein that is basic in overall charge.^{1,2} Negative-stranded RNA viruses have nucleocapsid proteins bind their RNAs through a positively charged cleft.³

In positive-stranded RNA viruses, a common motif in the viral coat proteins (CPs) is a relatively unstructured peptide tail rich in basic amino acid residues. These tails extend into the center of the virions and are largely unresolved in the crystallographic maps, implicating their interactions with the RNA genome.⁴⁻⁹ The nonspecific electrostatic interaction between these tails and RNA had been proposed to constrain the length of viral genome. For instance, Belyi and Muthukumar, as well as Hu *et al.*, reported a striking linear correlation between the total number of positive charges in the tails and the length of the encapsidated RNA.^{10,11} In studies of various viruses, partial deletions of the positively charged residues have resulted in virions that packaged a reduced amount of viral RNAs.¹²⁻¹⁵ Here, we use Brome Mosaic Virus (BMV) to systematically examine the influence of electrostatics on RNA packaging.

BMV is well suited for analysis of selective encapsidation of RNA.¹⁶ The BMV RNAs are composed of four RNAs, named RNA1 (3.2 kb), RNA2 (2.9 kb), RNA3 (2.1 kb), and RNA4 (0.9 kb). During infection, RNA1 and RNA2 encode, respectively, the 1a and 2a proteins that function in the replicase complex. The dicistronic RNA3 is used to translate movement protein for the systemic infection of BMV in plants. RNA4 is transcribed from a subgenomic promoter in minus-strand RNA3 and directs the translation of the CP.

Infectious BMV contains three separate particles.¹⁷ RNA3 is encapsidated at a 1:1 ratio with RNA4 while RNA1 and RNA2 are individually encapsulated into separate capsids. The three types of BMV particles are indistinguishable, with 180 copies of the CP arranged with a $T=3$ icosahedral symmetry. Furthermore, the division of the RNAs in the particles allows each to contain a total RNA length of between 2.9 and 3.2 kb (the sum of RNA3 and RNA4 is 3.0 kb). The ratios of the three particles, as identified by the relative abundances of the four BMV RNAs, are also

kept fairly constant in a typical infection; in *Nicotiana benthamiana* plants, particles containing RNA1 and RNA2 are approximately equal in abundance while the particles containing RNA3/RNA4 are in several fold excess relative to the first two particles.

The CP is 189 residues in length with 8 positively charged residues among the first 26. The density of the entire C subunit in an asymmetric trimer was elucidated, but residues 1-40 and 1-25 were respectively missing for the A and B subunits, presumably unresolved due to their interaction with the RNA.¹⁸ Deletion of residues 2-25 of the BMV CP failed to yield viable virions in plants, and a truncated CP without residues 1-35 only formed empty $T=1$ capsids *in vitro*.^{19,20} Other BMV mutants with selectively removed positively charged residues in the N-terminal tail have reduced packaging of RNA4, or RNA1.²¹⁻²³

In this study, we compared the stiffness of BMV virions that package different RNAs through atomic force microscopy (AFM) nanoindentation. We also examined the effects of altered capsid charge on BMV RNA packaging. In addition, we compared the empirical results to the theoretical optimal RNA lengths for the mutant virions under thermodynamic equilibrium. Our results suggest that the number of positive charges is an important factor, but not the sole factor, in controlling both the abundance and the species of the RNAs packaged.

Results

BMV exhibits heterogeneous stiffness among particles packaging different RNAs

The stiffness of virus, as measured by elastic constant, is the most important factor in evaluating the viral physical stability.^{24,25} Assuming the electrostatic interaction between the positively charged N-terminal tail of BMV CP and that RNA is the fundamental force in virus assembly and stabilization, we expect a narrow distribution of the stiffness of BMV as the RNAs packaged in each particle are 3 ± 0.2 kb. To compare the stiffness of the BMV particles that package different RNAs, we produced two forms of virus in the geranium tobacco plant (*N. benthamiana*) through agroinfiltration. The first used replication-competent BMV RNAs to yield the three subpopulations of particles that typified wild-type (WT) BMV (henceforth referred to as the three-particle virions). The second was made by

expressing the 1a and 2a replication proteins with RNA3 to produce virions that package only RNA3 and RNA4 (henceforth referred to as the R3/4 virions).²⁶

AFM imaging and nanoindentation were used, respectively, to measure the height over the surface support and the elastic constants of R3/4 particles and three-particle BMV. The two-dimensional histogram of the virions allowed testing for possible correlations between height and stiffness (Fig. 1a). The heights of both sets of particles were between 25 and 29 nm, the expected size for BMV, and ruled out the possibility of us analyzing incomplete or broken particles. However, a graphic comparison of elastic constant distributions regardless the particle diameter clearly showed the broader distribution of three-particle BMV with respect to that of R3/4 BMV (Fig. 1b), shifting toward stiffer particles. Such a significant difference in the distribution breadth of ~ 0.12 N/m far exceeded the estimated instrumental broadening (0.04 N/m) (Fig. S1) and argued against the difference being solely due to capsid stabilization by electrostatic interaction, which should be less than 0.003 N/m for BMV particles packaging different RNAs, if one assumes a direct proportionality between the number of charges and the force constant. These results suggested that force(s) in addition to electrostatics would affect the encapsidation of BMV RNAs.

Capsids with an additional 360 positive charges form virions of different densities

To better understand the role of the electrostatics in packaging of BMV RNAs, we increased the number of positive charges in the capsid and examined its effects on RNA packaging. One turn of an α -helix encompassing two positive-charged residues was inserted after residue 15 (Fig. 2a).

Using this design, we should minimally perturb the N-terminal tail that may form an α -helix upon interaction of BMV RNA.²⁷ Virions produced from this construct, named 1H, would increase the number of basic residues within the capsid by 360. For the 1H and the majority of the virus preparations, both the three-particle virion and the R3/4 particle were generated and characterized. However, the results for the R3/4 virions will be presented while the results for the three-particle virions are provided as supplemental information.

Strikingly, 1H R3/4 virions segregated into two bands of approximately equal abundance in a cesium chloride (CsCl) gradient (Fig. 2b), henceforth named 1H-L and 1H-H (for light and heavy, respectively). Northern blot analysis for BMV-specific RNAs among the total RNAs extracted from *N. benthamiana* confirmed that the insertion in 1H did not cause a defect in viral RNA replication (Fig. S2b). In negative-stained electron micrographs, 1H-L and 1H-H virions were indistinguishable from the WT in their diameters and shapes (Fig. S3a). The CPs from both 1H-L and 1H-H exhibited electrophoretic mobilities in SDS-PAGE that were commensurate with the presence of the four-residue insertion (Fig. S3b). In a differential scanning fluorometry (DSF) assay that examined the denaturation profile of the capsids as a function of increasing temperature,²⁸ both 1H-L and 1H-H had one dominant transition at 78 °C, the $T_{m,app}$ was comparable to the WT R3/4 virions (Fig. S4a and b). These results indicated that 1H-L and 1H-H shared the same capsid despite having distinct densities.

We postulate that the difference in density of the 1H-L and 1H-H particles is due to a change in the RNA packaged by the virions. The average amount of RNA in each virion population was examined by UV spectrophotometry.²⁹ The number of nucleotides (nt) present per virion was calculated from the

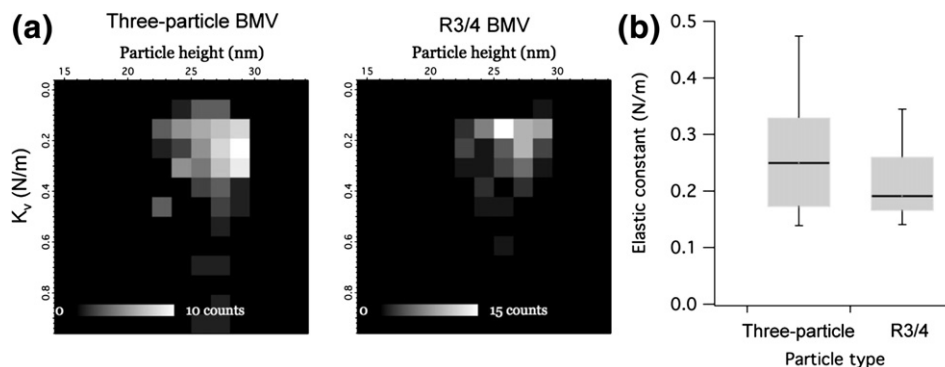


Fig. 1. BMV exhibits heterogeneous stiffness among particles packaging different RNAs. (a) Joined histograms of particle height and elastic constants measured from 108 three-particle virions and 80 R3/4 virions. The three-particle BMV contains a significant subpopulation of stiffer particles than R3/4 BMV and a few that are softer. (b) Boxplot representation of elastic constant distributions of three-particle virions and R3/4 BMV. The whiskers were set at 5% and 95% of the total range.

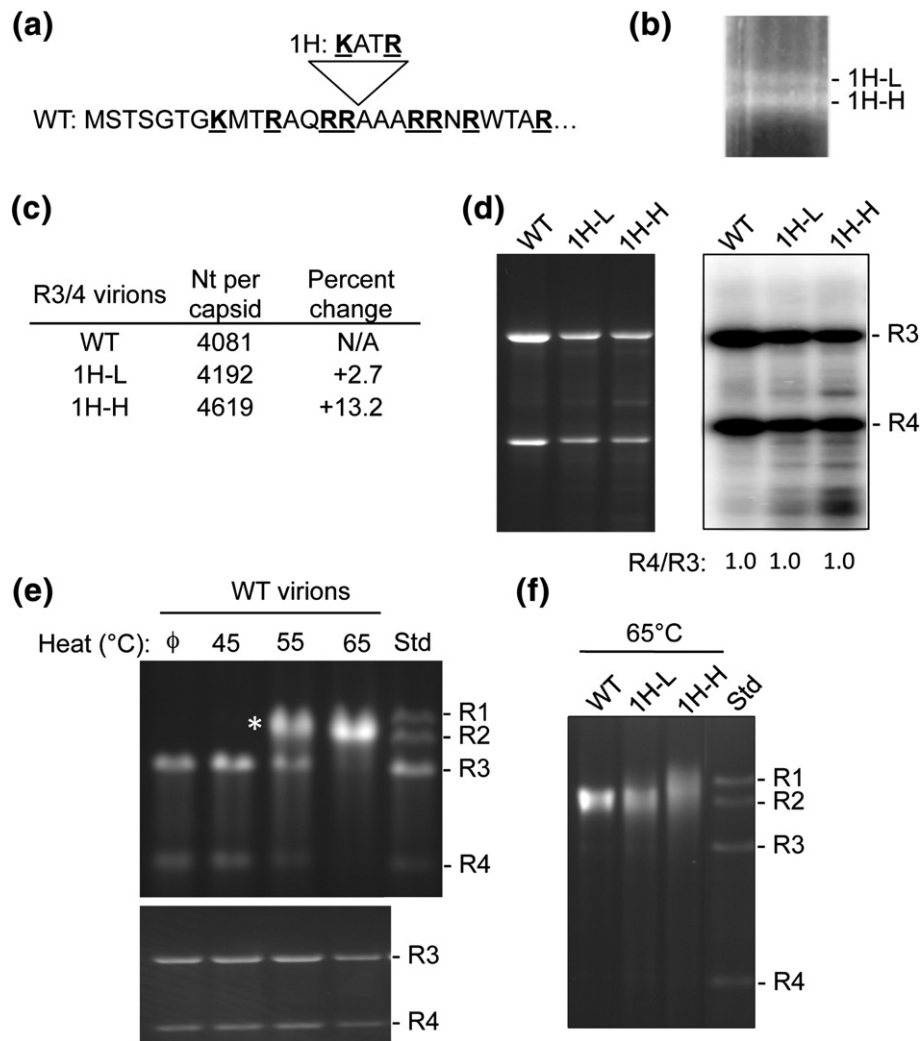


Fig. 2. Capsids with an additional 360 positive charges form two subpopulations of virions. (a) Sequence of the BMV capsid residues 1–26. All positively charged residues are in boldface and underlined. The triangle indicates the position of extra inserted residues in mutant 1H. (b) The R3/4 1H virions separated into two bands in a CsCl gradient. The names of the two preparations, 1H-L and 1H-H, are shown on the right of a photograph of the density gradient illuminated with white light. (c) Spectrophotometric analysis of the number of nucleotides per BMV capsid. The change in percentage was shown to allow normalization of the spectroscopic results. (d) RNAs encapsidated in the 1H-L and 1H-H virions. The left panel shows an ethidium-bromide-stained glyoxal gel. The right panel shows the RNA from the same gel blotted by the riboprobe specific to BMV 3' UTR. The quantification at the bottom shows the ratio of RNA4 to RNA3 in each sample. R3 and R4 are short for RNA3 and RNA4. (e) A demonstration that RNAs from the heated virions generate a noncovalent complex of RNA3 and RNA4. The upper panel shows a native TBE gel, and the lower panel shows a glyoxal gel. Both gels were stained with ethidium bromide. RNA from the WT three-particle virion serves as a molecular standard. The 3-kb RNA3–RNA4 complex is denoted by the asterisk. (f) Analysis of the RNA complex formed by heating of 1H-L and 1H-H virions. The image shows the ethidium-bromide-stained native TBE gel.

corrected absorption of virions at wavelengths of 260 and 280 nm using the theoretical extinction coefficients of the CP and RNA. The 1H-L virions contained nearly identical amounts of nucleotides as the WT, while the 1H-H virions packaged 13% more RNA (Fig. 2c). We noted that the calculations based on the spectroscopic readings for the WT R3/4 virions were ~1000 nt higher than what would be expected for the sum of RNA3 and RNA4, despite

adjustments for the buffer. This discrepancy is likely due to the theoretical extinction coefficients used. Nonetheless, the percentage changes unambiguously suggest that the 1H-H virions package more RNAs than a WT virion.³⁰

The RNAs extracted from 1H-L and 1H-H virions were subjected to gel electrophoresis followed by Northern blot analysis. Both 1H-L and 1H-H virions packaged approximately equal amounts of RNA3

and RNA4 as the WT, but 1H-H also contained a larger amount of truncated RNAs (Fig. 2d). The truncated RNAs were detected in Northern blots probed with a BMV-specific riboprobe that is complementary to the nearly identical 3' untranslated region (UTR) of all BMV RNAs, indicating that they were derived from BMV RNAs.³⁰

We sought to determine whether the truncated BMV RNAs are present in the same particles as RNA3 and RNA4 in the 1H-H virions. Krishna and Schneemann had shown that the co-packaged RNAs in Flock House Virus migrated as a complex in non-denaturing gels when the RNAs were extracted after a brief heating of the virions.³¹ We applied the same approach to the R3/4 WT BMV by heating it to 45, 55, or 65 °C for 10 min; extracted the RNAs; and subjected them to electrophoresis. In a non-denaturing gel, BMV virions heated to 55 or 65 °C unambiguously generated a ca 3-kb band, the size expected for an RNA3–RNA4 complex (Fig. 2e, top). In a denaturing gel, the RNA complex was resolved to the individual bands (Fig. 2e, bottom). This result demonstrates that heating BMV virions can provide information about whether RNAs are located within a virion.

Heating the 1H-H virions to 65 °C resulted in an RNA smear that migrated slightly slower than BMV RNA1 (3.2 kb), suggesting that the truncated RNAs were co-packaged in the 1H-H virions with BMV RNA3 and RNA4. In contrast, the 1H-L virions treated in the same manner resulted in a 3-kb RNA band (Fig. 2f). The results with the 1H virions suggest that 360 extra positive charges could lead to two equally probable outcomes of either packaging or not packaging additional truncated RNAs, a significant portion of which are recognized by BMV-specific probes.

Mutants with 720 additional positive charges packaged extra RNAs

The results with the 1H mutant prompted us to further examine the relationship between charges in the capsid and the amount and types of RNA encapsidated. Three mutants that increased the number of positive charges in the capsid by 720 were constructed, along with a control for each (Fig. 3a). Mutant 2H₁₅ contains an eight-residue insertion after residue 15 in the CP that could form two α -helical turns containing four positively charged residues. The control, 2HA₁₅, has an eight-residue insertion, but with alanines replacing the arginines and lysines. To examine whether the location of the insertions affected RNA encapsidation, we placed the insertions within 2H₁₅ and 2HA₁₅ after residue 7 to generate 2H₇ and 2HA₇. In addition, to keep the length of the N-terminal tail unchanged from the WT, we made 4R that had four arginines substituted for the uncharged residues at positions 9, 12, 16, and

17. Mutant 4S that has the same four residues replaced with serines serves as a control. All six mutants produced in the R3/4 particle form did not exert major adverse effect on viral RNA accumulation in *N. benthamiana*.

All six virions purified as a single band in CsCl density gradients and produced particles that were indistinguishable from the WT in negative-stained electron microscopy (EM) images (Fig. S3a and data not shown). The mutant CPs also exhibited the expected electrophoretic mobilities in SDS-PAGE (Fig. S3b). In Northern blots examining RNA accumulation, all of the mutants accumulated the expected BMV RNAs by 3dpi, suggesting that they are not significantly defective in RNA synthesis (Fig. S2). In the DSF assay, all mutants exhibited predominant denaturation at 78 °C, the same as the WT virions, and varying extents of a minor peak at 68 °C (Fig. S4a and b). The presence of a minor peak with a lower-than-normal denaturing temperature is likely associated with aberrant packaging of RNAs in these mutant virions.

Spectroscopic analysis showed that 2H₇, 2H₁₅, and 4R had, respectively, an average of 19%, 21%, and 9% higher amounts of nucleotides per capsid relative to the WT virions (Fig. 3b). In contrast, the RNAs within the three corresponding control mutants were within 3% of that of the WT virion. These results show that increasing positive charges on the capsid increased the amount of RNA encapsidated.

Gel electrophoresis and Northern blot analyses showed that 2H₇ and 4R have a 1:1 ratio of RNA3 to RNA4 (Fig. 3c). However, these virions packaged a higher abundance of truncated RNAs that annealed to the BMV-specific probe when compared to the corresponding controls or WT BMV. Mutant 2H₁₅ contained comparable amounts of the truncated RNAs as 2HA₁₅ but had a close to 2:1 ratio of RNA4 to RNA3, suggesting that an extra RNA4 (0.9 kb) was packaged in the 2H₁₅ virion. Ethidium-bromide-stained RNAs from the 4R virion also revealed a band slightly larger than that of RNA1, which is likely of cellular origin because it did not hybridize to the BMV-specific RNA probe (Fig. 3c, see asterisk). 4R also contained more truncated RNAs that were not recognized by a riboprobe specific to the BMV 3' UTR. This may be due to the truncated RNAs lacking the 3' UTR or to the 4R packaging a significant proportion of cellular RNAs. RNAs from heated virions confirmed that the truncated RNAs or the additional RNA4 were co-packaged with RNA3 and RNA4 (Fig. 4d), where 2H₁₅, 2H₇, and 4R all exhibited a band or smear larger than the 3-kb RNA3–RNA4 complex.

Curiously, mutants 4S and 2HA₁₅, which did not alter the number of positive charges within the capsid, packaged a reduced amount of RNA4 relative to RNA3 (Fig. 3c). Since this defect was

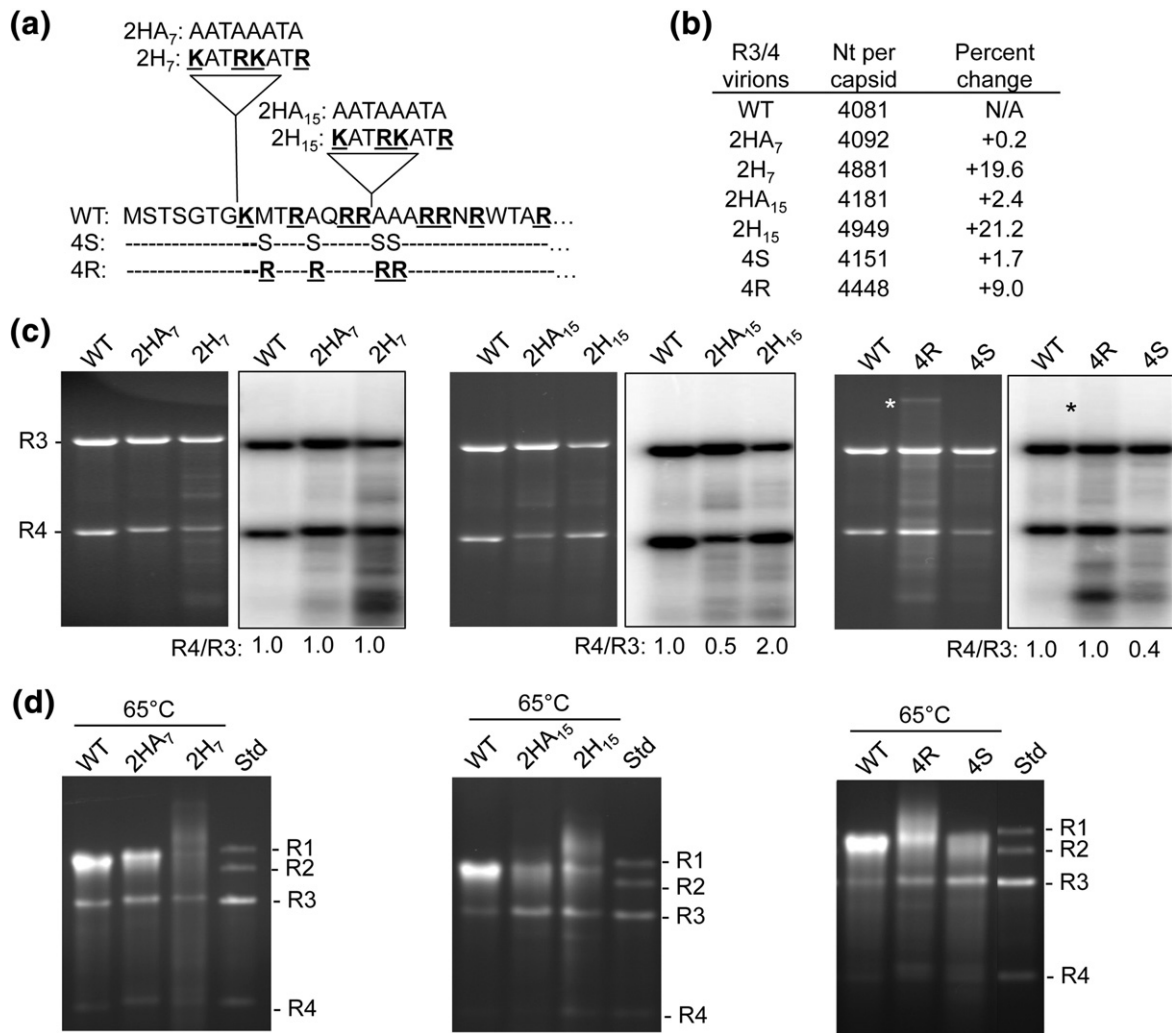


Fig. 3. Analysis of mutant virions that had an additional 720 positive charges. (a) Schematic of the N-terminal sequences of the six mutants characterized. The basic residues are in boldface and underlined. The triangles balanced on lines denote the locations of the insertions. Only the substituted residues were typed out in 4R and 4S. (b) Spectroscopic analysis of the number of nucleotides per virion. (c) RNAs from the six purified BMV mutants. For each set of mutants, the left image was ethidium-bromide-stained glyoxal gel, and the right was Northern blot images that detect BMV RNAs. The ratios of the full-length RNA4 to RNA3 are shown at the bottom of the Northern blot image. The white and black asterisks denote an unknown cellular RNA in the 4R virion. (d) RNA complex formed by heating the virions. The RNAs were separated on a TBE gel and stained by ethidium bromide.

not observed in 2H₇, and the substitutions in the 4S overlapped with the site of insertion in 2HA₁₅, it is likely that this location in the CP N-terminal tail affects RNA4 packaging. In an analysis of RNAs from heated virions, the control virions did not exhibit RNA with electrophoretic mobility higher than 3 kb (Fig. 4d). However, the ca 3-kb band from the 4S and 2HA₁₅ appeared more heterogeneous compared to that of the WT, suggesting the co-packaging of truncated RNAs with RNA3 in some of the particles. This scenario can also account for the observation by spectroscopic analysis that 4S and 2HA₁₅ virions contain the same overall number of nucleotides as do WT virions.

All six mutants and 1H described above were also produced in the three-particle form. The results were highly consistent with the results for the viruses in R3/4 form analyzed in this work. A summary of the results from the three-particle BMV is in Fig. S2a and Fig. S5.

We also constructed mutants 3H and 4H that had, respectively, either three or four repeats of the KATR motif inserted after residue 15 of the CP. In *N. benthamiana*, 3H and 4H produced the expected amount and ratios of the BMV RNAs, suggesting that their RNA replication was unaffected by the mutations (data not show). However, the yield of 3H and 4H particles was at less than 5% of the WT,

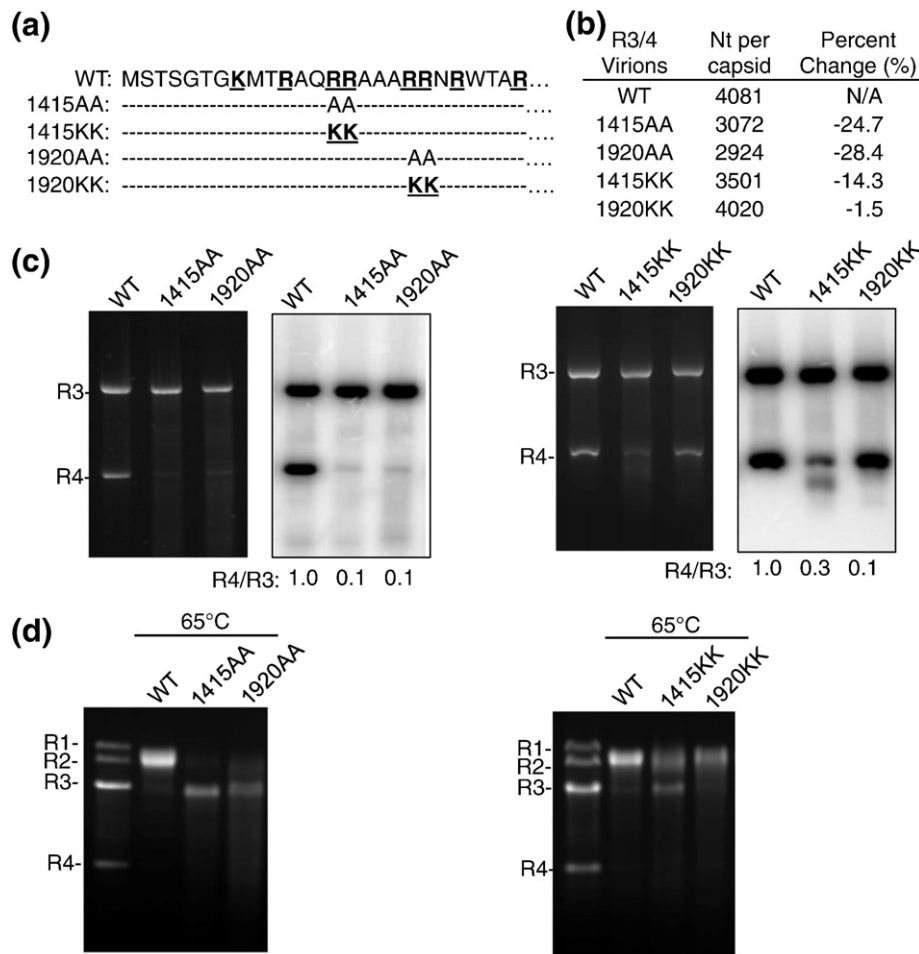


Fig. 4. Analysis of mutant virions that had 360 fewer positive charges. (a) Schematic of the N-terminal sequences of the four mutants characterized. The basic residues are in boldface and underlined. Only the substituted residues were typed out in the mutants. (b) Spectroscopic analysis of the number of nucleotides per virion. (c) RNAs from the four purified BMV mutants. For each set of mutants, the left image was ethidium-bromide-stained glyoxal gel, and the right was Northern blot image that detects BMV RNAs. The ratios of the full-length RNA4 to RNA3 are shown at the bottom of the Northern blot image. (d) RNA complex formed by heating the virions. The RNAs were separated on a TBE gel and stained by ethidium bromide.

thus precluding further analysis of the virions. These results suggest that there is an upper limit of either residues and/or positively charged residues that can be inserted in the N-terminal tail before affecting virion integrity.

Mutants with 360 fewer positive charges packaged less amounts of RNAs

To further examine the relationship between electrostatic charge and RNA encapsidation, we made mutants 1415AA and 1920AA that had pairs of arginines substituted with alanines (Fig. 4a). The number of positive charges thereby decreased by 360 within each mutant capsid. Substitutions of the pairs of arginine to lysines at the same positions generated mutants 1415KK and 1920KK.

All four mutant virions purified as single bands in CsCl density gradients. In negative-stained electron micrographs, the virions were indistinguishable from the WT. Furthermore, CP exhibited the expected electrophoretic mobilities in SDS-PAGE (Fig. S3a and b). Both mutants 1415KK and 1920KK had $T_{m,app}$ of 78 °C in the DSF assay, comparable to that of the WT virions (Fig. S4a and b). Mutants 1415AA and 1920AA had a predominant denaturation peak with a $T_{m,app}$ of 65.5 °C and a minor one with a $T_{m,app}$ of 72 °C (Fig. S4a and b). Spectroscopic analysis revealed that the 1415AA and 1920AA contained, respectively, 25% and 28% less RNA relative to the WT. Interestingly, although 1415KK was unaffected in its morphology and in its thermodenaturation, it packaged 14% less RNA (Fig. 4b). Mutant 1920KK had RNA levels that are comparable to that of WT BMV.

The RNAs encapsulated in the four mutants were examined by gel electrophoresis and Northern blots (Fig. 4c). Mutant 1920KK had the expected complement of RNA3 and RNA4. Mutant 1920AA, 1415AA, and (to a lesser extent) 1415KK had reduced amounts of RNA4 compared to RNA3. 1415KK also contained a truncated RNA that was slightly smaller than RNA4 and was detectable by the BMV-specific riboprobe. Heat-treated virions revealed that, where present, RNA4 and possibly the truncated BMV RNA in 1415KK were packaged together with RNA3 (Fig. 4d). These results reinforce our earlier observations that the amount of RNA packaged could be affected by both the number and the location of positive charges in the BMV CP.

Theoretical calculation of the optimal amounts of RNAs in BMV mutants

We sought to determine whether the variation observed in RNA packaging by BMV mutants could be understood by thermodynamic models that incorporate electrostatics, RNA and N-terminal tail conformational entropy, and excluded volume effects. Specifically, we calculated, as a function of the number of charges in the peptide tails, the free energy of a theoretical virion due to confinement of the RNA and interactions among the RNA, the N-terminal tail, and ions. We employed the method of Scheutjens and Fler in which electrostatics are treated with the nonlinear Poisson–Boltzmann equation with the spatial distributions of RNA, N-terminal tails, ions, and water solved numerically with a self-consistent field approximation on a lattice.^{32–35} With the exception of the N-terminal tails, the assembled capsid was treated as a sphere, which is impermeable to N-terminal tails and RNA but is permeable to water and small ions. The N-terminal tails were explicitly represented as flexible polymers, whose segments carry electric charge dictated by the amino acid sequence. The RNA molecule was treated as a linear polyelectrolyte, and we imposed spherical symmetry on the complex. We assumed that the capsid geometry is invariant to packaged RNA length, and hence, the free energy due to attractive protein subunit–subunit binding interactions is a constant in our model. However, the free energy due to interactions among the N-terminal tails depends on protein sequence and RNA length in our model.

We calculated encapsidation free energies under three sets of assumptions to determine their effect on the results. In the first case, we considered one charge per RNA nucleotide, and we modeled the virion in the presence of 130 mM monovalent salt. In the second case, we considered one charge per RNA nucleotide, and we modeled the virion in the

presence of a negatively charged macromolecule to which the capsid is impermeable. As noted by Ting *et al.*, negatively charged macromolecules are ubiquitous within the cell, and if they are selectively excluded from the interior of the capsid, their presence establishes a Donnan equilibrium and thus a potential gradient across the capsid shell.³⁶ Specifically, we considered the capsid in the presence of small permeable cations and ions with respective concentrations of 130 and 30 mM and impermeable negative charge (on the macromolecules) of 100 mM. In the third case, we considered the same ion concentrations as for case 2, but we assumed that the charge density of the RNA molecule is limited by counterion condensation.³⁷ Under that approximation, the bare charge on the RNA molecule (1 charge/0.5 nm in our model) is reduced to a linear density of 1 charge/ l_B with l_B as the Bjerrum length. Counterion condensation cannot be investigated explicitly within our model due to the continuum approximation and imposition of spherical symmetry.

The free energy is shown as a function of RNA length for the third case (100 mM impermeable macromolecule and counterion condensation) for the WT capsid and the five mutants, 1415AA, 1920AA, 1H, 2H₁₅, and 4R (Fig. 5a and b). 2H₇ was not examined since it has the same inserted sequence as 2H₁₅. We find that, in all cases, the free energy varies nonmonotonically with RNA length, with a minimum that corresponds to the length of RNA around which the assembled capsid is most thermodynamically stable. Under these assumptions, the optimal RNA length is approximately 2400 nt for the WT capsid. This value is comparable to but less than 2900–3000 nt typically packaged by BMV capsids. In addition, the volume fraction of RNA segments (Fig. 5d) is nonmonotonic with distance from the capsid surface, which is qualitatively consistent with the experimental observation.

The predicted encapsidation free energies are qualitatively similar for cases 1 and 2 described above. The predicted optimal packaged length is shown as a function of total capsid charge for each case and each sequence in Fig. 5c. In all cases, the optimal length increases approximately linearly with the amount of capsid charge, with small variations due to changes in capsid N-terminal tails. For case 1, the amount of charge on the packaged RNA corresponds to nearly the amount of positive charge on the capsid. For case 2, the amount of packaged RNA increases by approximately 300 nt for each sequence in comparison to case 1. The magnitude of this increase can be understood by simple arguments described in the supporting information. Finally, for case 3, the ratio of RNA charge to capsid charge is approximately the same as for case 2 after accounting for charge renormalization due to counterion condensation.

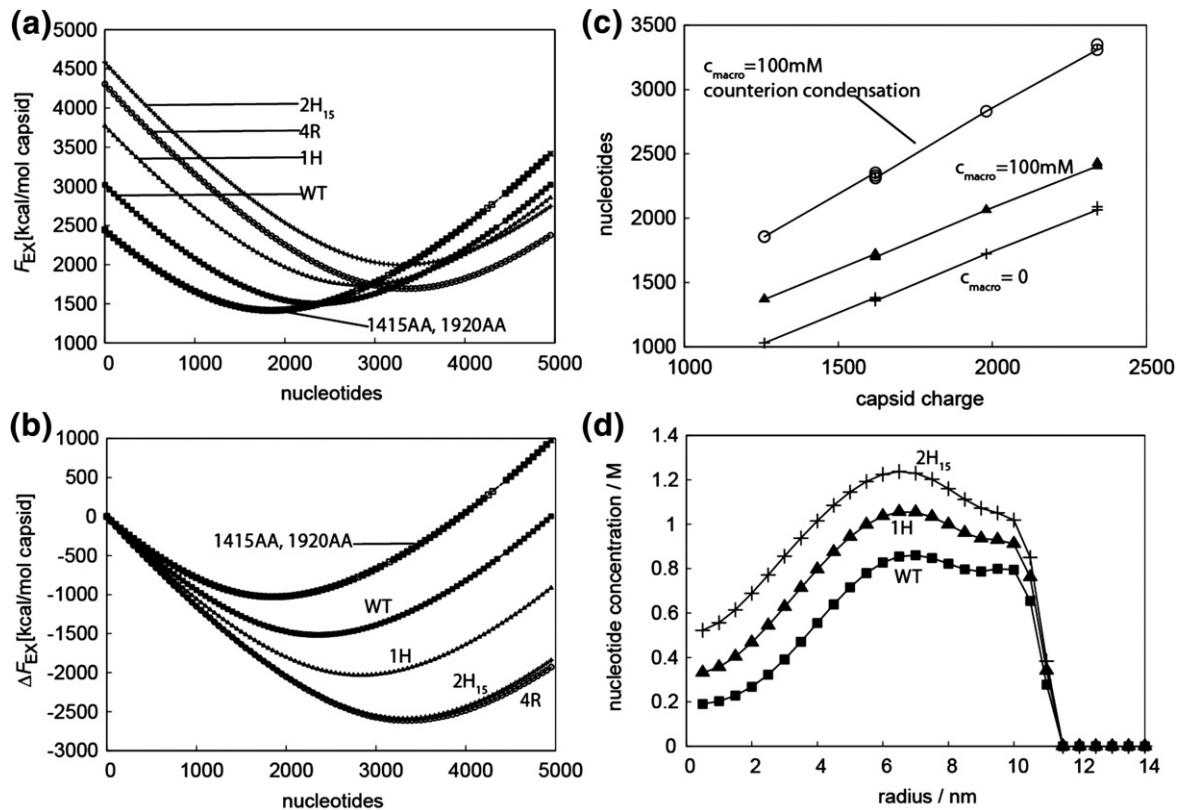


Fig. 5. Theoretical prediction of the relationship between the capsid charge and RNA length. (a) The free energy is shown as a function of RNA length for the WT and the five mutants in which additional charges were inserted, deleted, or substituted into the N-terminal tails. Results are shown for case 3 described in the text, where there is a concentration ($c_{macro}=100$ mM) of anionic macromolecules excluded from the capsid, and the RNA charge is renormalized to 0.7 charges per nucleotide by counterion condensation. (b) The same data are plotted to identify the driving force for RNA encapsidation, ΔF_{EX} , which is the free-energy difference between the capsid with the RNA molecules enclosed and the empty capsid. (c) The predicted optimal length of packaged RNA is shown as a function of capsid charge for all peptide sequences (1420AA, 1920AA, WT, 1H, 4R and 2H₁₅) for each of the three sets of assumptions considered in the text: $c_{macro}=0$, $c_{macro}=100mM$, and $c_{macro}=100$ mM with charge renormalization due to counterion condensation. (d) Theoretical prediction of the RNA density as a function of radial position within the capsid.

The predicted optimal length increases by approximately 20% for the 1H mutant and 40% for the 2H₁₅ mutant. These changes in optimal RNA length with respect to the WT capsid are about twofold larger than the experimental observations, which are an approximately 13% increase for 1H and a 21% increase for 2H₁₅. The predicted decreases for the 1415AA and 1920AA mutants of -21% are close to the experimental observations, which are -25% and -28% . Furthermore, the minimum free-energy RNA length for the control sequence, 2HA₁₅, is identical with that for WT (data not shown), consistent with experimental observations. In contrast, the predicted increase for the 4R sequence is approximately equal to that of 2H₁₅ (42%), while in experiments, the 4R sequence packaged only 9% more RNAs. In all cases, the amount of charge on the packaged RNA increases linearly with the amount of positive charge on the capsid N-terminal tails, indicating that

electrostatics is a dominant effect in these calculations. These results suggest, however, that additional effects accounting for the sequence of charges in the N-terminal tails would be required for a theoretical model to agree with the full set of experimental results.

BMV that packages extra RNAs did not change the capsid structure

Our observations with the mutant virions suggest that the BMV can sterically accommodate additional RNAs. To better analyze the densities of the RNAs within the virion, we selected 2H₁₅ for further characterization. Cryo-electron micrographs of virions frozen in vitreous were used to reconstruct a three-dimensional density map by applying icosahedral symmetry. The map was refined to a resolution of 5 Å with well-defined pentameric

and hexameric capsomeres (Fig. S6a). The lateral section of the density showed an outer diameter of approximately 286 Å (Fig. 6a), consistent with averaging diameter of the WT.¹⁸ The inner radius of the 2H₁₅ is about 96.4 Å, which is about 6 Å larger than that of the WT. This is probably because larger segment on the N-terminal tails in 2H₁₅ is disordered. The class average of the 2H₁₅ particles exhibited that the center of virion was filled with additional density when compared to the WT, which is likely to be associated with the additional RNA4 (Fig. S6b). Results from small-angle X-ray diffraction studies also revealed that 2H₁₅ had a thicker layer of the encapsidated RNAs and capsid (Fig. S7).

The cryo-EM map fitted well with the 3.3-Å crystal structure of the WT BMV and enabled the tracing of the C^α backbone for A, B, and C capsid subunits (Fig. 6b and c). With the guide of the crystal

structure, we assigned the model of pentameric subunit A to residues 41–189 and the model of hexameric subunits B and C to residues 25–189 and 25–189, respectively [the numbers do not include the inserted residues to allow comparison to the X-ray structure of the WT BMV (Fig. 6b)].¹⁸ The densities of the remainder of the N-terminal tail were not resolved in the models. These observations agree with the assumption that the N-terminal tails are complexed with the RNA. A further superposition of the backbone of the 2H₁₅ trimer on that of the crystal structure of the WT revealed that the globular domains of the capsid subunits were largely unchanged relative to the comparable structures from the crystal structure of BMV (rmsd < 2 Å), while the N-terminal and C-terminal tails displayed the most significant variations. We did not find any substantial change in diameter of the center pore of the quasi-3-fold axes at current

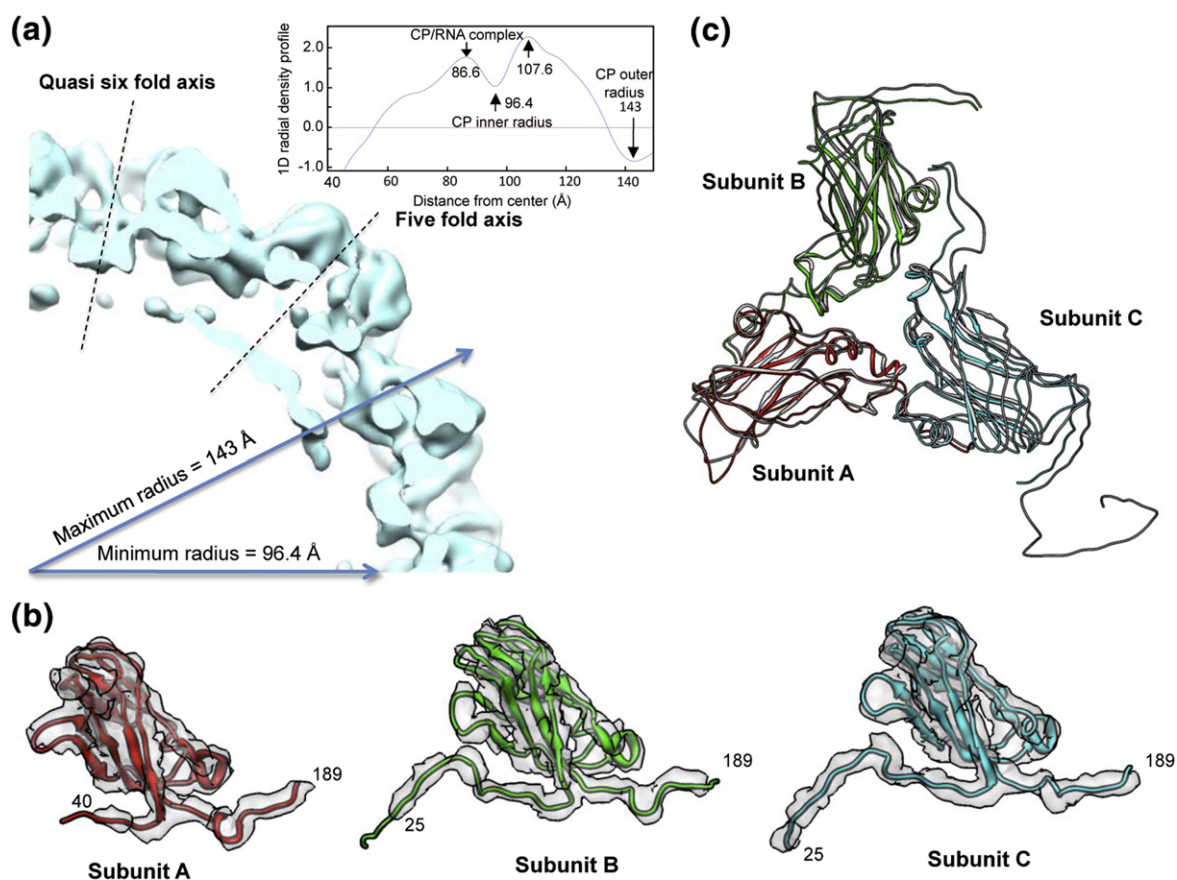


Fig. 6. The cryo-EM structure of 2H₁₅ BMV. (a) A quarter cross-section through the 2H₁₅ capsid. The plane of the cross-section includes the 5-fold and quasi-6-fold axes that cut in the centers of the pentameric and hexameric capsomeres. The inset is the averaged one-dimensional radial density profile of the 2H₁₅. (b) Models for subunits. Subunit A was shown in red (residues 40–189), subunit B was shown in green (residues 25–189), and subunit C was shown in blue (residues 25–189), respectively. The inserted eight residues were not included in the numbering of the residues. The density map is shown in transparent gray and displayed at slightly higher isosurface threshold. (c) Backbone representation of the 2H₁₅ trimer superimposed on the corresponding part from the crystal model of the WT BMV (Protein Data Bank ID 1js9), which is presented as a dark-gray overlay.

resolution level. Overall, our cryo-EM model of 2H₁₅ indicated that the inserted residues in the N-terminal tail and the packaging of additional RNA4 did not affect the folding or the arrangement of the capsid subunits. However, in this reconstruction of 2H₁₅, we were unable to visualize the density of RNA.

Discussion

A common motif in positive-strand RNA viruses is that their CPs contain highly positively charged tails that extend into the interior of the capsid. Presumably, the electrostatic attraction between the RNA and positively charged tails guides virion assembly and maintains virus integrity. This work seeks to examine whether electrostatic interaction is sufficient to explain BMV RNA packaging using a combination of mechanical and molecular genetic approaches. The results are also compared to models of RNA encapsidation based on thermodynamic modeling of capsid–RNA interaction. This comparison leads to the conclusion that, while electrostatics can guide RNA packaging, changes in the amount and the species of RNA encapsidated varied for different mutations with the same capsid charges and that there are sequence-specific effects on RNA recognition related to the location of positively charged residues in the BMV CP.

One prediction of an electrostatic model is that mechanical resistance to pressure will correlate with the length of the RNA encapsidated. However, the three-particle BMV, a natural mix of the three subpopulations of BMV, exhibits a rather wide range of elastic constants in the nanoindentation experiment as opposed to a narrow one observed for R3/4 BMV. The heterogeneity in elastic property of the three-particle BMV could not solely come from the different orientation of the particles or the packaging of non-genomic material from host cells, which should also be valid for R3/4 particles as well. Most importantly, the difference in elastic constant due to capsid stabilization by the electrostatic interaction of different RNAs and CP could not account for the observed drastic variation, suggesting that there are forces other than the electrostatics that also contribute substantially to the RNA–CP interaction within capsid.

Molecular genetic analysis of BMV mutants with altered N-terminal sequences further supported that the role for the positively charged N-terminal tail is more complex than a simple bearer of electrostatic charges. Twelve purified mutant BMV R3/4 virions with increased, decreased, or unchanged numbers of capsid charges were analyzed for the amount of encapsidated RNA by spectroscopic analysis. The correlation between the total number of positive changes on the N-terminal tails and the calculated

number of RNA nucleotides was plotted in Fig. 7. In general, the RNA length was affected by the capsid charge. For example, all four BMV mutants with increased capsid charges (1H, 2H₇, 2H₁₅, and 4R) packaged more RNA than the WT. However, several features were notable. First, the sum of the RNAs packaged by virions that have the same number of charges could differ by ca 400 nt (e.g., compare the results for the 4R and 2H₇ or those for the WT and 1415KK). This distribution is underscored by the results from the 1H virions, which formed two distinct particles between which the amount of RNA varied by an estimated 400 nt. Second, we note that the 4R, whose charge is increased by substitution instead of insertion, had less encapsidated RNAs than did 2H₇ and 2H₁₅, which had the same number of capsid charges. This may arise because 2H₇ and 2H₁₅ have longer extended N-termini with more evenly dispersed positive charges. Third, a nonlinear correlation for the 12 mutants was observed. The incorporation of additional RNA in response to increased numbers of positive charges in the capsid was lower than the loss of RNA that resulted when the number of capsid charges was decreased. While this effect might be attributed to excluded volume, its magnitude was not captured in any of the theoretical models that we considered, all of which do account for excluded volume.

We note that there appear to be other constraints on RNA encapsidation. For example, BMV virions with insertions of three or four repeats of the residues inserted in 1H were unable to efficiently form virions, although the accumulation of viral RNAs in *N. benthamiana* leaves was largely

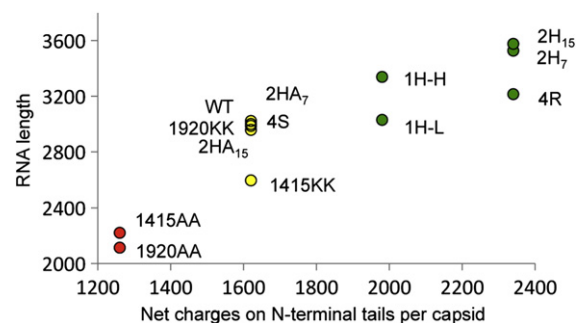


Fig. 7. The correlation between the charges in the N-terminal tails and the amount of RNA packaged. The net charges on N-terminal tails per capsid were calculated by multiplying the number of positive charges within the first 49 residues by 180, the number of CP subunits per $T=3$ capsid. The RNA length for each mutant was estimated by applying the corresponding percentage change from Figs. 2b, 3b, and 4b to 2987 nt, the expected size of RNA3 and RNA4 in WT. The mutants with increase, unchanged, and decreased capsid charges were colored green, yellow, and red, respectively.

unaffected by the mutations. Based on the currently available results, however, we cannot determine whether the defect in virion formation is due to steric issues that arose from the increased length of the capsid tails or from the increased amount of RNA that is associated with the increased capsid charges.

A position-dependent effect of residues in the N-terminal tail on RNA encapsidation was also suggested by our results. Mutant 1920KK is competent at packaging RNA4 while mutant 1415KK packaged reduced amounts of RNA4 and a truncated BMV RNA. It is possible that a bidentate contact involving arginines at residues 14 and 15 is required for proper interaction between the capsid and RNA4. Other mutants with unchanged capsid charges but sequence alterations close to residue 14 and 15, namely, 4S and 2HA₁₅, also reduced RNA4 encapsidation. A similar specific interaction between the N-terminal tail and RNA4 was also observed with mutant 2H₁₅, which have four positively charged residues inserted after position 15. 2H₁₅ packaged two copies of RNA4 per RNA3 in the virion. Choi and Rao have also observed that the N-terminal tail of BMV is involved in the selective encapsidation of BMV RNA4.²³

The mutations we analyzed had dramatic variations in the packaging of RNA4 in contrast to the minimal effect on the encapsidation of RNA3. These results suggest that the RNAs play a role beyond that of a simple polyelectrolyte in the assembly of virions. The BMV capsid may have evolved mechanisms to differentially recognize the four BMV RNAs. A greater flexibility for the encapsidation for RNA4 is logical for BMV since it could be produced from a copy of RNA3 and is not needed to initiate RNA synthesis.^{16,17,26} By contrast, RNA2 is preferentially encapsidated by several mutants when the three-particle virions were examined (Fig. S5). We have also demonstrated previously that the RNA2-containing virion is the most resistant to degradation.²⁸ Perhaps it is more important for BMV to protect RNA2 since it encodes the polymerase needed for propagation of the virus.

Although our experimental data argue against the existence of a fixed correlation between the genome size and capsid charge, the charge ratio and optimal RNA length predicted by thermodynamic models incorporating electrostatics still provide valuable guidance. We need to bear in mind that these charge ratios and optimal RNA lengths vary when different models are considered. For instance, Hu *et al.* assumed that RNA wraps around peptide tails and found that the free energy is minimized when the RNA contour length is approximately equal to the total contour length of peptide tails, resulting in an optimal charge ratio of 2;¹¹ Belyi and Muthukumar treated the RNA as a polyelectrolyte and the peptide tails as oppositely charged brushes and used the

ground-state dominance approximation to predict an optimal charge ratio of 1:1.¹⁰ Upon renormalizing the charge on RNA and the peptide tails according to counterion condensation theory, the predicted ratio of packaged nucleotides to capsid charges was 1.6:1. Calculations that place the capsid charge entirely at the surface have predicted charge ratios of <1, 1, and 2.^{38–40} Ting *et al.* found that the optimal ratio of packaged charge to capsid charge was sensitive to details such as the charge density on the peptides but was less than 1:1.³⁶ They showed, however, that the Donnan potential due to anionic macromolecules that are selectively excluded from the capsid could increase this ratio to larger than 1:1. In this work, we applied a continuum version of the self-consistent field theory with explicit representation of the N-terminal tail as flexible polymers with the sequences of charge considered in our experiments. We found that, for all sets of approximations that we considered, the optimal amount of packaged charge was relatively insensitive to the location of charges within the peptide sequence and increased linearly with the total positive charge on the capsid. Upon incorporating the Donnan potential and charge renormalization due to counterion condensation, our theory predicted an optimal charge ratio of approximately 2.

Based on our experimental data describing the amount of RNA packaged by each mutant, the amount of packaged RNA tends to increase with increasing positive charges on the capsid, and the ratio of packaged nucleotides to positive charge on the capsid ranges from 0.33 to 2.22. These results are in line with the predictions of the various theoretical models listed above. However, the fact that the amounts of RNA packaged vary nonlinearly with capsid charge suggests that the sensitivity of RNA recognition by the peptide sequence is not fully captured by any of the theoretical approaches.

There are several factors that could account for the imperfect agreement between theory and experimental observations. First, some discretion is required when counting the number of capsid charges when comparing the predicted charge ratio to that of real viruses. The theoretical calculation is based on a model in which the peptide arms protrude inward from capsid and interact with RNA. However, defining the part of the peptide arm that extends inward and hence interacts with the RNA is relatively arbitrary, which directly affects the counting of charges. Second, the disposition of charges across the secondary or tertiary structure of the peptide tails is not well captured in the theoretical models, including ours. This simplification could contribute to the fact that the theoretical calculations predict that the RNA packaging free energy depends strongly on the net charge of the peptide tails but is relatively insensitive to the location of the charges within this sequence. This consideration

may be especially relevant for the 4R mutant, where the mutated residues are more clustered than in other mutants or the WT. Furthermore, the theory does not include sequence-specific peptide structures that may affect the recognition of particular RNA sequences and/or structures (see below). Other approximations in the theory, such as the assumption of spherical symmetry within the virion or the neglect of ion-ion correlations, could also contribute to the discrepancy between theory and experimental results. Finally, it is possible that, in addition to the thermodynamic considerations described here, kinetics or other nonequilibrium effects also control the amount of RNA packaged in the experiments.^{41,42} The capsid assembly is an out-of-equilibrium process, so that equilibrium is only reached, if ever, once the final product has assembled.

There are several caveats in the experimental data that we want to highlight. First, the genome size derived from spectroscopic measurement only reflects the average length of RNA for a population of virions. Although all the mutants in our study, except 1H, purified as a single band in the CsCl gradient, the RNA content within each mutant is not necessarily homogenous given the gradient resolution, especially when truncated RNAs of various lengths are packaged. This variation is also evident in the profiles of RNAs from the heat-treated virions, where the co-packaged RNAs migrated in a smear. In other occasions, a minor population of virus that is probably too diluted to form a visible band in CsCl gradient may be included in the purified virion. For instance, a small fraction (less than 10%) of the 1415AA and 1920AA virions packaged RNA4 together with RNA3 while the majority packaged RNA3 alone.

Second, it is important to note that, with the exception of 4R, a BMV-specific riboprobe complementing to 3' UTR recognized most of the truncated RNAs encapsulated by the mutants. Although it is unclear how these truncated RNAs are produced, the prevalence of 3' UTR suggests that this region may contain a sequence or assume a conformation that can be specifically recognized by BMV CP. *In vitro* RNA encapsidation experiments have revealed that the BMV capsid preferentially encapsidates RNAs containing a tRNA-like structure, which is found in the BMV 3' UTR.⁴³ It is also possible that the tRNA-like structure, a motif required for BMV replication,⁴⁴ contributes to the specific recognition by the CP through coordinating RNA replication and encapsidation.⁴⁵⁻⁴⁷

Lastly, the BMV mutants in our study accommodated RNAs that ranged from 2000 to 4000 nt in sizes without any obvious change in the capsid sizes or conformations. The cryo-EM reconstruction of the 2H₁₅ virion yielded models comparable to the WT virions. These results indicate that BMV capsid is a highly adaptable cage for RNA. However, the

volume within the capsid is finite, as seen by the low yields of the 3H and 4H mutant virions in our studies. A self-assembly study using the CP of Cowpea Chlorotic Mottle Virus, a close relative of BMV, and RNA of different size found that when RNA exceeded 4500 nt, it was assembled into two or more connecting capsids.⁴⁸

In summary, comparison of theoretical predictions with our experimental observation virions with varying amounts of positive charge indicates that electrostatics plays an important role in the amount of RNA that is packaged. However, our results also indicate that the specific amino acid sequence of the capsid N-terminal tails contributes to RNA packaging in a manner that cannot be entirely described by models that treat them as simple polyelectrolytes, possibly due to secondary or tertiary structure of the peptide and its corresponding ability to recognize particular structural motifs found within the viral genome.

Materials and Methods

Plasmid constructs and agroinfiltration

An *Agrobacterium*-mediated gene delivery system was used to express the three BMV genomic RNAs and the individual BMV replication proteins.²⁶ cDNA for BMV RNA3 expressing the mutant CP was made by PCR-mediated site-directed mutagenesis with appropriate primers that will be available upon request. The sequences of the cDNAs were confirmed to have no other mutations. *N. benthamiana* plants used to produce virions were grown in an environmentally regulated growth chamber kept at a constant 25 °C, 70–75% humidity, and a 16/ 8-h light/dark cycle.

Virus purification

The purification of BMV was modified from the method of Bujarski.⁴⁹ Briefly, *N. benthamiana* leaves were homogenized in virus buffer I [250 mM NaOAc and 10 mM MgCl₂ (pH 4.5)], and the supernatants were clarified by 10% chloroform. The supernatant was then layered on a 10% sucrose cushion prepared in virus buffer I and centrifuged for 3 h at 28,000 rpm using a Beckman SW32 rotor to pellet the virus. The pellets were dissolved in virus buffer II [50 mM NaOAc and 10 mM MgCl₂ (pH 5.2)] with a 50% CsCl (w/v) and banded by centrifuging for 20 h at 65,000 rpm using a Beckman TLA110 rotor. The band of virions, which appeared opalescent white, was collected and dialyzed with three changes in virus buffer II and stored at –80 °C until use.

Nanoindentation by atomic force microscope

To mount the sample, we deposited a droplet of 50 µl virus on a freshly cleaved highly ordered pyrolytic graphite disk, mounted on the peizo holder, and incubated

for 10 min before probing. A droplet holder was used to minimize evaporation and maintain a stable buffer system.

AFM imaging and mechanical indentation measurements were recorded using a Cypher AFM (Asylum Research) instrument in liquid, at room temperature. Silicon cantilevers (BioLeverMini; Olympus) with a tip radius of ~ 9 nm and a spring constant of ~ 0.1 N/m were used. The cantilever spring constant was calibrated using the thermal oscillation method.⁵⁰ AFM images were obtained in the alternative contact mode using low set point amplitudes (<10 mV). This way, damage to the virus by the tip was minimized. Samples were first imaged at low resolution (500 nm \times 500 nm) to identify a single virus. Higher-resolution topographic images (80 nm \times 80 nm) were then recorded, followed by force mapping in contact mode. Force maps consisted of arrays of 6×6 force–distance curves on 80 nm \times 80 nm areas, using a maximum applied force of 700 pN and a loading rate of 1 μ m/s. After force mapping, a topographic image was always collected to check the integrity of the particle. When notable morphological differences were observed, the data were rejected.

To estimate the instrumental response, we collected a series of force curves at the same position (middle of the virus particle) from the same particle, until a large deviation of spring constant was observed. The variance in the elastic constants obtained this way characterizes the accuracy of measurements in the absence of heterogeneous broadening. Images and force maps were analyzed using Igor Pro 6.12 software (WaveMetrics).

Spring constant calculation

The compression of the virus by the tip can be represented as two springs connected in series represented by the virus and the cantilever.²⁴ The elastic constant of a virus is related to that of the cantilever by: $K_v = \frac{K_c K_{\text{eff}}}{K_c + K_{\text{eff}}}$ where K_{eff} is the effective spring constant obtained by a least-square fit of the linear region in the force–distance curve. K_c is the cantilever spring constant.

RNA extraction and Northern blot

Virion RNAs were extracted from the CsCl gradient purified virions by a 1:1 mixture of phenol/chloroform after treating virions with $2 \times$ lysis buffer [200 mM Tris, 2 mM ethylenediaminetetraacetic acid, and 2% SDS (pH 8.5)]. Northern blots were performed following the separation of glyoxal-treated RNA on a 1.2% agarose-gel containing BPTE buffer (10 mM PIPES, 30 mM Bis-Tris and 1 mM EDTA, pH 6.5).⁵¹ An *in vitro*-transcribed 3' UTR riboprobe was used to detect the positive-strand BMV RNAs.³⁰ Bands in the blots were quantified using ImageQuant software.

Analysis of encapsidated RNAs

To determine whether distinct RNAs were present in one virion, we heated purified virions to 65 °C for 10 min followed by an immediate cooling on ice for 10 min. Virion RNA was extracted as described above and analyzed by electrophoresis on a 0.9% agarose-TBE (Tris–borate–ethylenediaminetetraacetic acid) gel.⁵² Some samples

were also transferred to a Nylon membrane and probed with the riboprobe specific to the BMV 3' UTR described above to determine whether the RNAs contained the nearly identical 3' UTR of the BMV RNAs.

Differential scanning fluorimetry

DSF was performed in a 96-well plate in a Stratagene real-time PCR machine. Each reaction had a total volume of 25 μ l containing 4 μ g virion and SYPRO orange (Molecular Probes) at a $2.5 \times$ final concentration. The temperature ramp was set at a rate of 1.0 °C/min from 25 to 95 °C. The fluorescence intensity was measured with Ex470/Em550 and plotted as the first derivative of fluorescence versus temperature. The temperature with the fastest change in fluorescence intensity was extrapolated to be the apparent virus denaturing temperature.

Quantification of RNAs per virion

Spectrophotometry analysis of the virions was performed according to Porterfield and Zlotnick.²⁹ Briefly, the UV absorption spectra were collected using an Agilent 8453 UV–visible spectrophotometer after blanking with virus buffer II. The concentration of nucleotide ([RNA]) and CP monomer ([protein]) was calculated using the following equations:

$$A_{\text{corrected},260} = \epsilon_{\text{RNA}260}[\text{RNA}]L + \epsilon_{\text{protein}260}[\text{protein}]L$$

$$A_{\text{corrected},280} = \epsilon_{\text{RNA}280}[\text{RNA}]L + \epsilon_{\text{RNA}280}[\text{protein}]L$$

where $A_{\text{corrected},260}$ and $A_{\text{corrected},280}$ are the corrected absorbances of virus at wavelengths 260 and 280 nm, respectively, after subtracting the light-scattering part. L is the length of the light path. The values used for $\epsilon_{\text{RNA}260}$, $\epsilon_{\text{RNA}280}$, $\epsilon_{\text{protein}260}$, and $\epsilon_{\text{protein}280}$ were 8000 $\text{M}^{-1} \text{cm}^{-1}$, 4000 $\text{M}^{-1} \text{cm}^{-1}$, $14,400$ $\text{M}^{-1} \text{cm}^{-1}$, and $24,000$ $\text{M}^{-1} \text{cm}^{-1}$, respectively. The concentration of capsid was obtained from that of the CP monomer with the knowledge that each capsid was made up by 180 monomers. The concentration of the RNA divided by that of the capsid was extrapolated to be the length of the RNA per capsid.

EM and image processing

For negative staining EM, samples were applied to glow-discharged carbon-coated copper grids and stained with 1% uranyl acetate. The grids were visualized by a magnification of $30,000 \times$ on a JEOL JEM 1010 electron microscope at an operating voltage of 80 kV. Images were collected on a $4\text{K} \times 4\text{K}$ CCD Gatan Camera.

The WT BMV (1 mg/ml) or 2H_{15} (0.8 mg/ml) was applied onto a R1.2/1.3 Quantifoil grid. The grids were previously washed and glow-discharged. After blotting and being rapidly frozen in liquid ethane using a Vitrobot, we stored the grids in liquid nitrogen. JEOL JEM 3200FC electron microscope (JEOL, Tokyo, Japan), operated at 300 kV, and $4\text{K} \times 4\text{K}$ CCD Gatan Camera (Gatan, Pleasanton, CA) were used for imaging and data collection. The particles were selected from the electron micrographs using EMAN's Boxer routine.⁵³ The particles were translationally and rotationally aligned, classified, and averaged without applying symmetry using EMAN's Refine2d command.

Modeling

To calculate the thermodynamic driving force for RNA packaging, we employed the method of Scheutjens and Fleer.^{32,34,35,54,55} Our calculation is comparable to the self-consistent field calculations performed in Siber and Podgornik,³⁸ but we explicitly model the N-terminal arms. The lattice approach facilitates representation of different sequences. The capsid is modeled as a sphere with an inner radius of 11 nm, which is impenetrable to RNA and protein segments but permeable to water and small ions. The N-terminal tails are treated as flexible polymers grafted to the inner surface of the capsid. The N-terminal tail is considered to include the first 49 amino acid residues of the WT protein and any residues inserted into this region for mutant sequences. The WT sequence thus includes 10 positive charges and 1 negative charge for a net charge of 9. The RNA molecule is represented as a linear polyelectrolyte. For the N-terminal tails and the RNA, the statistical segment length and segment diameters are set equal to the lattice size, $l=0.5$ nm. This value is consistent with the average excluded volume of amino acids and small ions (including their solvent shells). It underestimates the excluded volume of RNA, but we find that the results are insensitive to varying l . The length l is roughly consistent with the distance between unpaired nucleotides in single-stranded RNA, and thus, we consider that each RNA segment corresponds to approximately 1.5 nt. We neglect the effects of base pairing and tertiary structure on the RNA architecture; repeating the calculation with the RNA represented as a branched polymer yielded almost identical free energies.

We consider one RNA molecule confined to the capsid interior; the results would change only slightly if we considered the case of several smaller RNA molecules with the same overall contour length. Except for charge, all species have identical interactions that account for excluded volume. For each protein sequence and RNA length considered, the spatial distribution of densities for each species is calculated by minimizing the semi-grand free energy,^{34,54,56} where the water and small ions are in equilibrium with a bulk solution. As described in the main text, we consider three sets of assumptions. In the first two cases, the RNA charge density is not renormalized. In the first case, the virion is in equilibrium with a bulk solution containing 130 mM 1:1 salt. In the second case, the bulk solution contains 130 mM monovalent cations, 30 mM monovalent anions, and 100 mM negative charges on macromolecules. The capsid is permeable to the monovalent anions and cations, but the macromolecule charges are selectively excluded from the capsid interior. In case 3, we consider the same bulk solution as for case 2, but we assume that the RNA charge density is renormalized by counterion condensation to a linear charge density of e/l_B with $l_B=0.714$ nm, the Bjerrum length.^{57–62} Our implementation follows those described in Refs. 34, 54, and 56; full details are given in the supporting online material.

Reconstructing and modeling cryo-electron microscopic map of 2H₁₅

A total of 11,576 particles were selected from the micrographs for CTF (contrast transfer function) correction and three-dimensional reconstruction. To generate an

initial model for reconstruction, we selected about 3000 particles by MPSA (Multi-Path Simulated Annealing) to build a subnanometer model of 2H₁₅.⁶³ The map was then refined by EMAN.⁵³ A 5.0-Å resolution of the cryo-reconstruction density map was determined by 0.5 FSC (Fourier shell correlation) curve provided by EMAN program (Fig. S6a). An atomic model of the BMV capsid derived from the 3.3-Å crystal structure was then fitted into this cryo-transmission EM map and used as a guide to manually trace the C^α backbone of the 2H₁₅ using program Coot.⁶⁴

Acknowledgements

We thank the Indiana University and Texas A&M Cereal Killers past and present. We thank A. Murali for help with the EMAN program and the Office of the Vice Provost for Research at Indiana University for supporting the small-angle X-ray scattering measurements, and the Indiana University Nanoscale Characterization Facility for access to instrumentation. This work was funded by the National Institutes of Health (NIH) National Institute of Allergy and Infectious Diseases grant 1R01AI090280 (C.K.), the NIH National Institute of Allergy and Infectious Diseases grant R01AI080791 (M.F.H.), and the NIH grant P41GM103832 and Robert Welch Foundation grant Q1242 (W.C.).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2012.03.023](https://doi.org/10.1016/j.jmb.2012.03.023)

References

1. Polisky, B. & McCarthy, B. (1975). Location of histones on simian virus 40 DNA. *Proc. Natl Acad. Sci. USA*, **72**, 2895–2899.
2. Sung, M. T., Cao, T. M., Coleman, R. T. & Budelier, K. A. (1983). Gene and protein sequences of adenovirus protein VII, a hybrid basic chromosomal protein. *Proc. Natl Acad. Sci. USA*, **80**, 2902–2906.
3. Ruigrok, R. W., Crépin, T. & Kolakofsky, D. (2011). Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr. Opin. Microbiol.* **14**, 504–510.
4. Smith, T. J., Chase, E., Schmidt, T. & Perry, K. L. (2000). The structure of cucumber mosaic virus and comparison to cowpea chlorotic mottle virus. *J. Virol.* **74**, 7578–7586.
5. Silva, A. M. & Rossmann, M. G. (1985). The refinement of southern bean mosaic virus in reciprocal space. *Acta Crystallogr., Sect. B: Struct. Sci.* **41**, 147–157.
6. Choi, H. K., Lu, G., Lee, S., Wengler, G. & Rossmann, M. G. (1996). Structure of Semliki Forest virus core protein. *Proteins*, **27**, 345–359.

7. Fisher, A. J. & Johnson, J. E. (1993). Ordered duplex RNA controls capsid architecture in an icosahedral animal virus. *Nature*, **361**, 176–179.
8. Tong, L., Wengler, G. & Rossmann, M. G. (1993). Refined structure of Sindbis virus core protein and comparison with other chymotrypsin-like serine proteinase structures. *J. Mol. Biol.* **230**, 228–247.
9. Speir, J. A., Munshi, S., Wang, G., Baker, T. S. & Johnson, J. E. (1995). Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure*, **3**, 63–77.
10. Belyi, V. A. & Muthukumar, M. (2006). Electrostatic origin of the genome packing in viruses. *Proc. Natl Acad. Sci. USA*, **103**, 17174–17178.
11. Hu, T., Zhang, R. & Shkrovskii, B. I. (2008). Electrostatic theory of viral self-assembly. *Phys. A (Amsterdam, Neth.)*, **387**, 3059–3064.
12. Kaplan, I. B., Zhang, L. & Palukaitis, P. (1998). Characterization of cucumber mosaic virus. *Virology*, **246**, 221–231.
13. Venter, P. A., Marshall, D. & Schneemann, A. (2009). Dual roles for an arginine-rich motif in specific genome recognition and localization of viral coat protein to RNA replication sites in flock house virus-infected cells. *J. Virol.* **83**, 2872–2882.
14. Dong, X. F., Natarajan, P., Tihova, M., Johnson, J. E. & Schneemann, A. (1998). Particle polymorphism caused by deletion of a peptide molecular switch in a quasispherical icosahedral virus. *J. Virol.* **72**, 6024–6033.
15. Marshall, D. & Schneemann, A. (2001). Specific packaging of nodaviral RNA2 requires the N-terminus of the capsid protein. *Virology*, **285**, 165–175.
16. Rao, A. L. N. (2006). Genome packaging by spherical plant RNA viruses. *Annu. Rev. Phytopathol.* **44**, 61–87.
17. Kao, C. C. & Sivakumaran, K. (2000). Brome mosaic virus, good for an RNA virologist's basic needs. *Mol. Plant Pathol.* **1**, 91–97.
18. Lucas, R. W., Larson, S. B. & McPherson, A. (2002). The crystallographic structure of brome mosaic virus. *J. Mol. Biol.* **317**, 95–108.
19. Sacher, R. & Ahlquist, P. (1989). Effects of deletions in the N-terminal basic arm of brome mosaic virus coat protein on RNA packaging and systemic infection. *J. Virol.* **63**, 4545–4552.
20. Larson, S. B., Lucas, R. W. & McPherson, A. (2004). Crystallographic structure of the $T = 1$ particle of brome mosaic virus. *J. Mol. Biol.* **346**, 815–831.
21. Calhoun, S. L., Speir, J. A. & Rao, A. L. N. (2007). *In vivo* particle polymorphism results from deletion of a N-terminal peptide molecular switch in brome mosaic virus capsid protein. *Virology*, **364**, 407–421.
22. Rao, A. L. & Grantham, G. L. (1996). Molecular studies on bromovirus capsid protein. II. Functional analysis of the amino-terminal arginine-rich motif and its role in encapsidation, movement, and pathology. *Virology*, **226**, 294–305.
23. Choi, Y. G. & Rao, A. L. (2000). Molecular studies on bromovirus capsid protein. VII. Selective packaging on BMV RNA4 by specific N-terminal arginine residuals. *Virology*, **275**, 207–217.
24. Michel, J. P., Ivanovska, I. L., Gibbons, M. M., Klug, W. S., Knobler, C. M., Wuite, G. J. & Schmidt, C. F. (2006). Nanoindentation studies of full and empty viral capsids and the effects of capsid protein mutations on elasticity and strength. *Proc. Natl Acad. Sci. USA*, **103**, 6184–6189.
25. Roos, W. H., Gertsman, I., May, E. R., Brooks, C. L., 3rd, Johnson, J. E. & Wuite, G. J. (2012). Mechanics of bacteriophage maturation. *Proc. Natl Acad. Sci. USA*, **109**, 2342–2347.
26. Gopinath, K., Dragnea, B. & Kao, C. (2005). Interaction between brome mosaic virus proteins and RNAs: effects on RNA replication, protein expression, and RNA stability. *J. Virol.* **79**, 14222–14234.
27. Argos, P. (1981). Secondary structure prediction of plant virus coat proteins. *Virology*, **110**, 55–62.
28. Hema, M., Murali, A., Ni, P., Vaughan, R. C., Fujisaki, K., Tsvetkova, I. *et al.* (2010). Effects of amino-acid substitutions in the brome mosaic virus capsid protein on RNA encapsidation. *Mol. Plant-Microbe Interact.* **23**, 1433–1447.
29. Porterfield, J. Z. & Zlotnick, A. (2010). A simple and general method for determining the protein and nucleic acid content of viruses by UV absorbance. *Virology*, **407**, 281–288.
30. Hema, M., Gopinath, K. & Kao, C. (2005). Repair of the tRNA-like CCA sequence in a multipartite positive-strand RNA virus. *J. Virol.* **79**, 1417–1427.
31. Krishna, N. K. & Schneemann, A. (1999). Formation of an RNA heterodimer upon heating of nodavirus particles. *J. Virol.* **73**, 1699–1703.
32. Scheutjens, J. & Fleer, G. J. (1980). Statistical theory of the adsorption of interacting chain molecules. 2. Train, loop, and tail size distribution. *J. Phys. Chem.* **84**, 178–190.
33. Scheutjens, J. M. H. M. & Fleer, G. J. (1979). Statistical theory of the adsorption of interacting chain molecules. 1. Partition function, segment density distribution, and adsorption isotherms. *J. Phys. Chem.* **83**, 1619–1635.
34. Bohmer, M. R., Evers, O. A. & Scheutjens, J. M. H. M. (1990). Weak polyelectrolytes between two surfaces: adsorption and stabilization. *Macromolecules*, **23**, 2288–2301.
35. Bohmer, M. R., Koopal, L. K. & Lyklema, J. (1991). Micellization of ionic surfactants: calculations based on a self-consistent field lattice model. *J. Phys. Chem.* **95**, 9569–9578.
36. Ting, C. L., Wu, J. & Wang, Z. G. (2011). Thermodynamic basis for the genome to capsid charge relationship in viral encapsidation. *Proc. Natl Acad. Sci. USA*, **108**, 16986–16991.
37. Manning, G. S. (1969). Limiting laws and counterion condensation in polyelectrolyte solutions. I. Colligative properties. *J. Chem. Phys.* **51**, 924.
38. Siber, A. & Podgornik, R. (2008). Nonspecific interactions in spontaneous assembly of empty versus functional single-stranded RNA viruses. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **78**, 051915.
39. Angelescu, D. G., Bruinsma, R. & Linse, P. (2006). Monte Carlo simulations of polyelectrolytes inside viral capsids. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **73**, 041942.
40. van der Schoot, P. & Bruinsma, R. (2005). Electrostatics and the assembly of an RNA virus. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **71**, 061928.
41. Elrad, O. M. & Hagan, M. F. (2010). Encapsulation of a polymer by an icosahedral virus. *Phys. Biol.* **7**, 045003.

42. Kivenson, A. & Hagan, M. F. (2010). Mechanisms of capsid assembly around a polymer. *Biophys. J.* **99**, 619.
43. Choi, Y. G., Dreher, T. W. & Rao, A. L. N. (2002). tRNA elements mediate the assembly of an icosahedral RNA virus. *Proc. Natl Acad. Sci. USA*, **99**, 655–660.
44. Sivakumaran, K., Kim, C. H., Tayon, R., Jr & Kao, C. C. (1999). RNA sequence and secondary structural determinants in a minimal viral promoter that directs replicase recognition and initiation of genomic plus-strand RNA synthesis. *J. Mol. Biol.* **294**, 667–682.
45. Annamalai, P. & Rao, A. L. N. (2005). Replication-independent expression of genome components and capsid protein of brome mosaic virus in planta: a functional role for viral replicase in RNA packaging. *Virology*, **338**, 96–111.
46. Annamalai, P. & Rao, A. L. N. (2006). Packaging of brome mosaic virus subgenomic RNA is functionally coupled to replication-dependent transcription and translation of coat protein. *J. Virol.* **80**, 10096–10108.
47. Annamalai, P., Rofail, F., Demason, D. A. & Rao, A. L. N. (2008). Replication-coupled packaging mechanism in positive-strand RNA viruses: synchronized coexpression of functional multigenome RNA components of an animal and a plant virus in *Nicotiana benthamiana* cells by agroinfiltration. *J. Virol.* **82**, 1484–1495.
48. Cadena-Nava, R. D., Comas-Garcia, M., Garmann, R. F., Rao, A. L., Knobler, C. M. & Gelbart, W. M. (2011). Self-assembly of viral capsid protein and RNA molecules of different sizes: requirement for a specific high protein/RNA mass ratio. *J. Virol.*, doi:10.1128/JVI.06566-11.
49. Bujarski, J. J. (1998). Bromovirus isolation and RNA extraction. *Methods Mol. Biol.* **81**, 183–188.
50. Proksch, R., Schäffer, T. E., Cleveland, J. P., Callahan, R. C. & Viani, M. B. (2004). Finite optical spot size and position corrections in thermal spring constant calibration. *Nanotechnology*, **15**, 1344–1350.
51. Sambrook, J. & Russell, D. W. (2006). Separation of RNA according to size: electrophoresis of glyoxylated RNA through agarose gels. *Cold Spring Harbor Protoc.* doi:10.1101/pdb.prot4057.
52. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
53. Ludtke, S. J., Baldwin, P. R. & Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97.
54. Israels, R., Leermakers, F. A. M. & Fleer, G. J. (1994). On the theory of grafted weak polyacids. *Macromolecules*, **27**, 3087–3093.
55. van Lent, B., Israels, R., Scheutjens, J. & Fleer, G. J. (1990). Interaction between hairy surfaces and the effect of free polymer. *J. Colloid Interface Sci.* **137**, 380–394.
56. Hagan, M. F. (2009). A theory for viral capsid assembly around electrostatic cores. *J. Chem. Phys.* **130**, 114902.
57. Manning, G. S. (1969). Limiting laws and counterion condensation in polyelectrolyte solutions. 2. Self-diffusion of small ions. *J. Chem. Phys.* **51**, 934–938.
58. Manning, G. S. (1969). Limiting laws and counterion condensation in polyelectrolyte solutions. 3. An analysis based on Mayer ionic solution theory. *J. Chem. Phys.* **51**, 3249–3252.
59. Manning, G. S. (2007). Counterion condensation on charged spheres, cylinders, and planes. *J. Phys. Chem. B*, **111**, 8554–8559.
60. Muthukumar, M. (1996). Double screening in polyelectrolyte solutions: limiting laws and crossover formulas. *J. Phys. Chem.* **105**, 5183–5199.
61. Patra, C. N. & Yethiraj, A. (1999). Density functional theory for the distribution of small ions around polyions. *J. Phys. Chem. B*, **103**, 6080–6087.
62. Kundagrami, A. & Muthukumar, M. (2008). Theory of competitive counterion adsorption on flexible polyelectrolytes: divalent salts. *J. Chem. Phys.* **128**, 244901.
63. Liu, X., Jiang, W., Jakana, J. & Chiu, W. (2007). Averaging tens to hundreds of icosahedral particle images to resolve protein secondary structure elements using a Multi-Path Simulated Annealing optimization algorithm. *J. Struct. Biol.* **160**, 11–27.
64. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.