Chapter 3
The In Vitro Packaging of "Overlong" RNA by Spherical Virus-Like Particles

Ana Luisa Duran-Meza, Abigail G. Chapman, Cheylene R. Tanimoto, Charles M. Knobler, and William M. Gelbart

Abstract Of the myriad viruses, very few have been shown to be capable of self-assembly in vitro from purified components into infectious virus particles. One of these is Cowpea Chlorotic Mottle Virus (CCMV), an unenveloped spherical plant virus whose capsid self-assembles around its RNA genome without a packaging signal. While heterologous RNA, not just cognate viral RNA, can be packaged into individual CCMV virus-like particles (VLPs), the RNA needs to fall within a certain range of lengths. If it is too short, it is packaged into particles smaller than wild type, or with two or more RNAs per capsid. If the RNA is too long, multiple capsids assemble around one RNA, and the RNA associated with these multiplet structures is not as RNase resistant. Further, as shown in the present work, 4200 nt appears to be the limiting length of RNA that can be packaged into single RNase-resistant CCMV VLPs. We explore the extent to which "overlong" RNA can be packaged more efficiently upon the addition of spermine, a polyvalent cation whose increasing concentration has been shown to compactify RNA. Finally, we show that the capsid protein of Brome Mosaic Virus (BMV), a bromovirus closely related to CCMV, also gives rise to multiplets when it is self-assembled with the same "overlong" RNA constructs, but with different distributions of multiplets.

Keywords Virus-like particles · In vitro assembly · Capsid multiplets · RNase-resistant capsid packaging · Spermine compaction

Introduction

The capacity for in vitro self-assembly of viruses from their purified constituents—shown by the icosahedral plant viruses CCMV (cowpea chlorotic mottle virus) and BMV (brome mosaic virus), bacteriophage MS2, and the rod-like TMV (Tobacco Mosaic Virus)—makes them highly promising platforms for biomedical applications,

A. L. Duran-Meza · A. G. Chapman · C. R. Tanimoto · C. M. Knobler · W. M. Gelbart (✉)
Department of Chemistry and Biochemistry, University of California Los Angeles, 607 Charles E. Young Drive East, Los Angeles, CA 90095, USA
e-mail: gelbart@chem.ucla.edu

M. Comas-Garcia and S. Rosales-Mendoza (eds.), *Physical Virology*, Springer Series in Biophysics 24, https://doi.org/10.1007/978-3-031-36815-8_3

among them vaccines and gene delivery [1]. Their utility depends on the ability of the capsid proteins of these ssRNA viruses to assemble in vitro around heterologous RNAs, forming virus-like particles (VLPs) of high stability, which: protect the RNA against attack by RNase; are capable of delivering their contents to the ribosomal (translational) machinery; and are free of biological contaminants that might arise from assembly in bacterial or yeast cells. Once formed, the VLPs may be functionalized by chemically linking proteins to them that recognize specific cells and enhance uptake [2, 3].

While the packaging capacity of these VLPs is not a major issue in determining their suitability for delivering cargos of small molecules such as miRNAs or mRNAs that code for short protein sequences, the maximum length of RNA that can be packaged in spherical capsids is a limiting factor to their employment for delivering constructs such as self-replicating RNAs (replicons) that code for a replicase and a gene of interest [4], because the strongly preferred curvature (radius) of the capsid protein dictates the volume available to packaged RNA. This is not a limitation for TMV whose genomic RNA is 6400 nt long, and whose cylindrical VLPs can accommodate any length of heterologous RNA because the capsid length is simply proportional to RNA length and its curvature is independent of its length. But faithful and efficient TMV VLP assembly requires that the RNA includes an origin-of-assembly (OAS) sequence to initiate packaging [5]. While a terminal TLS-like sequence has been identified for the packaging of BMV RNAs [6] a packaging signal has not been found for CCMV assembly [7] and heterologous RNAs lacking a TLS have been packaged without difficulty [8]. The packaging of RNA in capsids formed in vitro from MS2 CP (in the absence of the maturation protein that breaks symmetry and ensures infectivity in the wt virus) depends on the presence of a *pac* site, a 19-base stem loop that initiates packaging [9], and in vitro packaging of heterologous RNA in MS2 as long as 3000 nt requires the presence of multiple pac sites [10].

CCMV and BMV have capsids with Caspar-Klug triangulation numbers $T = 3$, composed of 180 proteins. While all RNA viruses with this same capsid structure have essentially identical outer diameters of ~30 nm, their nucleotide content varies significantly. Wild-type CCMV and BMV capsids contain ~3000 nt of ssRNA. In contrast, Cucumber Mosaic Virus (CMV), Bacteriophage MS2, Turnip Crinkle Virus (TCV), Turnip Yellow Mosaic Virus (TYMV), and NoroVirus (NoV), all with $T = 3$ capsids and approximately 30-nm diameters, have genomes with lengths 3300, 3569, 4034, 6400 and 7600 nt, respectively. The remarkably long genomes for TYMV and NoV are attributable to the presence in their capsids of agents that act to condense the RNA, such as the polyamine spermidine in the case of TYMV [11] and small basic proteins in the case of NoV [12]. Their RNAs otherwise appear to lack the high degree of branching that results in compaction [13, 14].

In the case of CCMV and BMV the major driving force for the packaging of RNA is the electrostatic interaction between the negatively charged phosphate backbone and the positively-charged N-terminus of the CP. Direct evidence for this non-specific interaction is found in self-assembly studies with mutant CCMV CPs in which the number of N-terminal cationic residues is successively decreased and shown to lead to decreasing lengths of RNA being packaged [15]. Further evidence comes from

the cryoelectron microscopy structural study by Beren et al. [16] of BMV virions in which the RNA is shown in high-resolution reconstructions to be highly disordered, in contrast to what is found, for example, in wt MS2 where the multiple packaging signals and interaction of the RNA ends with the maturation protein enforce a single well-defined structure [17]. The lack of any specific structure for the packaged RNA in BMV (and, presumably, CCMV) accounts for their ability to package (each individually) a broad range of RNAs—notably, the three very different molecules making up their genome, as well as heterologous RNAs.

It would appear then, from the standpoint solely of volume constraints, that CCMV and BMV should be capable of packaging RNAs with lengths well in excess of the 3000 nt found in the wild type. What then is their packaging limit? One might assume that this could be simply determined by a series of experiments in which assemblies with increasingly long RNAs are carried out until a limit is reached beyond which the RNAs are not packaged. However, it has been shown that attempts to package RNA longer than 3000 nt by CCMV CP leads to "multiplet" structures in which a single RNA is shared by two or more capsids [8]. Multiplets formed by CCMV CP and several heterologous "overlong" RNAs have also been observed in assemblies of homopolymeric polyU and CCMV CP [18], for a conjugated polyelectrolyte and CCMV CP [19], and for packaging by SV40 CP of RNA [20] and polystyrene sulfonate [21]. The existence of doublet structures has also been found by Elrad and Hagan [22] in molecular dynamics simulations of the packaging of long polymers by capsid protein in their coarse-grained modelling of virus-particle-like assembly [23].

To determine the packaging limit by experiment therefore requires analyses of the multiplet distributions as RNA length is increased. This is the strategy we employ here to determine the efficiency of packaging and maximum length of RNA that can be packaged by CCMV and BMV. In previous studies we examined the formation of multiplets in CCMV assemblies with CP purified from virus-infected plants, and in the present work we find the same multiplet distribution using CP expressed in E. coli (a protein source that is more readily scaled up). Further, by analyzing the RNA extracted from assembly mixes before and after RNase treatment, using RNA molecules with lengths up to 6400 nt, we observe that CCMV CP is unable to package RNA longer than 4200 nt in single capsids. The same in vitro-packaging-length limit is found for the CP from the closely related bromovirus BMV, but with a different distribution of singlets, doublets, triplets, and quadruplets.

If the amount of RNA that can be packaged is set by the volume and effective charge of the RNA, then the addition of compactifying (e.g., polyvalent cationic) agents would be expected to increase the packaging limit. Indeed, the compaction of TYMV RNA by polyamines has been demonstrated by their effect on the sedimentation rate, which is inversely proportional to the RNA hydrodynamic radius. The addition of 0.1 mM spermine resulted in a 40% decrease in the radius while addition of bis(3-aminopropyl amine) led to a 25% decrease [24]. More recently, fluorescence correlation spectroscopy determinations [25] of the effect of the addition of 1 mM spermine to MS2 RNA showed a 50% decrease in its hydrodynamic radius. Finally, recent dynamic light scattering measurements on a broad range of lengths

and sequences of RNA molecules, as a function of increasing spermine concentrations, have established a continuous (saturating at 50%) decrease in hydrodynamic radius in all cases [26]. Consistent with these results we report here the extent to which RNA compaction by polyvalent cations affects the efficiency of packaging of "overlong" RNA by CCMV and BMV capsid protein. We find that both the number of singlet particles and the packaged-length limit can indeed be enhanced by spermine. But the effect is small, and increasing the polyamine concentration eventually results in RNA aggregation, which limits the extent to which spermine can enhance the relative number of singlets or the length of RNA packaged into RNase-resistant singlets.

Results

To generate a wide range of lengths and sequences of RNA molecules, BMV RNA1 (3234 nt), NOV-EYFP (4196 nt), NOV-R.Luc (4413 nt), NOV-STING* (4638 nt), and TMV (6395 nt) were prepared by in vitro transcription of T7 DNA plasmids, and SIN-19 (8985 nt) and FL-SIN (11,703 nt) from SP6 plasmids (Thermo Fisher, USA), and all of the RNAs were purified with an RNEasy Mini Kit (Qiagen, DEU). Recombinant capsid proteins were grown in ampicillin-chloramphenicol-resistant cells of E. coli strain Rosetta 2 BL21, and in vitro reconstitutions of VLPs were carried out using our published protocols [8]. RNA was extracted from VLPs with a QIAamp Viral Mini Kit (Qiagen DEU) following the manufacturer's specifications. To assess VLPs for RNA protection they were mixed with RNaseA at a ratio of 0.1 g RNase A:1 g RNA, and incubated at 4 °C for 1 h. Digestion of RNA was stopped by the addition of RNase inhibitor and the sample was washed through a 100 kDa MW-cutoff Amicon filter to purify the remaining RNase-resistant VLPs.

Multiplet VLP Products Are Observed for RNAs Longer Than 3000 nt

Using a Tecnai G2 TF20 electron microscope (FEI, USA), assembly mix samples were prepared and imaged with negative stain as reported elsewhere [8, 18]. Electron micrographs were manually analyzed to count the number of singlets, doublets, triplets, and higher-order multiplets in each of the images acquired; ImageJ (US National Institutes of Health) was used to measure the geometric mean of orthogonal measurements of the diameter of the particles.

First we confirmed that bacterial-expressed capsid proteins give rise to VLP multiplet distributions similar to those found earlier [8] for infected-plant-derived proteins, when self-assembled with RNAs of different lengths. When 3200 nt RNA is packaged by CCMV CP, mostly singlet VLPs are formed (see Fig. 3.1a). Assemblies with 4200 nt RNA (Fig. 3.1b) contain some multiplet particles—mostly doublets—where one RNA is shared by two or more capsids. Packaging of 6400 nt-RNA (Fig. 3.1c) leads to a mixture of singlet, doublet, triplet, and higher-order multiplet particles, but again mostly doublets. The histogram of sizes of singlets in Fig. 3.1d is consistent with that measured in the earlier experiments conducted with plant-derived CCMV

CP and heterologous RNA [8]. The size distributions of singlet particles—measured for each of the multiplet assemblies associated with the three different RNA lengths—have widths of several nm, centered around 29–31 nm. Slightly larger singlet sizes are observed (Fig. 3.1d) for the longer-RNA (4200 and 6400 nt, versus 3200 nt) assemblies.

This is consistent with the fact that the longer RNA samples include a significant number of shorter, prematurely-terminated, transcripts and degradation products, as seen in Fig. 3.2. More explicitly, the 4200 and 6400 nt samples include RNA lengths in the range 3200–4200 nt that are short enough to be self-assembled into single T = 3 capsids but that result in capsids that are less stable and less ordered—and hence somewhat *larger*—than those for 3200 nt-RNA, the length preferred by CCMV CP.

As suggested by studies that demonstrated no structural difference between recombinant CP expressed in E. coli and wt CP derived from plants [27], we observe that the measured multiplet distribution (see Fig. 3.3) is the same as reported earlier [8] for these RNA + CCMV CP assembly mixes, even though plant-derived CP is used in the latter case and recombinant CP in the former. For a 3200 nt RNA, 85% of the particles are singlets; for 4200 nt RNA, 65% of particles are singlets, and the rest are mostly doublets; and for 6400 nt RNA, 32% appear to be singlets while 56% percent are doublet particles. Upon serial dilution of the assembly mix, the multiplet

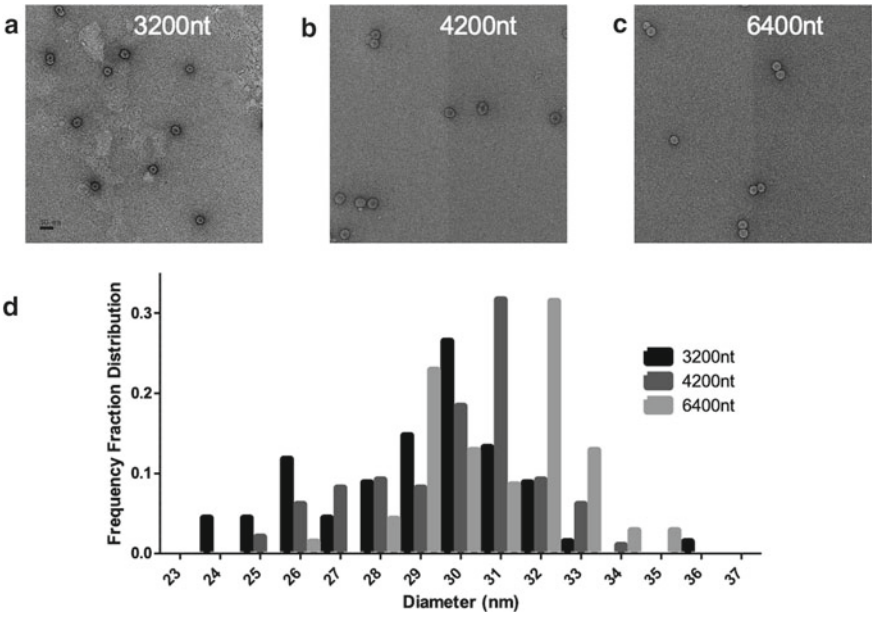

Fig. 3.1 Characterization of CCMV VLPs assembled with different lengths of RNA. Negative-stain electron micrographs of self-assembly mixes involving CCMV CP and: ~3200 nt-long RNA (**a**), ~4200 nt-long RNA (**b**), and ~6400 nt-long RNA (**c**). **d** shows the singlet capsid-size-distribution plots measured for each of these self-assemblies before RNase treatment

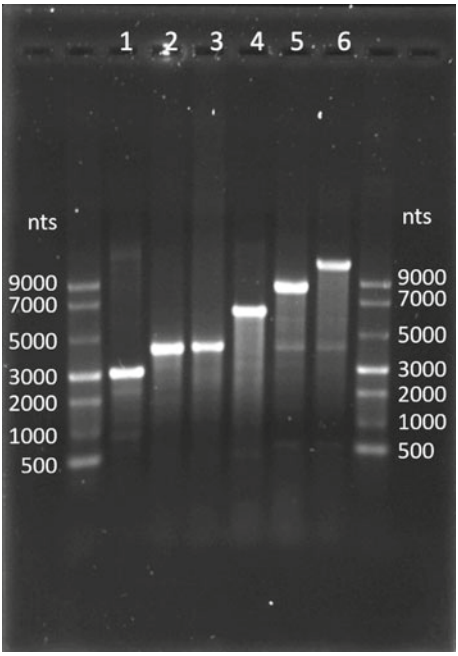

Fig. 3.2 In vitro transcribed RNAs of increasing length. Agarose gel (0.83%) showing different-length RNA transcript products. Lane 1: 3234 nt RNA, Lane 2: 4413 nt RNA, Lane 3: 4638 nt RNA, Lane 4: 6395 nt RNA, Lane 5: 8985 nt RNA, Lane 6: 11703 nt RNA. 1 μg of RNA was mixed with RNA loading dye and denatured by heating to 65 °C at a rate of 1 °C per sec, holding at 65 °C for 10 min, then cooling to 4 °C at a rate of 1 °C per sec. The gel was run at 100 V for 1.5 h, and visualized with GelRed

particles continue to be observed and the distribution remains unaltered, indicating that they are not a result of particle crowding or aggregation. Assembly of overlong RNAs into VLPs with *BMV* CP also results in multiplets, but gives rise to distributions that contain significantly larger fractions of higher-order multiplets than in the case of CCMV.

Importantly, as discussed below, most of the singlets that appear in the 4200 nt-RNA assemblies, and all of the singlets observed in the 6400 nt-RNA case, contain RNA molecules shorter than 4200 and 6400 nt, respectively. These shorter RNAs arise as premature transcripts and hydrolysis degradation products of the in vitro transcribed RNAs. As seen in Fig. 3.2, an agarose gel analysis of RNA molecules with lengths ranging from ~3200 to ~12000 nt, each lane shows a strong band (the corresponding full-length transcript) accompanied by broad streaks associated with faster-running, shorter, premature transcripts and degradation products. As discussed below in the context of Figs. 3.4 and 3.5, it is these shorter RNAs—up to lengths of 4200 nt—that are packaged into singlets when transcripts of RNAs longer than 3200 nt are mixed with CCMV CP under assembly conditions.

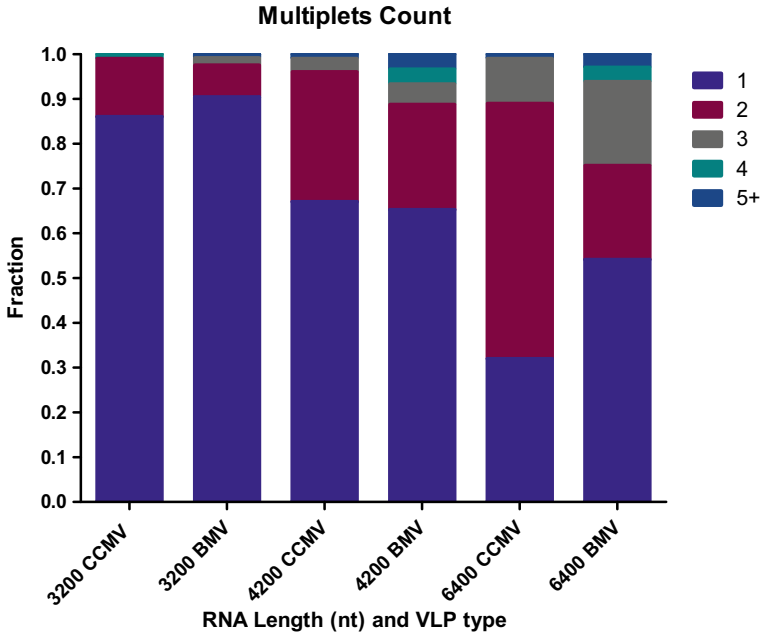

Fig. 3.3 Characterizing CCMV and BMV VLP particles by negative-stain EM. Counts of the frequency of singlets, doublets, triplets, quadruplets and larger particles in samples of CCMV and BMV VLP assemblies (before RNase treatment)

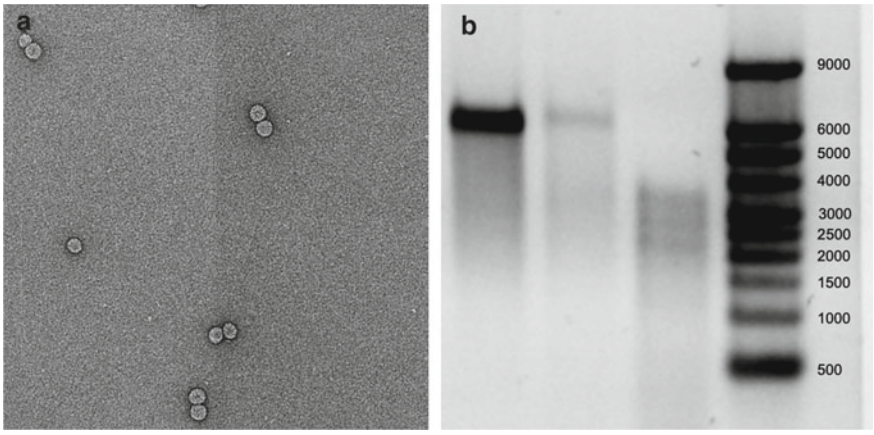

Fig. 3.4 6400 nt RNA in CCMV VLPs. **a** Negative-stain electron micrograph of CCMV VLPs assembled with in vitro transcribed TMV-(6400 nt-)RNA, showing singlet, doublet and triplet particles. **b** 0.83% agarose gel. Lane 1: in vitro transcribed 6400 nt RNA. Lane 2: RNA extracted from CCMV-VLPs assembled with 6400 nt RNA. Lane 3: RNA extracted from RNase-treated CCMV-VLPs assembled from 6400 nt RNA. Lane 4: RNA Ladder

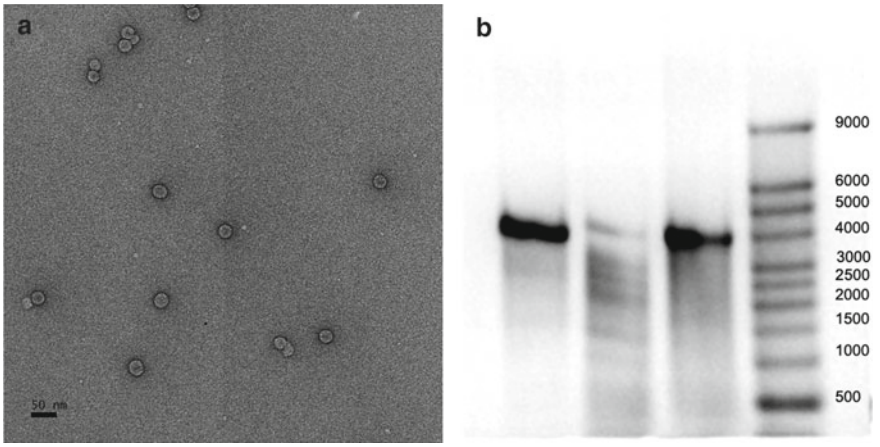

Fig. 3.5 4200 nt RNA in CCMV VLPs. **a** Negative-stain electron micrograph of 4200 nt-RNA CCMV-VLPs, showing singlet, doublet and triplet particles. **b** 0.83% agarose gel. Lane 1: RNA extracted from CCMV-VLPs assembled with 4200 nt-RNA. Lane 2: RNA extracted from RNase-treated CCMV-VLPs assembled with 4200 nt-RNA. Lane 3: in vitro transcribed 4200 nt RNA, Lane 4: RNA Ladder

After RNase A Treatment, Only Singlet VLPs Are Observed

When treated with RNase, all multiplet VLPs become sets of singlets, as evidenced by electron micrographs (not shown), i.e., no doublets, triplets, or quadruplets remain, from which we conclude that the multiplets are not RNase resistant: the portions of the RNA connecting the capsids are not protected. However, the singlets remaining after disassembly of multiplets are RNAase resistant (see Fig. 3.4b and discussion below), evidence that after disassembly of multiplets the capsids are intact.

RNAs Longer Than 4200 nt Cannot Be Packaged into RNase-Resistant Singlets

To show in particular that 6400 nt RNA cannot be packaged into singlet CCMV VLPs, we extracted RNA from samples of 6400 nt-RNA CCMV-VLPs before and after their treatment with RNase. Recall that assemblies of this length of RNA with CCMV CP result predominantly in doublets, but also in a significant number of singlet VLPs; see Figs. 3.1c and 3.4a. Lane 1 of Fig. 3.4b, loaded with in vitro-transcribed 6400 nt RNA, features a strong band where expected (see RNA ladder in lane 4), but also a smear of shorter lengths down to ~2000 nt. The RNA extracted from VLPs assembled with this in vitro transcribed RNA and CCMV CP is run in lane 2 and shows a *weak* band at 6400 nt along with, again, a smear of shorter lengths down to ~2000 nt. Lane 3 contains the RNA extracted from the assembly mix *after RNase treatment*. Most significantly, no band in this lane is present at 6400 nt and the shorter-length smear begins at ~4000 nt and runs down to ~2000 nt, from which it can be concluded that *no RNAs longer than ~4000 nt were packaged into singlet VLPs*. Rather, the predominant doublet VLPs involve the sharing by two capsids of > 4000 nt-long RNAs that are digested into smaller

molecules upon RNase treatment. The strong signal in lane 3 from RNAs around 3000 nt in length (the average size of CCMV viral RNA) is consistent with these molecules being the ones that are preferentially packaged into singlet VLPs and hence RNase-resistant.

4200 nt RNA Can Be Packaged by CCMV CP, but Inefficiently

To establish whether it is possible to package 4200 nt RNA in CCMV VLP singlets, we assembled in vitro-transcribed 4200 nt RNA with CCMV CP and extracted the RNA from the resulting VLPs before and after treatment with RNase. The typical structures that are found when packaging the 4200 nt RNA are shown in Fig. 3.5a. The agarose gels in Fig. 3.5b show in vitro-transcribed 4200 nt RNA along with shorter transcript RNA (lane 3). Lane 1 in the agarose gel shown in Fig. 3.5b contains RNA extracted from 4200 nt-RNA CCMV-VLPs before RNase treatment, and lane 2 contains RNA extracted from RNase-treated 4200 nt-RNA CCMV-VLP samples. The in vitro-transcribed RNA used in these self-assemblies is run in lane 3, showing a strong band at 4200 nt corresponding to the full-length RNA transcript—but also showing a significant amount of faster-running shorter RNA that arises from hydrolysis degradation and premature transcripts. From the relatively weak 4200 nt band in lane 2 (involving the RNA extracted from RNase-treated VLP assemblies), along with the accompanying RNA intensity there associated with faster-running shorter RNA, we conclude that only a small percentage of the full-length RNA has been packaged into RNase-resistant singlets, i.e., most of the singlets observed for assemblies with in vitro-transcribed 4200 nt RNA correspond to the packaging of shorter RNA. By comparing the intensity profile in lane 1 (involving the RNA extracted from VLP assemblies that have *not* been treated with RNase) with the profile in lane 3 (in vitro-transcribed RNA), we conclude further that the full-length RNA packaged in multiplets has been cleaved by RNase to give the bands at shorter lengths in lane 2 which also include contributions from direct packaging of the shorter transcripts in the 4200 nt-RNA sample.

Size Differences in Doublet Pairs Suggest Preference for Packaging 3000–4000 nt RNA

It had been noted earlier [17] that while the diameters of capsids contained within doublets and triplets for assemblies of 9000 nt RNA are most likely to be closely similar, the distribution of size differences has a half-width of about 5 nm and differences as large as 10 nm are observed. As shown in Fig. 3.6, we have found this as well for the doublets formed in capsids assembled around 6400 nt RNA. Similarly, the largest size differences observed for the assemblies around 3200 and 4200 nt RNA are also as large as 10 nm, and the maxima in the distributions occur at around 5 nm. These results are consistent with the strong preference of CCMV for packaging 3000–4000 nt of RNA, which can be achieved in triplets or doublets for 9000 nt-RNA assemblies and in doublets for a 6400 nt RNA. For the 3200 and 4200 nt RNA assemblies, however, size differences in doublets must necessarily be larger in order to accommodate cargo sizes compatible with T = 2– and T = 1–sized capsids.

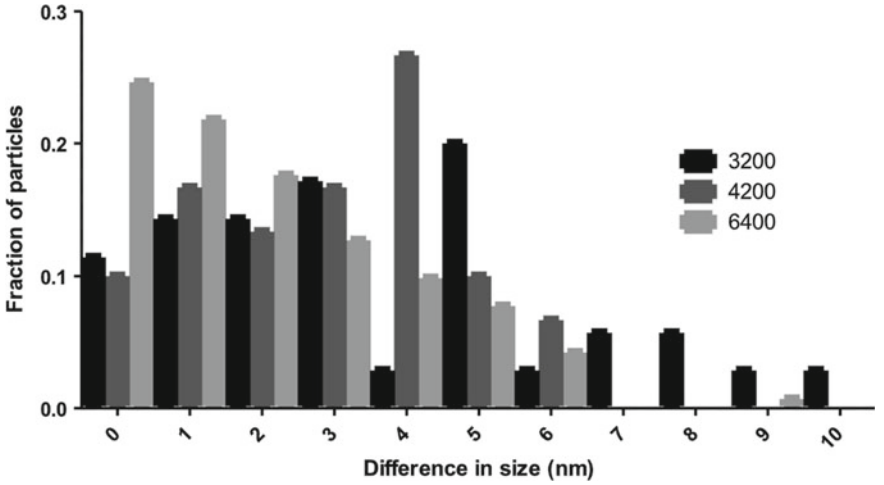

Fig. 3.6 Distribution of size differences between particles in doublets, for assemblies of CCMV CP with 3200, 4200, and 6400 nt RNA

Spermine Has a Small Enhancing Effect on Packaging

The ability of the polyamine spermine to compactify RNAs has been demonstrated for a variety of molecules [24, 25]. The control parameter is the ratio of total cationic (spermine) charge to total anionic (nucleotide/phosphate) charge, with each spermine having a +4 charge and each RNA nucleotide corresponding to a −1 charge from the phosphate backbone. In the case of the 4196 nt RNA, for example, the compaction by spermine is clear from the fact that RNA samples incubated (at room temperature for 30 min) at different spermine:RNA ratios are seen to run progressively faster in an agarose gel (not shown) as spermine concentration is increased, even though the effective charge on the RNA is decreased; for charge ratios above 1, however, the spermine also aggregates the RNA. Continuous compaction of the RNA with increasing cationic:anionic charge ratio is documented more directly and systematically by dynamic light scattering measurements [26], suggesting that there should be a corresponding effect of spermine on the distribution of multiplets in overlong-RNA assemblies and on the maximum length that can be packaged in RNase-resistant singlets.

Particle counts of VLPs made with spermine-compacted RNA show (Fig. 3.7) that for a 4200 nt RNA at a spermine:RNA charge ratio of 0.25:1 there is a 10% increase in the number of singlet particles observed. In assemblies at charge ratios of 0.5:1 and 0.75:1 there is no further increase in singlets; rather, at these higher charge ratios there is an increase in the number of aggregated RNAs and hence higher-order multiplets. For all of these experiments spermine was stored in degassed water overlaid with argon gas and frozen at −80 °C until used. Frozen RNA was thawed on ice and thermally denatured by heating to 90 °C at a rate of 1 °C s^{-1}, holding at

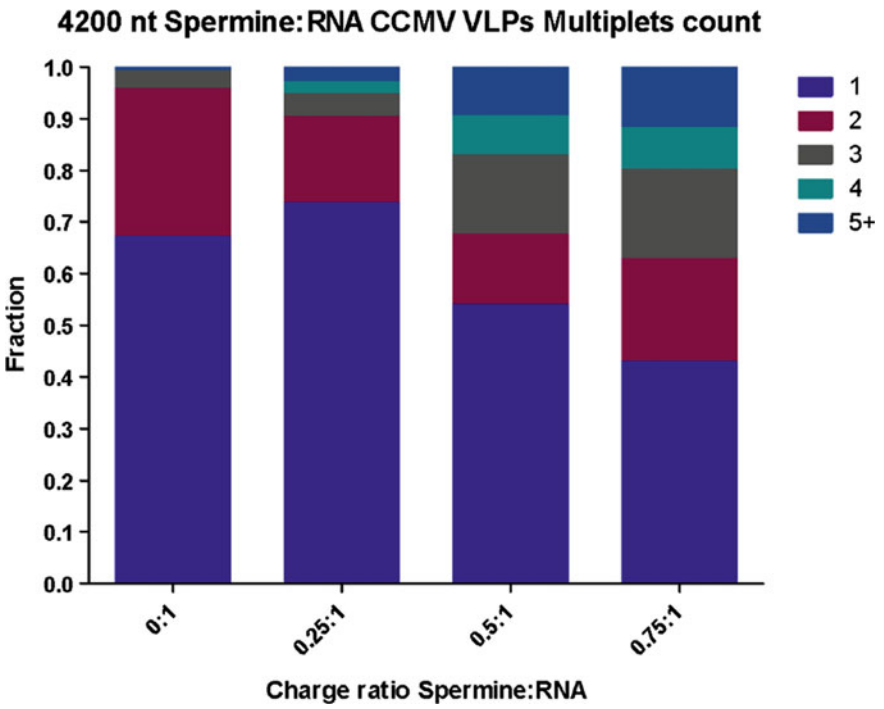

Fig. 3.7 Effect of spermine on RNA and its packaging by CCMV CP

90 °C for 1 s, then cooling to 4 °C at a rate of 1 °C s^{-1}. Thermally denaturing and slowly cooling the RNA allows for disruption of any duplexes formed between RNA molecules in the freezing process and for refolding of the RNA. The RNA was then mixed with spermine at a specified charge ratio and incubated at 4 °C for 30 min.

Fraction of singlets (blue), doublets (red), triplets (grey), quadruplets (green) and larger particles (light blue) in samples of CCMV VLP particles formed from 4200 nt RNA compacted with spermine at different charge ratios.

That the spermine does not prevent the full-length of 4200 nt RNA from being packaged by singlet CCMV VLPs is confirmed by extracting the RNA from a CP assembly carried out for a 0.25:1 spermine:RNA ratio and running it in a gel (not shown).

Conclusions/Perspectives

An examination of the distribution of multiplets formed when RNA longer than 3000 nt is mixed under assembly conditions with recombinant CCMV CP has shown that there is a limit to the length of RNA that can be packaged into a single, RNase-resistant, capsid. Consistent with results obtained previously for infected-plant-derived CCMV CP, the frequency of doublet and higher-order multiplets increases with increasing length of RNA.

More explicitly, with gel electrophoresis analyses of RNA extracted from RNase-treated assembly mixes, we have shown that CCMV CP cannot form VLPs that protect a significant fraction of RNA longer than about 4200 nt. In particular, gels of RNA extracted from CCMV VLPs indicate that in an assembly mix of CP and 6400 nt RNA there are no capsids containing RNAs longer than 4200 nt that are RNase resistant. This suggests that the majority of the singlets observed in EM contain only fragments shorter than 4200 nt, which are a natural result of in vitro RNA transcription of the 6400 nt-TMV-RNA. Multiplets also occur when recombinant CP of BMV is used to package long RNA.

As remarked in the Introduction, the compaction of viral-genome-length RNA by polyamines has been demonstrated by several different techniques, including sedimentation velocity measurements [24], fluorescence correlation spectroscopy [25], and dynamic light scattering [26]. We find that this effect of spermine can be used to improve the packaging efficiency of "overlong" RNA by CCMV and BMV capsid protein, e.g., the number of RNase-resistant singlet particles can be increased when RNA is incubated with spermine prior to assembly. However, this effect is small, and introducing spermine also results in RNA aggregation at spermine:RNA charge ratios greater than one, thereby limiting increase in the number of singlets.

Our work establishes the inability of CCMV or BMV CP to package RNA longer than 4200 nt into particles which are resistant to RNase. In doing so, we highlight the strongly-evolved preference of CCMV to form $T = 3$ particles. The formation of multiplets is evidence of the dominant role of the spontaneous curvature of the CP in the assembly of VLPs of flexible anionic polymers such as RNA in determining the diameter of the capsids. In contradistinction, as in the case of charged nanoemulsion particles [28] that are incompressible, multiplets cannot form and the size of the capsid therefore must increase to accommodate increasing size of the cargo. The larger fraction of multiplets in the assemblies with BMV as compared to CCMV is likely attributable to the lower positive charge on the N-termini of the capsid protein (+9 for BMV rather than +10 for CCMV) resulting in a greater degree of overcharging of the RNA compared to that of the CP and therefore a smaller amount of RNA that can be stabilized by a single capsid. Moreover, electrophoretic mobility measurements show that at the pH of assembly, 4.5, the magnitude of the charge on CCMV capsids is about twice that on BMV capsids [29], making assembly of capsids into multiplets more difficult.

These observations demonstrate both the promise and limitations of CCMV and BMV VLPs as gene-delivery platforms. The in vitro self assembly provides highly

monodisperse and pure VLPs in good yields with easily obtained recombinant CP. However, these positive features come with the disadvantage of a major length restriction of 4200 nt in packaging capacity, which with the use of spermine may be only modestly increased. If, for example, the aim is to deliver self-amplifying RNA, the gene of interest must be genetically fused to an RNA-dependent-RNA-polymerase replicase, the shortest of which (that for the Nodamura virus) is 3000 nt long, thereby limiting the length of the gene of interest to only 1200 nt. If one wishes to retain the advantages of in vitro self assembly, then, the choice for packaging of longer RNA is TMV CP, despite the necessity of employing RNA constructs that encode an origin of assembly sequence (OAS) [30]. Here, in contrast to the situation with spherical viruses, the highly-evolved preference of capsid protein to form a hollow *cylinder* with fixed inner and outer radii can be satisfied by any length of RNA because the capsid curvature is independent of length.

Acknowledgements We are grateful to Jerrell Tisnado for expression and purification of recombinant capsid protein, and to all members of our research group for many helpful discussions about viral self-assembly over the past several years. This work has been supported financially by the National Science Foundation (Molecular and Cellular Biosciences Division, Genetic Mechanisms Program, Grants 1716975 and 2103700 to WMG).

References

1. Yildiz I, Tsvetkova I, Wen AM, Shukla S, Masarapu MH, Dragnea B, Steinmetz NF, Fu Y, Li J (2016) A novel delivery platform based on Bacteriophage MS2 virus-like particles. Virus Res 211:9–16
2. Vaidya AJ, Solomon KV (2022) Surface functionalization of rod-shaped viral particles for biomedical applications. ACS Appl Bio Mater 5(5):1980–1989
3. Smith MT, Hawes AK, Bundy BC (2013) Reengineering viruses and virus-like particles through chemical functionalization strategies. Curr Opin Biotech 24:620–626
4. Biddlecome A, Habte HH, McGrath KM, Sambanthamoorthy S, Wurm M et al (2019) RNA vaccines in in vitro reconstituted virus-like particles. PLoS One 14(6):e0215031
5. Saunders K, Thuenemann EC, Peyret H, Lomonossoff GT (2022) The Tobacco Mosaic virus origin of assembly sequence is dispensable for specific viral RNA encapsidation but necessary for initiating assembly at a single site. J Mol Biol 434:167873
6. Choi Y, Dreher TW, Rao ALN (2002) tRNA elements mediate the assembly of an icosahedral virus. Proc Natl Acad Sci USA 99:655–660
7. Annamalai P, Rao ALN (2005) Dispensability of 3' tRNA-like sequence for packaging cowpea chlorotic mottle virus genomic RNAs. Virol 332:650–658
8. Cadena-Nava RD, Comas-Garcia M, Garmann RF, Rao ALN, Knobler CM, Gelbart WM (2012) Self-assembly of viral capsid protein and RNA molecules of different sizes: requirement for a specific high protein. RNA mass ratio. J Virol 86:3318–3326
9. Stockley PG, Rolfsson O, Thompson GS, Basnak G, Francese S, Stonehouse NJ, Homans SW, Ashcroft AE (2007) A simple, RNA-mediated allosteric switch controls the pathway to formation of a T = 3 viral capsid. J Mol Biol 369:541–552
10. Zhan S, Li J, Xu R, Wang L, Zhang K, Zhang R (2009) Armored long RNA controls or standards for branched DNA assay for detection of human immunodeficiency virus type 1. J Clin Microbiol 47:2571–2576

11. Cohen SS, Greenberg ML (1981) Spermidine, an intrinsic component of turnip yellow mosaic virus. Proc Natl Acad Sci USA 78:54470–55474
12. Clarke IN, Lambden PR (2000) Organization and expression of calicivirus genes. J Infect Dis 181:S309–S316
13. Gopal A, Ececioglu D, Yoffe AM, Ben-Shaul A, Rao ALN, Knobler CM, Gelbart WM (2014) Viral RNAs are unusually compact. PLoS ONE 9:e105875
14. Erdemci-Tandogan G, Wagner J, van der Schoot P, Podgornik R, Zandi R (2014) RNA topology remolds electrostatic stabilization of viruses. Phys Rev E 89(3):032707
15. Garmann RF, Comas-Garcia M, Koay MST, Cornelissen JJLM, Knobler CM, Gelbart WM (2014) The role of electrostatics in the assembly pathway of a single-stranded RNA virus. J Virol 88:10472–10479
16. Beren C, Cui Y, Chakravarty A, Yang X, Rao ALN, Knobler CM, Zhou ZH, Gelbart WM (2020) Genome organization and interaction with capsid protein in a multipartite RNA virus. Proc Natl Acad Sci USA 117:10673–10680
17. Dai X, Li Z, Lai M, Shu S, Du Y, Zhou ZH, Sun R (2017) In situ structures of the genome and genome-delivery apparatus in an ssRNA virus. Nature 541(7635):112–116
18. Thurm AR, Beren C, Duran-Meza AL, Knobler CM, Gelbart WM (2019) RNA homopolymers form higher-curvature virus-like particles than do normal-composition RNAs. Biophys J 117(7):1331–1341
19. Brasch M, Cornelissen JJLM (2012) Relative size selection of a conjugated polyelectrolyte in virus-like protein structures. Chem Commun 48(10):1446–1448
20. Kler S, Wang JC, Dhason M, Oppenheim A, Zlotnick A (2013) Scaffold properties are a key determinant of the size and shape of self-assembled virus-derived particles. ACS Chem Biol 8(12):2753–2761
21. Li C, Kneller AR, Jacobson SC, Zlotnick A (2017) Single particle observation of SV40 VP1 polyanion-induced assembly shows that substrate size and structure modulate capsid geometry. ACS Chem Biol 12(5):1327–1334
22. Hagan MF, Elrad OM (2010) Understanding the concentration dependence of viral capsid assembly kinetics—the origin of the lag time and identifying the critical nucleus size. Biophys J 98(6):1065–1074
23. Perlmutter JD, Qiao C, Hagan MF (2013) Viral genome structures are optimal for capsid assembly. eLife 2:e00632
24. Mitra S, Kaeseberg P (1965) Biophysical properties of RNA from turnip yellow mosaic virus. J Mol Biol 14(2):558–571
25. Borodavka A, Dykeman EC, Schrimpf W, Lamb DC (2017) Protein-mediated RNA folding governs sequence-specific interactions between rotavirus genome segments. Elife 6:e27453
26. Duran-Meza AL, Oster L, Sportsman R, Phillips M, Knobler CM, Gelbart WM (2023) Long ssRNA Undergoes Continuous Compaction in the Presence of Polyvalent Cations. Biophys J. https://doi.org/10.1016/j.bpj.2023.07.022
27. Zhao X, Young MJ (1995) In vitro assembly of cowpea chlorotic mottle virus from coat protein expressed in *Escherichia Coli* and in vitro transcribed viral cDNA. Virology 207:486–494
28. Chang CB, Knobler CM, Gelbart WM, Mason TG (2008) Curvature dependence of viral protein structures on encapsidated nanoemulsion droplets. ACS Nano 2(2):281–286
29. Duran-Meza AL, Villagrana-Escareño MV, Ruiz-García J, Knobler CM, Gelbart WM (2021) Controlling the surface charge of simple viruses. PLoS ONE 16(9):e0255820
30. Smith ML, Corbo T, Bernales J, Lindbo JA, Pogue GP, Palmer KE, McCormick AA (2007) Assembly of trans-encapsidated viral vectors engineered from Tobacco mosaic virus and Semliki Forest virus and their evaluation as immunogens. Virology 358(2):321–333