

RNA Homopolymers Form Higher-Curvature Virus-like Particles Than Do Normal-Composition RNAs

Abby R. Thurm,¹ Christian Beren,¹ Ana Luisa Duran-Meza,¹ Charles M. Knobler,¹ and William M. Gelbart^{1,2,3,*}

¹Department of Chemistry and Biochemistry, ²Molecular Biology Institute, and ³California NanoSystems Institute, University of California, Los Angeles, California

ABSTRACT Unlike double-stranded DNA, single-stranded RNA can be spontaneously packaged into spherical capsids by viral capsid protein (CP) because it is a more compact and flexible polymer. Many systematic investigations of this self-assembly process have been carried out using CP from cowpea chlorotic mottle virus, with a wide range of sequences and lengths of single-stranded RNA. Among these studies are measurements of the relative packaging efficiencies of these RNAs into spherical capsids. In this work, we address a fundamental issue that has received very little attention, namely the question of the preferred curvature of the capsid formed around different RNA molecules. We show in particular that homopolymers of RNA—polyribouridylic acid and polyriboadenylic acid—form exclusively $T = 2$ -sized (~ 22 -nm diameter) virus-like particles (VLPs) when mixed with cowpea chlorotic mottle virus CP, independent of their length, ranging from 500 to more than 4000 nucleotides. This is in contrast to “normal-composition” RNAs (i.e., molecules with comparable numbers of each of the four nucleotides and hence capable of developing a large amount of secondary structure because of intramolecular complementarity/basepairing); a curvature corresponding to $T = 3$ -size (~ 28 nm in diameter) is preferred for the VLPs formed with such RNAs. Our work is consistent with the preferred curvature of VLPs being a consequence of interaction of CP with RNA—in particular, the presence or absence of short RNA duplexes—and suggests that the equilibrium size of the capsid results from a trade-off between this optimum size and the cost of confinement.

SIGNIFICANCE We address the question of the preferred curvature of the capsid formed around different RNA molecules. We show in particular that homopolymers of RNA—polyribouridylic acid and polyriboadenylic acid—form exclusively $T = 2$ -sized (~ 22 nm-diameter) virus-like particles when mixed with cowpea chlorotic mottle virus capsid protein, independent of their length, ranging from 500 to more than 4000 nucleotides. This is in contrast to “normal-composition” RNAs, which are capable of developing a large amount of secondary structure due to intramolecular complementarity/basepairing. Our work is consistent with the preferred curvature of virus-like particles being a consequence of the interaction of capsid protein with RNA—in particular, the presence or absence of short RNA duplexes—and suggests the need for further research on the microscopic basis for these effects.

INTRODUCTION

Since pioneering studies over 50 years ago (1,2), much has been learned about the *in vitro* self-assembly of viruses and virus-like particles (VLPs) from purified RNA and capsid protein (CP) (3–6). The first example of the reconstitution of a spherical virus is that of cowpea chlorotic mottle virus (CCMV) (2) and arguably the most systematic investigations of *in vitro* self-assembly of VLPs have been done with CCMV CP and a wide range of RNA sequences and

lengths (4,6,7). What all of these RNA molecules have in common is that they have “normal composition” (i.e., comparable numbers of each of the four nucleotides, nt), thereby giving rise to a significant amount of secondary structure—intramolecular self-complementarity/basepairing.

One qualitative result (8) of these studies is that the VLPs self-assembled around RNAs with lengths less than 2500 nucleotides (nt) are ~ 22 nm in diameter, a size typical of 120-subunit “ $T = 2$ ” capsids, which are not in the Caspar and Klug (9) hierarchy but which had been identified by Krol et al. (10). Longer RNAs form VLPs whose diameter is 28 nm as measured by TEM or 26 nm as measured by cryoelectron microscopy and are associated with the wild-type (WT) $T = 3/180$ -subunit structure (11). For lengths less than

Submitted February 28, 2019, and accepted for publication August 7, 2019.

*Correspondence: gelbart@chem.ucla.edu

Editor: Gijs Wuite.

<https://doi.org/10.1016/j.bpj.2019.08.012>

© 2019 Biophysical Society.



1500 nt, two or more copies of RNA are packaged in each capsid, and for lengths greater than 4000 nt, one finds two or more capsids sharing a single RNA (4). It is striking that RNA molecules two times longer than the viral length form doublets of capsids associated with a single molecule, RNAs three times longer form triplets, and so on rather than the progressively “extra” length simply being accommodated by bigger capsids. This is because $T = 3$ curvature—involving 180 subunits forming a 28-nm diameter capsid—is the strongly preferred curvature of CCMV CPs interacting with RNA.

Another qualitative result of *in vitro* reconstitution studies of CCMV-VLPs comes from experiments in which two RNAs of different lengths and/or different sequences compete for a limited amount of CCMV CP, allowing the relative efficiencies of assembly to be determined quantitatively (12). In competitions of the 3200-nt viral RNA1 with RNAs obtained by truncation or augmentation, it is found that the viral length is most efficiently packaged, but the efficiency of assembly depends on sequence as well as on length. In particular, an experiment in which a 3200-nt CCMV RNA was allowed to compete with a virtually identical length of RNA associated with the related brome mosaic virus (BMV) showed that the heterologous BMV RNA outcompeted the CCMV RNA. These results can be understood in terms of the role of RNA secondary structure; on the one hand, local motifs like stem loops with specific sequences in the single- and double-stranded portions—“packaging signals”—bind CP with special affinity, and on the other hand, large-scale motifs involving basepairing between distant parts of the sequence give rise to an overall size for the RNA that is most consistent with the preferred curvature of the CP. In fact, measurements of the three-dimensional size of CCMV RNA1 using cryoelectron microscopy (13,14) reveal that it is roughly the same size as the 28-nm capsid into which it is packaged.

To test our understanding of the role of sequence-dependent secondary structure on the relative packaging efficiencies of different RNA molecules, we recently investigated (15) the *in vitro* reconstitution of VLPs from CCMV CP and polyribouridylic acid (polyU), an RNA molecule for which secondary structure formation—self-complementary basepairing—is not possible. This homopolymer is singular among the four RNA homopolymers by virtue of its having the weakest base-stacking interactions between neighboring nt (16), so that it does not undergo a helix-coil transition; it is the most structureless, disordered, RNA molecule of all. In comparing these studies with our previous investigations with normal-composition RNA, we found two surprises. First, only $T = 2$ -sized VLPs formed, even as the length of polyU was varied from 1000 to more than 7000 nt and even though the hydrodynamic diameter of each polyU molecule was larger than that of a normal-composition RNA of the same length. Second, viral length (~ 3000 -nt long) polyU was packaged more

efficiently than viral RNA when competing for insufficient CP to package both of them.

In this work, we study polyriboadenylic acid (polyA), a more structured RNA homopolymer because of its strong base-stacking interactions (17). Like polyU, because of the absence of self-complementarity/secondary structure formation, polyA does not exhibit the effective branching typical of viral RNAs (18,19), but unlike polyU, it is not a simple linear polymer because it adopts a helical conformation as a result of the interactions between adenine moieties (20). This allows the effect of this additional degree of structuring on the efficiency of packaging to be examined experimentally. Using the two-step assembly conditions that we have previously found to be optimum for *in vitro* packaging (4)—allowing the CP to interact with the RNA at a neutral pH and low ionic strength and then lowering the pH to 4.75 to strengthen CP-CP interactions—we find that CCMV CP is unable to package polyA RNA effectively under these conditions as negative-stain electron microscopy (EM) shows the formation of VLPs of irregular size and shape; they also do not protect the packaged RNA from attack by RNase. This failure in the packaging of polyA can be attributed to the formation of a double-helical structure at a low pH and temperature (21), arising from hydrogen bonding between the helically ordered adenines. VLPs can, however, be formed successfully when the second step of assembly is carried out at a pH no lower than 5.8, so that helical duplexes of polyA do not occur. The VLPs formed under these conditions are slightly irregular in shape but are closed shells protective of their encapsidated RNA. Like the assemblies with polyU (15), the VLPs have sizes that are typical of $T = 2$ capsids, independent of the length of the packaged RNA, and are expected to form multiplets (4) of $T = 2$ -sized capsids for lengths greater than ~ 3000 nt. But unlike polyU, polyA is less efficiently packaged than viral RNA. These results can be understood in terms of the competing demands on the CP as it forms a VLP—more explicitly, minimization of the protein-protein interaction energy (as determined by the preferred curvature of CP interacting with the kind of RNA involved (e.g., $T = 2$ for the homopolymers and $T = 3$ for the heteropolymers)) and of the free energy cost associated with confinement of the RNA in the capsid (again determined by the kind of RNA involved).

MATERIALS AND METHODS

Buffers

The following buffers were employed: buffer B (protein buffer; 20 mM Tris-HCl (pH 7.2), 1 M NaCl, 1 mM EDTA, 1 mM dithiothreitol, and 1 mM phenylmethylsulfonyl fluoride); RNA assembly buffer (RAB; 50 mM Tris-HCl (pH 7.2), 50 mM NaCl, 10 mM KCl, and 5 mM $MgCl_2$); virus storage buffer (VSB; 50 mM sodium acetate (pH 4.5) and 8 mM magnesium acetate); high-pH VSB (50 mM sodium acetate (pH 6) and 8 mM magnesium acetate); 10 mM Tris-HCl (pH 7.5) and 1 mM EDTA (TE buffer); disassembly buffer (50 mM Tris-HCl (pH 7.5),

0.5 M CaCl₂, 1 mM EDTA, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride); virus running buffer (VRB; 100 mM sodium acetate (pH 5.5) and 1 mM EDTA); and buffer containing 40 mM Tris-HCl (pH 8), 20 mM acetic acid, and 1 mM EDTA (TAE).

PolyA and polyU synthesis and fractionation

PolyA was purchased from Sigma-Aldrich (St. Louis, MO) as a solid potassium salt with molecular weight varying from 0.1 to 2 MDa (many lengths included). This polydisperse mixture was dissolved in TE buffer at a concentration of 1 mg/mL and run on a denaturing 1% TAE agarose gel. Using an ssRNA ladder as a guide for length, the RNA was separated into discrete-length segments—fractions of 500–1500, 1500–2500, 2500–3500, and > 3500 nt polyA—and purified from the gel by electroelution. The amount of polyA purified was quantified by ultraviolet-visible absorbance measurements with a NanoDrop Spectrophotometer. The absorbance ratio at 260/280 nm was used as a measure of RNA purity ($A_{260}/A_{280} > 2.0$). To verify successful fractionation, the different length samples of polyA were run on a denaturing 1% TAE agarose gel (see Fig. S1). The synthesis and fractionation of polyU were carried out as described by Beren et al. (15).

In vitro transcription of BMV RNA1

BMV RNA1 (B1) was made by in vitro transcription of the linearized DNA plasmid pT7B1 (4) with T7 polymerase (provided by Dr. Feng Guo). Transcription reactions were carried out for 3 h at 37°C, followed by digestion using DNase I (New England BioLabs) for 1 h at 37°C, and the RNA was then purified from the transcription mix by washing through a 100-kDa Amicon filter with TE buffer. The concentration of RNA was determined by ultraviolet-visible absorbance using a NanoDrop spectrophotometer, and a A_{260}/A_{280} ratio of > 2.0 was used to verify RNA purity.

Purification of CCMV CP

CCMV CP was obtained and purified as described previously (2). Briefly, the virus was obtained from infected California cowpea plants (*Vigna unguiculata* black-eyed pea) and after an overnight dialysis into disassembly buffer, the CP was isolated as described by Annamalai and Rao (22). Matrix-assisted laser desorption/ionization-time of flight mass spectrometry (Voyager-DE-STR; Applied Biosystems) was employed to ascertain that the N-terminus was not cleaved during purification.

In vitro packaging of RNA by CCMV CP

RNA and CP were mixed in buffer B at a 4.2:1 CP/RNA mass ratio to a final RNA concentration of 30 ng/L and then dialyzed overnight into RAB at 4°C. It has been shown that this is the minimal mass ratio required to ensure that all of the RNA is packaged (4). After 12–16 h, the assembly contents were transferred to VSB at either a pH of 4.75 (for normal-composition RNA and polyU) or 6 (for polyA) and dialyzed at 4°C. After 6 h, the VLPs were collected for analysis.

Electrophoretic mobility shift assay

The products of assembly reactions were run on 1% native agarose gels in VRB. 5 μ g of virus was loaded per lane, and all samples were run against the same mass of WT CCMV as a control. Virions containing B1 were visualized using GelRed Nucleic Acid Stain (catalog number 41001; Biotium), whereas polyA VLPs were stained using Coomassie protein stain because polyA—because of its lack of secondary structure—is stained less efficiently by GelRed. Gel bands were visualized on a Pharos FX Plus Molecular Imager.

TEM

Negative-stain EM grids were prepared by depositing 6 μ L of assembly reaction on glow-discharged copper grids (400 mesh), which had previously been coated with Parlodion and carbon. After a 1-min deposition, grids were blotted and stained with 6 μ L of 2% uranyl acetate for 1 min and then blotted dry. Grids were stored in a desiccator and imaged on an FEI Tecnai TF20 EM (Electron Imaging Center for NanoMachines, California NanoSystems Institute at University of California, Los Angeles) at 50,000 $\times$ magnification. Particles were sized manually from images taken with negative-stain transmission EM using ImageJ (23). At least 200 particles were measured per sample.

Dynamic light scattering measurements

40 μ L aliquots of 300 ng/ μ L fractionated 2500–3500-nt polyA in RAB were measured using a Malvern Zetasizer Nano ZSP (Malvern Instruments, Malvern, UK).

RNase A digestion of VLPs

RNase digestion of virions and VLPs was used to confirm that their capsids were fully closed shells. VLPs were incubated with RNase A (PureLink RNase A; Thermo Fisher Scientific, Waltham, MA) for 1 h at 4°C at a ratio of 2.5 ng RNase A to 1 μ g VLP. WT CCMV, when subjected to the same treatment, is unaffected, but this low-concentration RNase treatment is capable of digesting at least 10 g of unpackaged RNA. After RNase treatment, the mixture was incubated with 40 units of RNase inhibitor (Protector RNase Inhibitor; Sigma-Aldrich) for 15 min at 4°C. The digestion products were then washed through a 100-kDa Amicon filter at least three times with VSB buffer to remove the RNase and the RNase inhibitor.

Fluorescence hybridization assay

200 pmol of fluorescein amidite-labeled polyT 15-nt DNA oligos (Integrated DNA Technologies, Skokie, IL) was incubated with 10 μ g samples of both WT CCMV and polyA VLPs at 4°C for 1 h. Without excess oligo being washed away, the finished incubation products were run on a 1% agarose gel in VRB and imaged using appropriate fluorescence excitation on a Pharos FX Plus Molecular Imager (see Fig. S5).

Competition experiments

The protocols used for the competition experiments are similar to those previously employed (12). Equal masses of two different types of RNA were incubated in buffer B at a total RNA concentration of 30 ng/ μ L and a final volume of 200 μ L, with only half the amount of CP required to package the RNA completely, a final CP/RNA mass ratio of 2.1:1. This assembly mixture was then dialyzed into RAB overnight at 4°C, followed by another 6 h dialysis in VSB at a pH of 6.

RESULTS

None of the polyA fractions are packaged completely by CCMV CP under canonical conditions

Following the procedure employed successfully for the packaging of a wide range of RNAs (4), fractionated lengths of polyA (Fig. S1) were mixed with purified CCMV CP at a 4.2:1 CP/RNA mass ratio in buffer B, a neutral-pH and

high-salt buffer, and dialyzed overnight in RAB, a neutral-pH and low-salt buffer, to establish initial RNA-protein interaction. After 12 h, the mixture was moved to low-salt, low-pH (4.8) VSB and dialyzed for an additional 6 h to strengthen protein-protein interactions and complete capsid formation.

The reaction mixture was run on a native 1% agarose gel, and the VLPs were visualized with Coomassie Blue protein stain (Fig. S2); WT CCMV was run as a control and visualized in the same manner. Although a single band is visible for each polyA fraction, the bands are faint, do not all run identically next to each other, and run more slowly than either WT CCMV or polyU VLPs (not shown in same gel). PolyU VLPs run like the “typical” VLP as a narrow band slightly slower than WT CCMV. The retarded and broader bands for the polyA assemblies suggest inconsistent packaging and are an initial indication of irregularities in the capsid structure of or lack of formation of polyA VLPs.

VLPs assembled around 500–1500, 1500–2500, 2500–3500, and > 3500 nt polyA were visualized in negative-stain transmission electron microscopy (TEM) before digestion with RNase A (Fig. S3). The particles of the 500–1500-nt length VLPs were highly irregular, but a sufficient number of the micrographs could be analyzed to enable a size determination to be performed (see bottom of Fig. S3). 1500–2500 nt and longer polyA VLPs were more obviously irregular, and insufficiently many well-formed capsids were visible to allow a size determination. The particles did not survive treatment with RNase A for 1 h: no polyA VLP bands were observed when the digested assemblies were run on an agarose gel as above, and none were seen in negative-stain TEM (Fig. S4).

Further evidence of the incomplete packaging of polyA was provided by an experiment in which 20 pmol of 20 nt polyT DNA oligos labeled with fluorescein amidite were incubated with 1 μ g of polyA VLPs for 1 h (this corresponds to a mole ratio of VLP/dT oligo of 1:100). When the mixture was run on an agarose gel, extensive binding of the oligo to the polyA was evident (Fig. S5). In contrast, no association was observed when the oligo was incubated with WT CCMV, showing that the malformed polyA VLPs fail to protect the RNA from hybridization with polyT oligos.

PolyA undergoes significant structural changes at a low pH

Difficulty in packaging polyA by a procedure involving the lowering of the pH to 4.8 is consistent with early measurements of polyA structures. In particular, viscosity and sedimentation experiments on polyA were interpreted by Fresco and Doty (20) as evidence of the formation of a rigid rod-like structure, which they identified with the double-helical form seen in fiber x-ray measurements by Rich et al. (21) under acidic conditions. The helical structure is made up of pairs of parallel chains of stacked protonated adenine

bases that are held together by hydrogen bonds between apposing bases and by dispersion (“stacking”) interactions between neighboring bases. To examine this, we have performed dynamic light scattering (DLS) measurements of 1500–2500-nt polyA in RAB as a function of pH by the addition of 0.08 M hydrochloric acid to the solution. A large jump in R_h to ~500 nm was seen as the pH was decreased to 5, in contrast to the 25 nm R_h at a pH of 6 and above (Fig. 1). This change was also seen to be temperature independent as 1500–2500-nt polyA at a pH of 7.2 was measured to have a R_h of 25 nm at both 4 and 25°C. The DLS measurements demonstrate that the challenges in encapsidating polyA at typical low-pH conditions come from the stiff double-stranded structures induced by lowering the pH of the solution below the 5.8 threshold. In contrast, both viral B1 and 1500–2500-nt polyU show no significant change in R_h with changes in pH (see Fig. S6).

PolyA can be encapsidated completely at a higher pH

When the pH in the second step of the assembly procedure was 6 instead of 4.75, polyA VLPs were formed that were RNase resistant, although irregular in shape. The presence of the particles was verified by agarose gel electrophoresis in pH 7 running buffer so as to maintain neutral pH (Fig. 2). PolyA VLPs assembled from four lengths (500–1500, 1500–2500, 2500–3500, and > 3500 nt) ran as narrow bands in the expected position when visualized with Coomassie Blue, indicative of more complete particles than seen for assemblies carried out at a low pH. The capsids were then treated with RNase A as before, and gel analysis

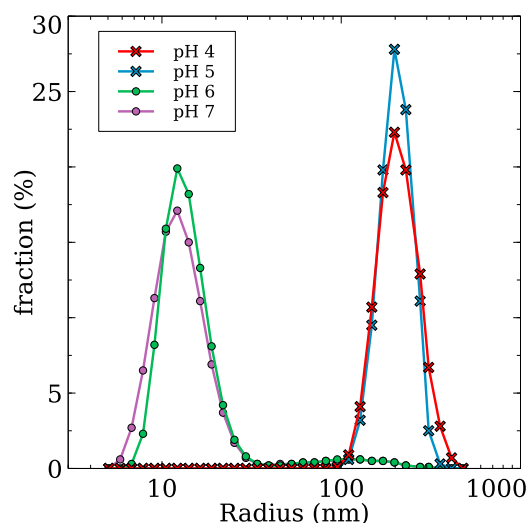


FIGURE 1 Effect of pH on volume-weighted hydrodynamic radius distributions of 1500–2500 nt polyA as measured by DLS. Measurements were made at 25°C. Clear size distinction is seen between radii measured at a pH of 6 and above (green and purple curves with circle symbols) and those at a pH of 4 and 5 (red and blue curves with cross symbols). To see this figure in color, go online.

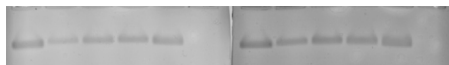


FIGURE 2 Gel analysis of polyA VLPs assembled at pH 6, pre- (left gel) and post- (right gel) treatment with RNase A. Samples are, in both gels, as follows: lane 1, WT CCMV; lane 2, 500–1500-nt polyA VLP; lane 3, 1500–2500-nt polyA VLP; lane 4, 2500–3500-nt (viral length) polyA VLP; lane 5, > 3500 nt polyA VLP; and lane 6 is empty. In the right gel, all samples were treated identically with RNase A. Both gels were run in TAE plus 2 mM CaCl_2 to maintain a neutral pH and prevent capsid swelling.

showed that each of the samples survived digestion (Fig. 2; Fig. S7) and were stable at a low pH.

TEM analysis of the VLPs revealed irregularly shaped capsids that were nevertheless sufficiently uniform to allow

size distributions to be measured. 200 particles of each sample were analyzed, with the distributions for each length of polyA peaking between 23 and 25 nm in diameter—larger than polyU $T = 2$ -sized capsids (10) but smaller than WT CCMV 28 nm $T = 3$ capsids (Fig. 3). After treatment with RNase A, the capsids appeared more regular as many of the more malformed particles were digested by the RNase. The distributions were centered around 23 nm and narrowed, with the exception of that for VLPs for the longest-length (>3500 nt) polyA, which was centered at 25 nm.

Axial ratios for the particles were determined from EM images by measuring the ratio of the vertical and horizontal particle diameters for at least 200 particles per sample. In cases where the ratio is less than 1, we took the inverse so that all

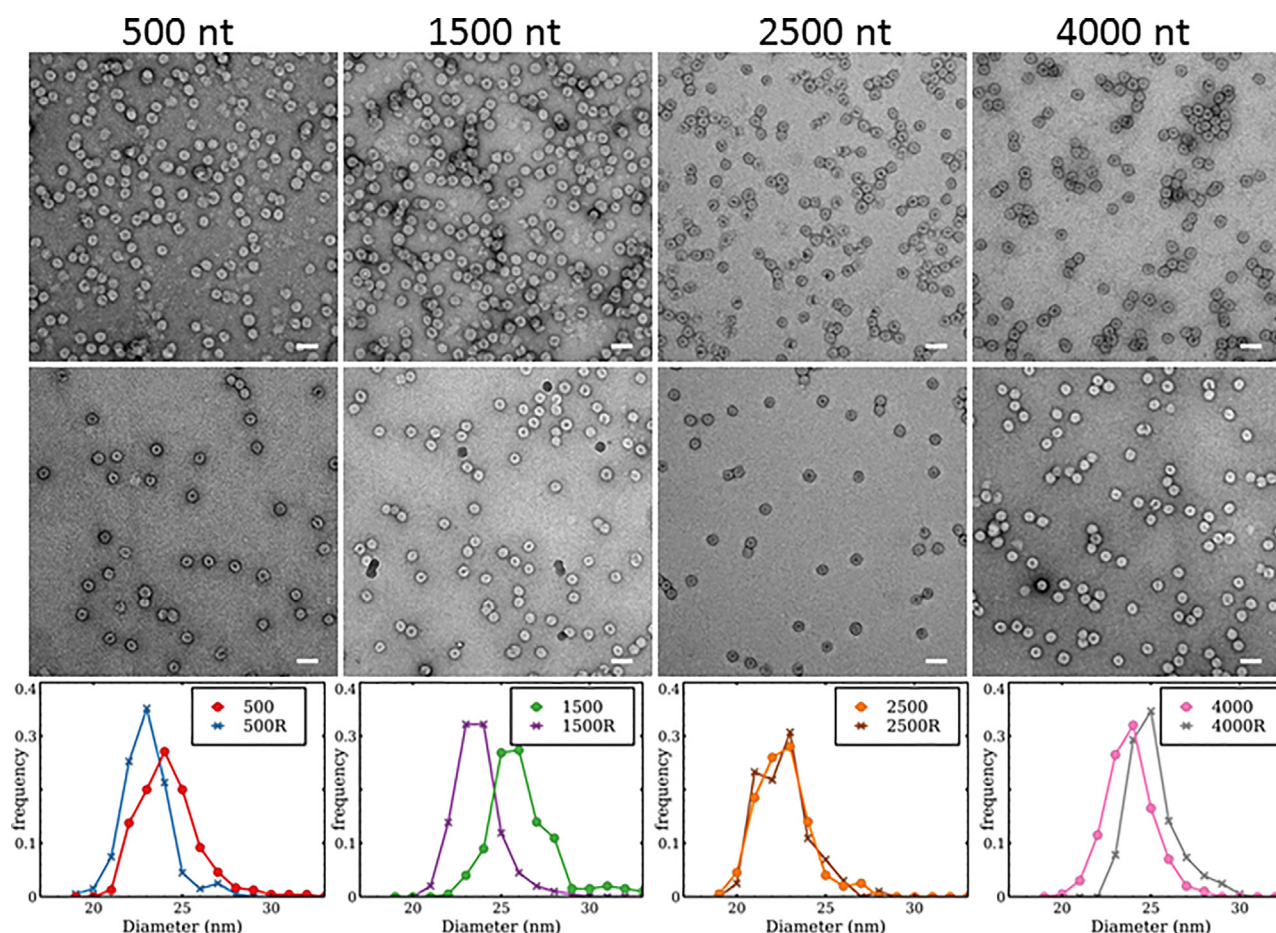


FIGURE 3 Top: electron micrographs of VLPs assembled at pH 6 from (left to right) 500–1500, 1500–2500, 2500–3500, and > 3500 nt polyA. The capsids are more uniform than those assembled at a low pH. Although still irregular in size with wide size frequency distributions, the particles are sufficiently well formed and spherical in shape for their diameters to be determined. Middle: electron micrographs of these same VLPs after being subjected to RNase. The particles that survived the RNase treatment appear to be qualitatively better formed and more consistent with size distributions peaking around 23 nm except for those formed from the longest-length polyA. Bottom: the effect of RNase treatment on polyA VLP size distributions was determined from the measurement of the diameters of 200 well-formed particles of each sample, before and after RNase treatment. VLPs containing shorter-than-viral lengths of polyA appear to become slightly more homogeneous and slightly smaller on average, viral length (2500–3500 nt) VLPs show no change, and longer-length (> 3500 nt) VLPs become slightly more homogeneous yet also slightly larger in average size. In all cases, the size changes are no more than 1–2 nm, comparable to the uncertainty in our measurements, consistent with VLPs that effectively protect against RNase. Left to right: VLPs formed from 500–1500, 1500–2500, 2500–3500, and > 3500 nt polyA are shown. Scale bars represent 50 nm. To see this figure in color, go online.

values are larger than 1. As shown in Fig. 4, polyA VLPs display a wider range of axial ratios than do WT CCMV, typical T = 3 VLPs made with B1, “T = 2” polyU VLPs, or “T = 2” VLPs containing short (500 nt) normal-composition RNA. And the distribution shows a smaller percentage of particles with ratios close to or ~ 1 , even when compared to “T = 2”-size capsids that tend to be more irregular in shape. The polyA axial ratios are also not discretized as found in the case of polyU and short RNA “T = 2” VLPs, which display axial ratios only below 1.1 and above 1.15, consistent with irregularities in shape that can be caused by additional lines of protein hexamers across the middle of the VLP (24). The presence of polyA axial ratios continuously up to ~ 1.25 implies that larger axial ratios in this case are caused by randomly irregular capsids and not by any special structures as in the case of polyU and short RNA VLPs.

B1 is more efficiently packaged than polyA

Competition experiments in which the second step in the assembly protocol was carried out at a pH of 4.75 have shown that when equal masses of 3000-nt polyU RNA and B1 are mixed with only enough CP to package one of them, polyU

VLPs are formed preferentially (15). By carrying out the second assembly step at a pH of 6, we have been able to include polyA in the comparison. Two new competition assays were performed, in which 2500–3500-nt polyU or 2500–3500-nt polyA competed against B1 (Fig. 5). (A-U complementarity precludes a direct competition between the two homopolymeric RNAs).

The electron micrographs (Fig. 5) show a small minority of capsids that are broken or malformed and some partial “secondary” shells that arise from the cationic N-termini of excess protein binding to the anionic outer surfaces of VLPs (25). Measurements taken of intact particles from electron micrographs of both sets of assembly experiments (Fig. 5) yielded size-distribution histograms from which the numbers of 22–23- and 28-nm particles (polyU versus B1) or 24–25- and 28-nm particles (polyA versus B1) could be compared. These analyses showed that polyU outcompetes B1 by a ratio of 3:2 at a pH of 6, whereas polyA is outcompeted by B1 by a factor of 3:1 (Fig. 6). The polyU results are consistent with those seen previously with a final assembly at a pH of 4.75. PolyA is outcompeted decisively even at adjusted assembly conditions that ensure its complete encapsidation by CCMV CP. One can speculate that

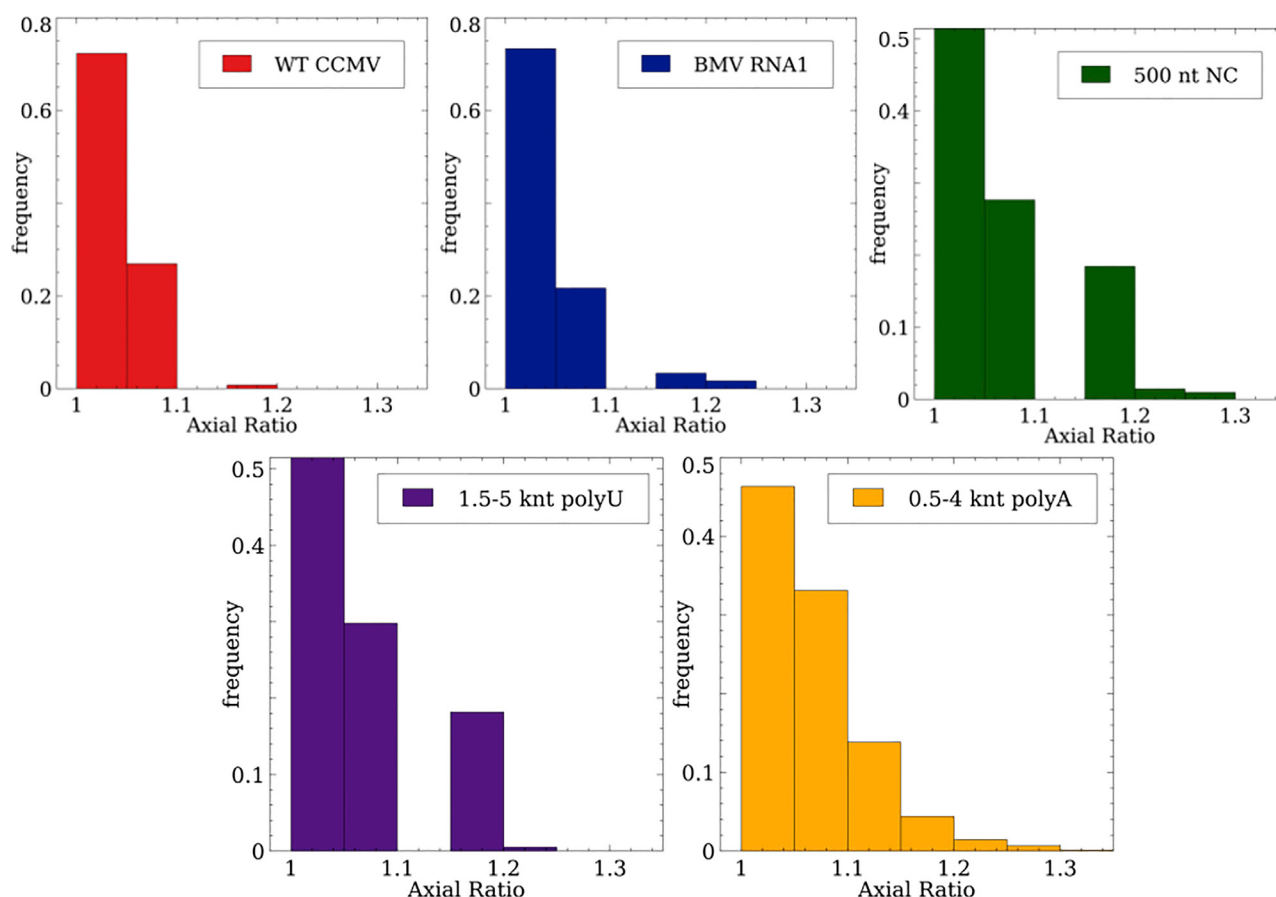


FIGURE 4 Histograms showing axial ratios for (top, left to right) WT CCMV, B1 VLPs, 500-nt normal-composition RNA VLPs, (bottom, left to right) an average of various lengths of polyU VLPs, and an average of various lengths of polyA VLPs. PolyA VLPs have both a very wide range of axial ratios as well as a continuous one, confirming irregularities in the VLP shapes seen by eye. To see this figure in color, go online.

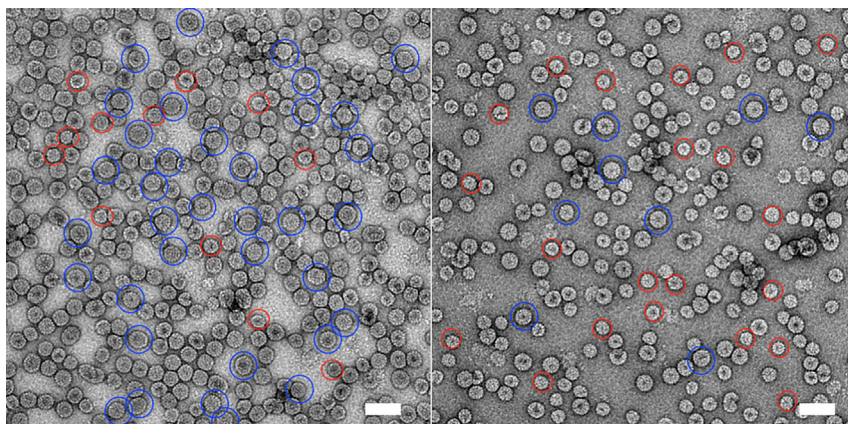


FIGURE 5 Electron micrographs of pH 6 competition experiments with 2500–3500 nt polyA and B1 (left) and 2500–3500 nt polyU and B1 (right). PolyA (red circles) is shown to be outcompeted by B1 (blue circles), whereas polyU (red circles) outcompetes B1 (blue circles). The smaller 24–25- or 22–23-nm particles in each image contain polyA and polyU, respectively, whereas the larger 28-nm particles contain B1. Scale bars represent 50 nm. To see this figure in color, go online.

this is one of the reasons the RNAs of viruses like CCMV have evolved to have pseudoknots (transfer RNA-like sequences) rather than polyA “tails” at their 3′ ends.

DISCUSSION

As emphasized in the [Introduction](#), CCMV is arguably the best studied viral self-assembly system. It was the first spherical virus to be reconstituted in vitro from purified components—CP and viral RNA (2). And subsequent work established that the viral genome could be replaced

by heterologous RNA (26) and even by nanocrystalline gold particles (27), anionic synthetic polymers (28,29), and nanoemulsion droplets (30), resulting in VLPs with the same capsid structure but of course no infectivity. This investigation completes a series of in vitro self-assembly experiments—all carried out with RNA and with the WT CP from CCMV—designed to explore the effect of RNA secondary structure by comparing and contrasting the VLPs formed by RNA molecules chosen to be as different as possible from one another. In particular, by looking on the one hand at normal-composition RNAs that involve comparable numbers of the four nt and hence a significant degree of secondary structure formation and at “single-letter” homopolymeric RNAs on the other hand, we are able to elucidate the role of intramolecular basepairing and of overall RNA size on the competitive binding of protein and on the relative yields and stabilities of capsids. PolyU, the RNA homopolymer with minimal base-stacking interactions, results in $T = 2$ -sized (versus WT $T = 3$) VLPs and yet is packaged more efficiently than viral (CCMV) RNA. Conversely, as reported above, polyA—the RNA homopolymer that undergoes a coil-helix transition at a low pH and hence is packageable only at a higher pH (> 5.8)—competes less well than viral RNA for CP and yet also results in $T = 2$ -sized VLPs.

These results lay the groundwork for a qualitative understanding of the roles of RNA secondary structure, flexibility, and size in determining the mode and extent of RNA-CP interaction and VLP formation. More explicitly, we have found the following:

- 1) PolyA can only be packaged into VLPs at pHs high enough so that it is in its coil (versus helix) state;
- 2) At these higher pHs (e.g., 6), polyA is less efficiently packaged than normal-composition RNA of a comparable nt length, which in turn is less efficiently packaged than polyU of the same length.
- 3) PolyA, like polyU, forms only $T = 2$ -sized VLPs, independent of length, ranging from RNAs considerably

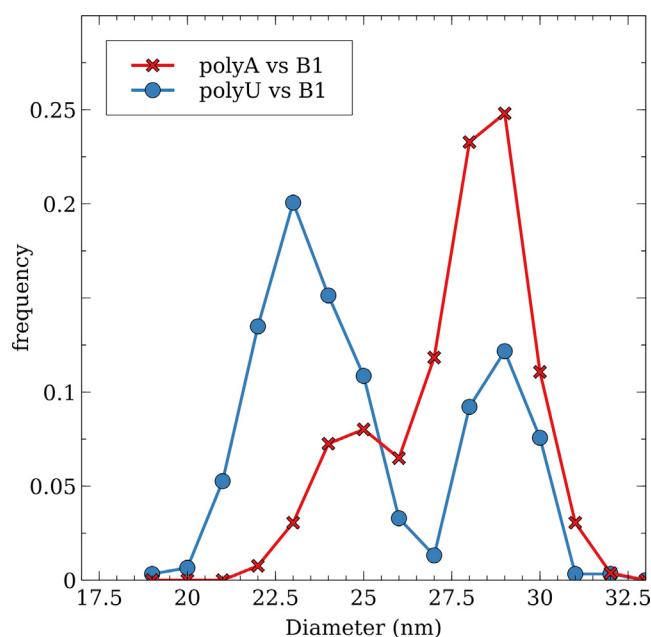


FIGURE 6 Size distributions in competition assembly experiments involving B1 RNA versus 3000-nt length polyU or polyA. The second step in the assemblies were carried out at pH 6. (Blue with circle symbols) PolyU outcompetes B1 for CP by a 3:2 ratio, consistent with the results seen before at pH 4.75 (15). (Red with cross symbols) B1 outcompetes the same length polyA for CP by a 3:1 ratio. To see this figure in color, go online.

shorter-than-viral length (~ 3000 nt) to those considerably longer, whereas the VLPs formed by normal-composition RNA are $T = 2$ -sized for shorter lengths and $T = 3$ for longer; and

- 4) PolyU, and likely polyA, forms multiplets involving $T = 2$ -sized VLPs when its length begins to approach twice the viral length, whereas the multiplets formed by “over-long” normal-composition RNAs continue to involve $T = 3$ VLPs.

The above findings can be understood in terms of several general concepts that are fundamental to the spontaneous formation of VLPs from RNA and CP. By keeping the CP the same and varying instead the nt length and composition of the RNA sequence, we are able to focus on RNA as the determinant of the competing structures and relative stabilities of VLPs. Several free-energy contributions are identified as the thermodynamic—versus kinetic—bases for these outcomes and are considered one-by-one below.

Preferred curvature of CP in the presence of different types of RNA

The notion of preferred curvature for a two-dimensional (2D) system has been most thoroughly developed in the context of many-component lipid bilayers in aqueous solutions and of surfactant monolayers in oil-water systems (31,32). There, the idea is relatively straightforward to define because it is an intrinsic property of the monolayer or bilayer (i.e., independent of the oil-water or water-water bulk phases that are separated by the 2D layer). Accordingly, one can write $\kappa(c - c_o)^2$ for the curvature energy per unit area of the layer, where κ is the 2D bending elastic constant (in units of energy), c its local curvature, and c_o its preferred curvature (associated with zero curvature energy cost)—with the values of κ and c_o determined completely by the nature and composition of the amphiphilic components of the monolayer and bilayer involved. Although this is true as well for the case of empty capsids, the bending modulus and the preferred curvature are not intrinsic properties of the protein for VLPs; they depend instead on the presence and particulars of the RNA with which the CP is interacting.

The fact that $T = 2$ -sized VLPs are the only ones observed in polyU and polyA assemblies, independent of the length of RNA, suggests that the values of c_o for CP interacting with both polyU and polyA are close to the reciprocal radius of a $T = 2$ capsid. This scenario would provide a natural thermodynamic basis for the exclusive appearance of $T = 2$ VLPs. (Although polyU has been shown to form multiplets of $T = 2$ capsids for lengths up to 10,000 nt, we were unable to obtain sufficient amounts of long (> 5000 nt) polyA to allow for the dilution and RNase digestion experiments needed to distinguish coincidentally touching

capsids from those that are parts of multiplets—two or more capsids sharing a single RNA molecule).

The situation for normal-composition RNA vis-à-vis preferred curvature is a bit more complicated. For RNAs with lengths from 500 to 1500 nt we find $T = 2$ -sized VLPs, for 2000–2500 nt we find a mix of $T = 2$ and $T = 3$, and for 3000 nt and all longer lengths—up to 12,000 nt—we find one or more $T = 3$ capsids per RNA. This is consistent with $c_o^{T=3}$ being the preferred curvature, but not so strongly favored over $c_o^{T=2}$ as to rule out $T = 2$ sizes for short (< 2500 nt) RNAs for which the confinement-energy cost (see below) is lower in smaller capsids. Note that there is also a kinetic advantage to the formation of smaller capsids. More explicitly, 500-nt-long RNAs, say, are not sufficiently long to bind enough CP to form a $T = 3$ -sized VLP nor even a $T = 2$ -sized capsid, so they have to wait, diffusing through solution, until they encounter a sufficient number of other CP-bound RNAs to form a stable capsid; this happens first for the $T = 2$ -sized VLPs (8). This also accounts for why short RNAs can compete effectively for CP (see below) against viral length RNAs. But the fact that multiplets of $T = 3$ -sized capsids are observed for RNAs two or more times longer than viral length (~ 3000 nt)—as opposed to larger VLPs appearing to accommodate the “extra” RNA—confirms that $T = 3$ is the preferred curvature for CCMV CP interacting with normal-composition RNA and therefore the thermodynamic VLP product.

Relative protein-binding ability/affinity of different types of RNA

Separate from the issue of preferred protein-shell curvature (CP-CP interactions) in the presence of RNA is the issue of the strength and nature of protein-RNA interactions. Very little is known about how CP binds RNA except in the case of a few RNA viruses with well-studied “packaging signals” (33) in which a precise primary sequence and its associated secondary structure—e.g., a particular stem loop—can be shown to have a singularly strong affinity for CP. In the case of MS2 (34) and tobacco mosaic virus (35), for example, specific packaging signals are the sites where CP binds; upon initiation of nucleation, subsequent CPs join the critical “embryo” from solution. For tobacco mosaic virus, all of the RNA is strongly interacting with the nucleating capsid, whereas for MS2, a succession of RNA sites with high affinity for CP are involved in organizing the genome within the capsid interior. In cases like CCMV, however, in which the dominant interaction between RNA and CP is the nonspecific electrostatic attraction between anionic backbone phosphates and cationic N-terminal residues, proteins saturate the RNA all along its length before nucleation of the capsid occurs, with the nonspecific interaction augmented by secondary-structure-dependent contributions that favor one RNA over another when two or more compete for CP. This scenario accounts for the

relative packaging efficiencies of the different RNAs we have studied, in which the same CP—CCMV CP—is involved in all of the self-assembly reactions. At a neutral pH, the RNAs are saturated by CP, and the high concentration of bound CP results in capsid formation as soon as the pH is lowered and CP-CP interactions are enhanced. The T numbers of the resulting capsids are determined by the preferred curvature associated with the RNA-CP interaction— $T = 2$ for homopolymeric RNA and $T = 3$ for heteropolymeric. Finally, in the head-to-head competition experiments, the relative packaging efficiencies are decided by the slightly different affinities of CP for local-sequence and secondary structure as well as by different RNA sizes (see below), again providing a thermodynamic basis for one RNA being preferred over the other for incorporation into VLPs.

Effect of RNA confinement (e.g., size)

The first two effects discussed above are local, in the sense that they involve either CP-CP or CP-RNA interactions on the scale of a few nm, roughly the size of the proteins and of their RNA “footprints.” In other words, CPs at a neutral pH bind RNA and then interact with other bound CPs in a way that is determined exclusively by local primary/secondary RNA and CP structure. But when an ordered shell is nucleating at a lower pH, the free-energy of the capsid embryo will be determined by the extent to which the overall RNA is able to interact maximally with the inside of the shell (which is positively charged) while having its configurational entropy minimally affected—and this balance depends on the large-scale size and compressibility of the RNA. From the fact that normal-composition RNA of viral length (3000 nt) is roughly the same size as the $T = 3$ capsid (13), it is clear that $T = 3$ VLPs ensure not just the local demands of CPs (i.e., $c \approx c_o^{T=3}$) but also the larger-scale requirements on RNA for maximal interaction with the inner capsid surface and minimal loss of configurational entropy. And, relatedly, the fact that $T = 2$ -sized VLPs arise in the case of shorter normal-composition RNAs—below 2000 nt in length—confirms that there is a significant free-energy cost associated with the overall size of the RNA not being big enough to be consistent with the $T = 3$ capsid. More explicitly, $T = 3$ curvature might be preferred insofar as it results in lower energy for the CP-CP interactions with RNA on a local (few nm) scale but not enough to offset the larger-scale (10s of nm) reorganization costs of the RNA. A similar effect likely underlies the progression from $T = 2$ to $T = 3$ VLPs in the case of progressively larger gold nanoparticles (27) and longer poly(styrenesulfonate) (PSS) molecules (28,29).

These RNA size/confinement effects provide then still another thermodynamic basis for $T = 2$ -sized VLP formation, separate from the $c_o^{T=2}$ versus $c_o^{T=3}$ preferred curvature effects. They also explain why polyA is less efficiently

packaged than both polyU and normal-composition RNA, even in its coil state at a high pH, because it pays the biggest confinement free energy cost because of its bigger size (see Fig. 7) and its larger persistence length (36,37), both resulting in a larger overall/large-scale packaging free-energy cost (38,39); for the same reasons, these VLPs involve the broadest distribution of sizes and asymmetries. Relatedly, one can understand why, for example, polyA in its low-pH state does not form VLPs; the helical order is sufficient to result in a significant amount of double-stranded RNA, whose incompressibility prohibits confinement in a spherical capsid, even a larger $T = 3$ one. And yet the coil disorder—involving single-stranded regions between duplexes (40)—is sufficient to interfere with confinement in tubular capsids of CCMV CP, such as those observed for fully complementary double-stranded DNA (41) in which bundles of long nucleic acid duplex can interact optimally with the tube interior without any major loss of configurational entropy.

The three thermodynamic ingredients outlined above are consistent with a three-parameter theory treating viral self-assembly that has been proposed by Zlotnick et al. (5), in which CP-CP interactions are accounted for by an energy ω (not including explicit angle-dependent, preferred curvature, contributions), CP-RNA interactions by an association constant K_{NA} (in inverse molar units), and “the work of fitting the RNA into the constraints of a growing capsid”—our size and confinement effects—by a free energy α . Phenomenological phase diagrams of nucleocapsid formation are then worked out for varying combinations of ω , K_{NA} , and α , with suggestions for how these parameters can be related to particular nucleic acid and CP combinations in different viruses and VLPs. In our present investigations, we have used a single CP (CCMV) and a “rationally

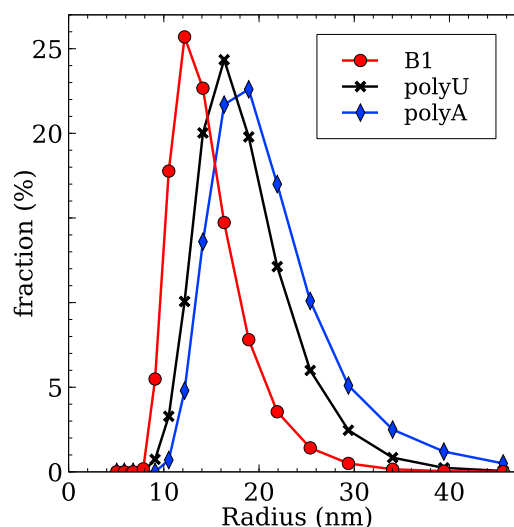


FIGURE 7 Volume-weighted distributions of hydrodynamic radii of B1 RNA (red with circle symbols), ~3000-nt long polyU (black with cross symbols), and polyA (blue with diamond symbols) at a pH of 8 and 4°C, determined by DLS. To see this figure in color, go online.

designed” series of nucleic acids (normal-composition—effectively “branched”—RNA (19) of different lengths and single-letter—effectively linear—RNAs involving either weak (polyU) or strong (polyA) base-stacking interactions) to highlight the different VLP curvatures and packaging efficiencies associated with each type of RNA.

All of the effects discussed above are implicitly based on self-assembly products corresponding to thermodynamic equilibrium. More explicitly, we argue the following: single-letter RNAs give rise to a strong energetic preference for $T = 2$ curvature and normal-composition RNAs to $T = 3$; $T = 2$ VLPs arise for short heteropolymer RNAs because of the importance of size/confinement contributions to the free energy of packaging; and differences in CP-RNA binding affinities and in RNA size and flexibility account for the relative efficiencies of RNAs competing to be packaged by CP. A similar scenario arises for CCMV CP VLP formation associated with the synthetic anionic polymer PSS, in which an evolution from $T = 2$ -sized to $T = 3$ structures is observed upon increasing the molecular weight from thousands to millions, again reflecting an interplay between RNA confinement and CP curvature free energies. It is also possible that $T = 2$ VLPs arise for shorter normal-composition RNAs and PSS because $T = 2$ -sized VLPs have the kinetic advantage of forming at shorter times because of their not needing to wait for the binding of additional proteins or for association with other protein-bound RNAs. We believe, however, that the VLP curvatures and yields observed for a wide variety of homo- and heteropolymeric RNA molecules—all involving the same CP—are most simply and consistently accounted for in terms of the several different free energy contributions we have outlined above. In particular, these findings for polyAVLP formation, in conjunction with our earlier studies with polyU and normal-composition RNAs, suggest strongly that the trade-off between curvature-preference and size-confinement effects is the dominant factor in determining the relative yields of thermodynamic products, in particular that $T = 2$ and $T = 3$ capsids are associated with the absence and presence of RNA secondary structure.

SUPPORTING MATERIAL

Supporting Material can be found online at <https://doi.org/10.1016/j.bpj.2019.08.012>.

AUTHOR CONTRIBUTIONS

All authors were involved in designing the experiments. A.R.T., C.B., and A.L.D.-M. performed the experiments and analyzed the data. A.R.T., C.B., W.M.G., and C.M.K. wrote the manuscript together.

ACKNOWLEDGMENTS

We thank Dr. Feng Guo for providing the T7 RNA polymerase, Dr. A. L. N. Rao for providing the complementary DNA used to generate B1 RNA, and

Dr. Hong Zhou, Dr. Ivo Atanasov, and Dr. Yanxiang Cui for helpful discussions related to EM. Electron microscope images were acquired at the California NanoSystems Institute Electron Imaging Center for NanoMachines.

This research was supported by grant MCB1716925 from the National Science Foundation.

REFERENCES

1. Fraenkel-Conrat, H., and R. C. Williams. 1955. Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components. *Proc. Natl. Acad. Sci. USA*. 41:690–698.
2. Bancroft, J. B., and E. Hiebert. 1967. Formation of an infectious nucleoprotein from protein and nucleic acid isolated from a small spherical virus. *Virology*. 32:354–356.
3. Borodavka, A., R. Tuma, and P. G. Stockley. 2012. Evidence that viral RNAs have evolved for efficient, two-stage packaging. *Proc. Natl. Acad. Sci. USA*. 109:15769–15774.
4. Cadena-Nava, R. D., M. Comas-Garcia, ..., W. M. Gelbart. 2019. Correction for Cadena-Nava et al., “self-assembly of viral capsid protein and RNA molecules of different sizes: requirement for a specific high protein/RNA mass ratio”. *J. Virol.* 93:e01939-18.
5. Zlotnick, A., J. Z. Porterfield, and J. C. Wang. 2013. To build a virus on a nucleic acid substrate. *Biophys. J.* 104:1595–1604.
6. Garmann, R. F., M. Comas-Garcia, ..., W. M. Gelbart. 2016. Physical principles in the self-assembly of a simple spherical virus. *Acc. Chem. Res.* 49:48–55.
7. Zlotnick, A., R. Aldrich, ..., M. J. Young. 2000. Mechanism of capsid assembly for an icosahedral plant virus. *Virology*. 277:450–456.
8. Comas-Garcia, M., R. F. Garmann, ..., W. M. Gelbart. 2014. Characterization of viral capsid protein self-assembly around short single-stranded RNA. *J. Phys. Chem. B*. 118:7510–7519.
9. Caspar, D. L., and A. Klug. 1962. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27:1–24.
10. Krol, M. A., N. H. Olson, ..., P. Ahlquist. 1999. RNA-controlled polymorphism in the in vivo assembly of 180-subunit and 120-subunit virions from a single capsid protein. *Proc. Natl. Acad. Sci. USA*. 96:13650–13655.
11. Speir, J. A., S. Munshi, ..., J. E. Johnson. 1995. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure*. 3:63–78.
12. Comas-Garcia, M., R. D. Cadena-Nava, ..., W. M. Gelbart. 2012. In vitro quantification of the relative packaging efficiencies of single-stranded RNA molecules by viral capsid protein. *J. Virol.* 86:12271–12282.
13. Gopal, A., Z. H. Zhou, ..., W. M. Gelbart. 2012. Visualizing large RNA molecules in solution. *RNA*. 18:284–299.
14. Garmann, R. F., A. Gopal, ..., S. C. Harvey. 2015. Visualizing the global secondary structure of a viral RNA genome with cryo-electron microscopy. *RNA*. 21:877–886.
15. Beren, C., L. L. Dreesens, ..., W. M. Gelbart. 2017. The effect of RNA secondary structure on the self-assembly of viral capsids. *Biophys. J.* 113:339–347.
16. Richards, E., C. Flessel, and J. Fresco. 1963. Polynucleotides. VI. Molecular properties and conformation of polyribouridylic acid. *Biopolymers*. 1:431–446.
17. Friedman, R. A., and B. Honig. 1995. A free energy analysis of nucleic acid base stacking in aqueous solution. *Biophys. J.* 69:1528–1535.
18. Fang, L. T., W. M. Gelbart, and A. Ben-Shaul. 2011. The size of RNA as an ideal branched polymer. *J. Chem. Phys.* 135:155105.
19. Singaram, S. W., R. F. Garmann, ..., A. Ben-Shaul. 2015. Role of RNA branchedness in the competition for viral capsid proteins. *J. Phys. Chem. B*. 119:13991–14002.

20. Fresco, J., and P. Doty. 1957. Polynucleotides. I. Molecular properties and configurations of polyriboadenylic acid in solution. *J. Am. Chem. Soc.* 79:3928–3929.
21. Rich, A., D. R. Davies, ..., J. D. Watson. 1961. The molecular structure of polyadenylic acid. *J. Mol. Biol.* 3:71–86.
22. Annamalai, P., and A. L. Rao. 2005. Dispensability of 3' tRNA-like sequence for packaging cowpea chlorotic mottle virus genomic RNAs. *Virology*. 332:650–658.
23. Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. 9:671–675.
24. Nguyen, H. D., V. S. Reddy, and C. L. Brooks, III. 2009. Invariant polymorphism in virus capsid assembly. *J. Am. Chem. Soc.* 131:2606–2614.
25. Prinsen, P., P. van der Schoot, ..., C. M. Knobler. 2010. Multishell structures of virus coat proteins. *J. Phys. Chem. B*. 114:5522–5533.
26. Hiebert, E., J. B. Bancroft, and C. E. Bracker. 1968. The assembly in vitro of some small spherical viruses, hybrid viruses, and other nucleoproteins. *Virology*. 34:492–508.
27. Chen, C., E. S. Kwak, ..., B. Dragnea. 2005. Packaging of gold particles in viral capsids. *J. Nanosci. Nanotechnol.* 5:2029–2033.
28. Hu, Y., R. Zandi, ..., W. M. Gelbart. 2008. Packaging of a polymer by a viral capsid: the interplay between polymer length and capsid size. *Biophys. J.* 94:1428–1436.
29. Cadena-Nava, R. D., Y. Hu, ..., W. M. Gelbart. 2011. Exploiting fluorescent polymers to probe the self-assembly of virus-like particles. *J. Phys. Chem. B*. 115:2386–2391.
30. Chang, C. B., C. M. Knobler, ..., T. G. Mason. 2008. Curvature dependence of viral protein structures on encapsidated nanoemulsion droplets. *ACS Nano*. 2:281–286.
31. Gelbart, W., A. Ben-Shaul, and D. Roux. 1994. *Micelles, Membranes, Microemulsions, and Monolayers*. Springer, New York.
32. Marsh, D. 2006. Elastic curvature constants of lipid monolayers and bilayers. *Chem. Phys. Lipids*. 144:146–159.
33. Stockley, P. G., R. Twarock, ..., R. Tuma. 2013. Packaging signals in single-stranded RNA viruses: nature's alternative to a purely electrostatic assembly mechanism. *J. Biol. Phys.* 39:277–287.
34. Rolfsson, Ó., S. Middleton, ..., P. G. Stockley. 2016. Direct evidence for packaging signal-mediated assembly of bacteriophage MS2. *J. Mol. Biol.* 428:431–448.
35. Sleat, D. E., P. C. Turner, ..., T. M. Wilson. 1986. Packaging of recombinant RNA molecules into pseudovirus particles directed by the origin-of-assembly sequence from tobacco mosaic virus RNA. *Virology*. 155:299–308.
36. Seol, Y., G. M. Skinner, and K. Visscher. 2004. Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.* 93:118102.
37. Seol, Y., G. M. Skinner, ..., A. Halperin. 2007. Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. *Phys. Rev. Lett.* 98:158103.
38. Perlmutter, J. D., C. Qiao, and M. F. Hagan. 2013. Viral genome structures are optimal for capsid assembly. *eLife*. 2:e00632.
39. Li, S., G. Erdemci-Tandogan, ..., R. Zandi. 2018. The effect of RNA stiffness on the self-assembly of virus particles. *J. Phys. Condens. Matter*. 30:044002.
40. Holcomb, D., and J. I. Tinoco. 1965. Conformation of polyriboadenylic acid: pH and temperature dependence. *Biopolymers*. 3:121–133.
41. Mukherjee, S., C. M. Pfeifer, ..., A. Zlotnick. 2006. Redirecting the coat protein of a spherical virus to assemble into tubular nanostructures. *J. Am. Chem. Soc.* 128:2538–2539.