Effect of RNA self-interactions on the self-assembly of virions

Tomica Kralj

Table of contents:

Introduction

Viruses are small replicating particles. They are parasites by nature (most of the viruses cause diseases), they don't have their own metabolism and cannot reproduce without the machinery of a host cell. They can survive outside of a cell but only as inactive spores.

Genetic material of viruses can be either DNA or RNA, encapsulated in protein container. Besides protein shell, some viruses also have lipid coating – for example, modified form of the host cell outer membrane or internal membranes such as nuclear membrane or endoplasmic reticulum. The lipid membrane itself and any carbohydrates present originate entirely from the host while some proteins are coded for by the host and some by the viral genome. Most viruses that have this kind of coating are dependent on this envelope for their infectivity.
After the initial assembly of a virus, the capsid proteins are often modified, a process known as maturation.

Capsid T-number

One of the remarkable properties of capsids is that they self-assemble in a regular and well-defined structure. The capsids need to be able to effectively protect the packaged genome, but they also need to be able to open at the right place and time to release the genome. Therefore, the demands on the capsids mechanics and function are highly specific.

Viral capsids are composed of one or several types of proteins. When capsid is composed of several different proteins they are organized in dimers or trimers which tile the surface forming a regular pattern. There are two basic shapes of viruses – icosahedral and helical. There is also prolate shape, which can be consider as a subclass of icosahedral, prolonged in the direction of its five-fold axis. In icosahedral capsids the proteins form pentagons and hexagons, and these capsids can be described by assigning to them a triangulation number (in further text T-number), which can be defined as follows:

On a surface tiled with hexagons we can establish a coordinate system such that one hexagon is the centre, and the axes are perpendicular to four of its six sides. Then we can assign two numbers for a hexagon at any arbitrary position by counting its distance on the axes from the first one.
Viruses can be visualized as covered in hexagonal grip, with the exception of having exactly 12 pentagons at equal distance from each other. If we measure the distance between two nearest pentagons, with h being the number of steps on one axis, and k on the other, and with h greater then or equal to k, then we can get the T number through following formula:

$$T = h^2 + hk + k^2 = (h+k)^2 - hk \qquad (1)$$

Simply put, the larger the T-number is, the more there are hexagons in respect to pentagons. T-number adopts certain integer values - 1, 3, 4, 7, 9, 12, and so on. The number of proteins constituting icosahedral shells, q, is 60 times the T number and this is called a structural index of viral capsids.
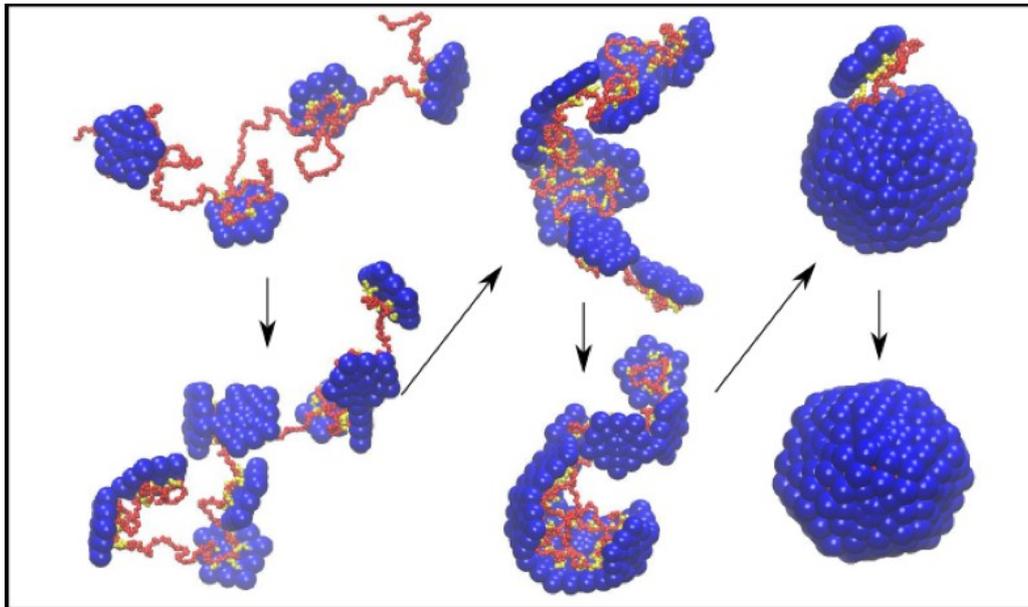
Figure 1: Illustration of an assembly of subunits into a complete capsid. Beads are colored as follows: blue=globular portions of proteins, yellow=positive arginine rich motive beads, red=polyelectrolyte.

Mechanical properties of capsids

The mechanical properties of viral capsids are crucial for their proper functioning, because rupture of the capsid due to defects occurring during maturation or as a result of external influences outside the host will, in general, result in a loss of infective capability [1].

Bulk experiments have provided knowledge about the averaged properties of viruses, and by making use of single-particle techniques such as atomic force microscopy (AFM) it was possible to take measurements of and manipulate with individual viruses. AFM was used to focus point forces on a single particles [2], making it possible to measure the strength of the protein-protein interactions in the capsid. Furthermore, AFM is perfectly suited for studying mechanical properties of small biological entities such as viruses due to its high sensitivity in applying and measuring forces combined with the possibility to work under physiological conditions. AFM has been used to image a range of viruses [3], but here the focus is on those experiments measuring the mechanical properties of viral capsids.

After a virus or an empty viral shell has assembled, we can inquire how resilient it is in terms of its response to external force and other perturbations. Capsids need to meet conflicting demands: they should be sufficiently stable to protect their genome in the extra-cellular environment, but sufficiently unstable that they can release their genome molecules into host cells [4]. Various bulk and single-particle assays have been developed to measure the mechanical properties of viruses, the growing field of mechanical virology.

Experiments with osmotic-shock were used to observe the stability of bacteriophage viruses under pressure against rupture and the mechanical properties of crystals and films composed of viruses

were analyzed by Brillouin light scattering. A disadvantage of these multiparticle techniques is that they represent an average over large numbers of viruses and also, they represent a rotational average, so any directionality of the mechanical properties with respect to the shell orientation is lost. The mechanics of single particles and their directionality can however be probed with the atomic force microscopy (AFM) based nanoindentation techniques [5]. The relation between the applied force and the resulting change in shell diameter is called the force–deformation curve (FDC). Depending on whether or not the capsid returns to its original state after the probe force is removed ('unloading'), we call this a reversible, respectively irreversible, deformation. The force measured by a nanoindentation probe results, at a fundamental level, from the fact that the probe forces the viral shell away from a state of minimum free energy.

To interpret measured FDCs, including irreversibility effects, we can compare them with the deformation free energy obtained from the continuum elasticity theory of thin elastic shells ('thin-shell theory' or TST) that we have already mentioned. TST is used extensively by engineers to predict the effects of external forces on thin-walled, hollow macroscopic structures, such as airplanes or oil tanks. In the simplest application of TST we model a viral shell as a thin spherical shell of uniform thickness and radius R. If the viral shell encloses genome molecules, then an internal osmotic pressure Π must be included, which can be as large as ~50 atm. Let $\zeta(r)$ be the indentation profile of the shell generated, for example, by a force probe. Specifically, $\zeta(r)$ is defined as the radial inward displacement of the surface of the sphere expressed in terms of a two-dimensional coordinate system that covers the shell. In the limit of small $\zeta(r)$, the TST deformation free energy 1F is a simple functional of $\zeta(r)$ in the form of an integral over the shell surface:

$$\int dS \left(\frac{1}{2}\kappa(\Delta\zeta)^2 + \frac{1}{2}\tau(\nabla\zeta)^2 + \frac{1}{2}Y\left(\frac{2\zeta}{R}\right)^2 \right) \qquad (2)$$

The first term of equation (2) describes the bending-energy cost of the indentation — note that $\Delta\zeta$ is the shell curvature — where the bending modulus κ has units of energy. The second term represents the work by the probe against the genome osmotic pressure Π with $\tau = \Pi R/2$ an effective surface tension. The third term measures the stretching of the layer induced by the force with the two-dimensional Young modulus Y of the layer. A dimensionless number $\gamma = YR^2/\kappa$ - the Föppl–von Kármán number - and a characteristic length scale l_B = buckling radius — can be constructed from the stretching and bending moduli, which will play an important role.

Alternatively, we can also apply three-dimensional elasticity theory to compute the elastic response of an elastic shell with a finite thickness h. We recover the TST result in the limit h < R with a spring constant

$$\kappa \propto E_{3D} h^2 / R \qquad (3)$$

where E_{3D} is the three-dimensional Young modulus. For larger indentation forces equations (2) and (3) should not be used. The calculation of the FDC of TST in the nonlinear regime requires the solution of a pair of somewhat challenging nonlinear differential equations, known as the Föppl–von Kármán (FvK) equations (they resemble Einstein's equations of general relativity). Instead of trying to solve the FvK equations analytically or numerically, it is more practical to numerically minimize the elastic energy directly using finite-element modeling (FEM). If a shell is indented by a hemispherical tip, the deformation of the sphere does not deviate much from the linear harmonic spring for deformation ratios $\zeta(0)/R$ up to 0.6. Then, for slightly larger values of $\zeta(0)/R$, a discontinuous drop takes place in the FDC. This is due to the fact that for larger deformations the elastic energies of two different shapes of the deformed shell cross each other. In the engineering

literature, singularities in the FDC of this type are known as 'buckling' transitions. They are identified with the well-known catastrophic failures of hollow structures subject to external loads, that is, failures without any visible precursor 'warning' in the FDC. Comparison with the FDC suggests a relation between the buckling instabilities of TST and the irreversible nonlinearities of the FDCs of viral shells [6]. However, mathematically, the buckling discontinuities of TST are quite similar to first-order phase transitions and, like first-order phase transitions, they could be nucleated by local structural defects. This indicates that the elastic response of the non-uniform icosahedral shells might differ from that of uniform spherical shells, which must be discussed before we can compare with experiment. The FDC of icosahedral shells was obtained by starting from a perfect icosahedron as the initial trial state. The sharp folds linking the 12 vertices of a perfect icosahedron are not compatible with the bending-energy term in equation (2).

However, as long as the FvK parameter $\gamma = YR^2/\kappa$ threshold value of the order of 10^2 still remains icosahedrally facetted. For FvK numbers less than this threshold, however, the shell adopts a nearly spherical shape (confusingly, this also is known as a buckling transition, but we shall not use this terminology). The FvK number of a viral shell can be estimated by comparing computed shapes of undeformed shells with those measured, for example, by cryo-transmission electron microscopy. For lower values of γ, the FDC remains quite close to the harmonic spring prediction. For larger values of γ, the relation is increasingly nonlinear, and then develops the buckling discontinuity. The size of the discontinuity increases with increasing γ and the critical value of the indentation for the buckling discontinuity decreases.

In the buckled state, the shell is detached from the tip at the centre, which is not the case in the small-force regime. The five-fold-symmetry sites thus indeed seem to act as structural defects that trigger buckling. The discontinuity of the FDC of a spherical shell with the same elastic moduli takes place at a much larger indentation. How do the predictions of TST compare with the AFM nanoindentation experiments? For small applied forces, the measured FDC is indeed linear in many cases [7]. Comparing the three-dimensional Young moduli (equation (3)) of various particles shows that sphere-like viruses that package their genome into preformed capsids, such as phage $\Phi29$, phage λ, HSV1 (herpes simplex virus type 1) and MVM (minute virus of mice) have a Young modulus that is at least double that of sphere-like viruses that self-assemble around their genome such as CCMV and HBV [8]. The FvK numbers were, incidentally, not obtained by comparing with measured shell shapes but, instead, were estimated assuming the TST relation

$$\gamma = 12(1-v^2)\left(\frac{R}{h}\right)^2 \qquad (4)$$

with v Poisson's ratio. An interesting application is the use of TST to explain measured differences in spring constants of 'nuclear' and 'viral' HSV1 capsids. The latter are stiffer than the former because they possess an extra protein layer, the inner tegument. Using equation (3), and assuming that the E_{3D} values for the capsid and inner tegument are similar, it follows that this extra protein layer should have a thickness of ~0.8 nm, a prediction that is verifiable by electron microscopy. For smaller viral particles, when the shell thickness h is not negligible compared with the radius R, TST is no longer expected to apply. The simplest extension is to use FEM to compute the FDC of a homogeneous elastic shell with a finite thickness. The elastic energy of a solid elastic sphere that is indented scales as $\zeta^{5/2}$, which is known as a 'Hertzian' response. The FDC of a thick-walled shell is expected to show, as a function of h, scaling crossover from the TST result for larger applied forces to a nonlinear Hertzian-type FDC for smaller applied forces. FEM studies of the indentation of elastic shells by point forces realistically shaped models for the AFM tip, were carried out. It was indeed observed that Hertzian nonlinearities occur at the onset of deformation of thick-shelled

particles. The next step is to use information on the heterogeneous geometry of the viral particles available from X-ray diffraction and cryo-electron microscopy studies, while still maintaining a uniform elastic modulus. Such an approach was also followed to investigate CCMV and HBV. By comparison with the measured FDC, a Young modulus of 0.22 GPa was found for CCMV, which happens to lie between the estimates obtained by the previous two methods. A comparable Young modulus, namely 0.26 GPa, was determined for HBV, which is a little lower than that obtained by using a TST approximation. Determining the Young modulus thus depends to some extent on the model that is used to analyse the FDC. Another example was a detailed FEM study of MVM that predicted stabilizing interactions between the encapsulated DNA and specific sites at the capsid interior, which was later experimentally confirmed. Furthermore, the orientation-dependent indentation, as well as by behaviour of HBV was determined by comparing experiments with detailed FEM simulations.
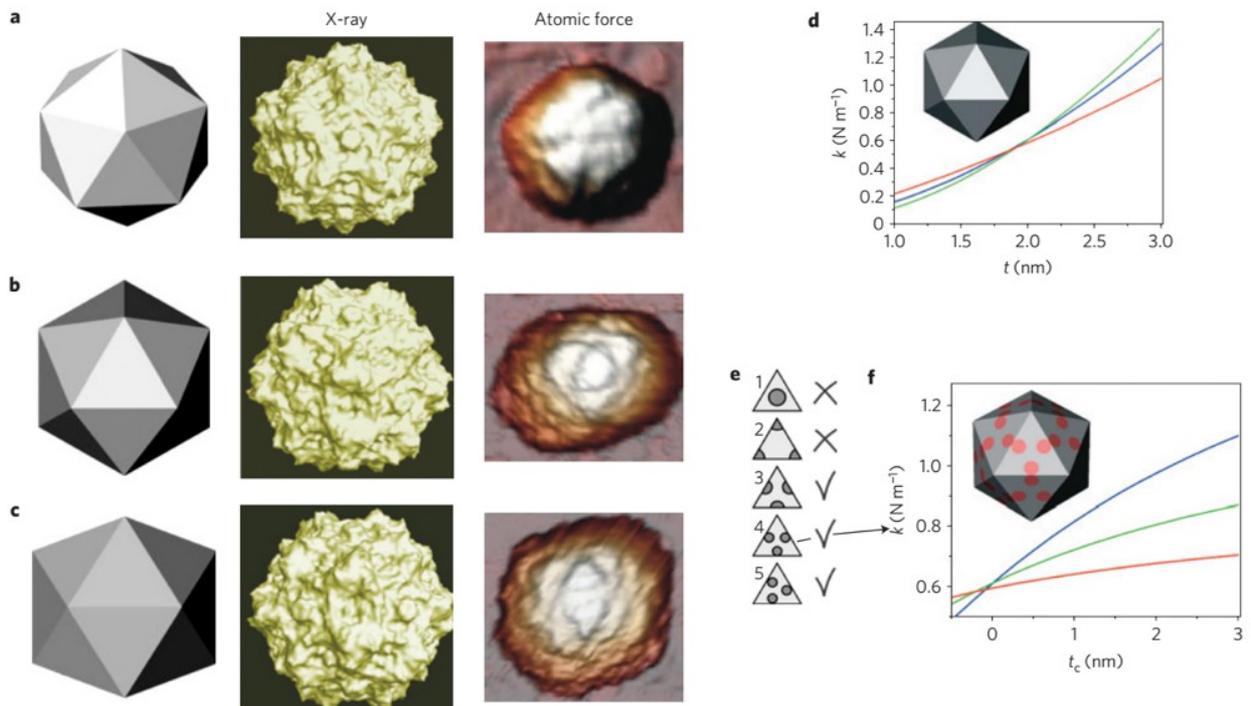

Figure 2: orientation dependence of MVM. a–c, From top to bottom: particles as seen along the five-, three- and two-fold symmetry axes. From left to right: schematic images of icosahedrons, reconstructions of MVM capsids, AFM images of MVM capsids. d, FEM analysis along the five-fold (red), three-fold (green) and two-fold (blue) symmetry axes as a function of shell thickness t. The experiments yield similar spring constants along all three axes of the empty particles (data not shown). These results match best with simulations for t ~ 2 nm. e, Reinforced shell models with patches of extra thickness tc at various sites. Only Models 3–5 predict the correct anisotropic reinforcement of DNA-filled MVM capsids as observed experimentally. Importantly, the patches in these three models coincide roughly with the locations where ordered DNA is bound to the shell, whereas this does not coincide in Models 1 and 2. f, FEM analysis result for Model 4 in e. [9]

6

VLP formation

Formation of virus-like particles (VLPs) and packaging of poly(styrene sulfonate) (PSS) by the protein of cowpea chlorotic mottle virus (CCMV)

In figure 3, a and b, we can see transmission electron microscopy (TEM) images from two packaging reactions, along with images of the empty CCMV capsid (c) and wt CCMV (d) for comparison. We can see that VLPs with spherical morphology are readily formed. The different preparations are clearly distinguishable. Prominent dark centers in all the empty capsids arise from penetration of the stain into the largely vacant interior. Similarly, the smaller dark centers in the images of wt CCMV capsids are consistent with the small central void that is known to exist in the packaged RNA in Bromoviruses. In contrast, there is markedly less stain penetration into the cores of the VLPs or, at most, a small apparently void region, which suggests that the PSS occupies the interior.
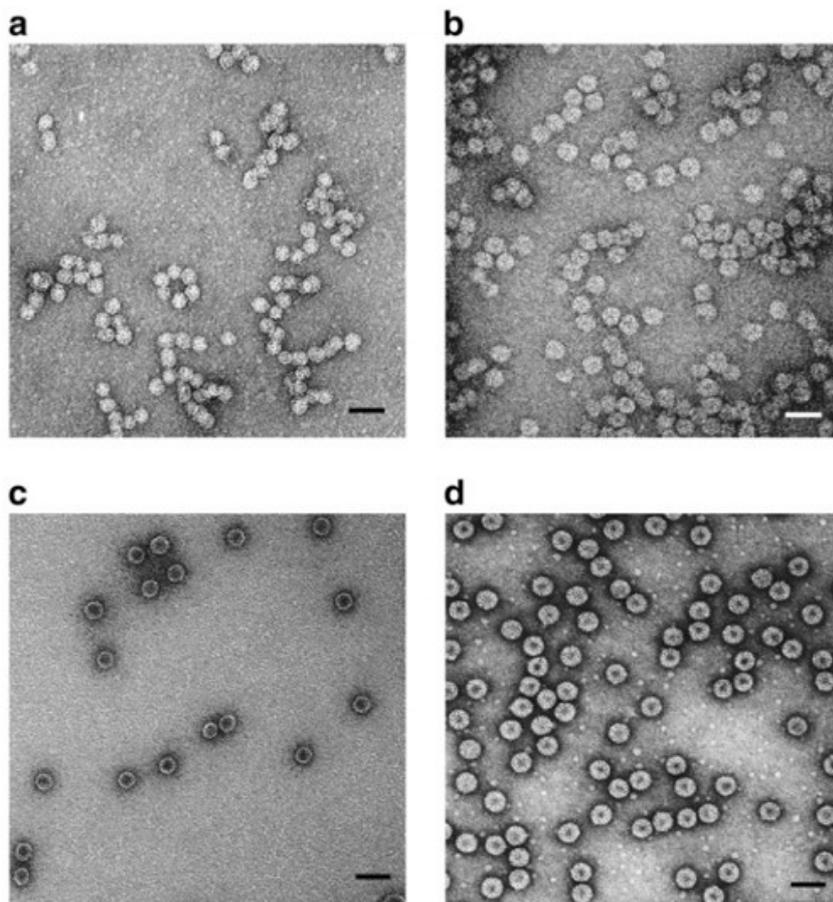

Figure 3: TEM images of capsids formed in self-assembly reactions. Samples were stained with 2% uranyl acetate. (a) VLPs formed with 700-kDa PSS. The mean capsid size for VLPs is 22 nm. (b) VLPs formed with 3.4-MDa PSS. The mean capsid size is 27 nm. (c) Empty CCMV capsids formed by dialysis of CP in buffer with high salt and low pH. The dark core in the center indicates the penetration of stain into ''void'' (aqueous solution) space, which is notably absent in the interiors of VLPs filled with PSS. (d) wt CCMV capsids in virus suspension buffer. Scale bars are 50 nm. [10]

Additional evidence of PSS encapsidation by CP was provided by UV absorbance measurements of VLP 2M that was fractionated on a sucrose gradient. Transition of the absorbances at 270 nm, where the absorbance of PSS is stronger, and at 290 nm, where only the CP has significant absorption, demonstrates that the protein and polymer (CP and PSS) are found in the same fractions. Further, the strongest absorbances are located mainly in fractions 7–13. In contrast, on the same gradient, the wild-type (wt) CCMV formed a band between fractions 11 and 16 and peaked at fraction 13. The width of the VLP 2M band was comparable to that of the wt CCMV, suggesting that the VLP population was relatively homogeneous, which is confirmed by direct measurements of the size distributions with EM.

The encapsidation of PSS by CCMV CPs has also been reported for a much lower-molecular-mass (9.9 kDa) polymer that was labeled with dansyl chloride. In those experiments, fast performance liquid chromatography was used to separate the products and it was showed that the CP and labeled PSS comigrated, consistent with the PSS being packaged inside the VLPs. Studies of the packaging of fluorescent (rhodamine-labeled) PSS under the same assembly conditions show that when the assembly products are fractionated on a sucrose gradient, the fluorescent PSS and CP comigrate, as in the case of the unlabeled PSS shown in figure 4. Moreover, the rhodamine fluorescence is not markedly diminished in the presence of the quencher methyl viologen, demonstrating that the labeled PSS was protected from the quencher by the capsid. In contrast, the fluorescence of a control sample containing only free fluorescent PSS was significantly diminished at the same quencher concentration. Similarly, fluorescence is quenched when free labeled PSS is added to a solution of wt CCMV, indicating that it is not sufficient for the PSS to adsorb on the outside of capsids but rather that it must be packaged inside to be protected against quenching. These results suggest that the PSS in the assembly reactions reported in this work is indeed packaged inside the capsids and not adsorbed outside.
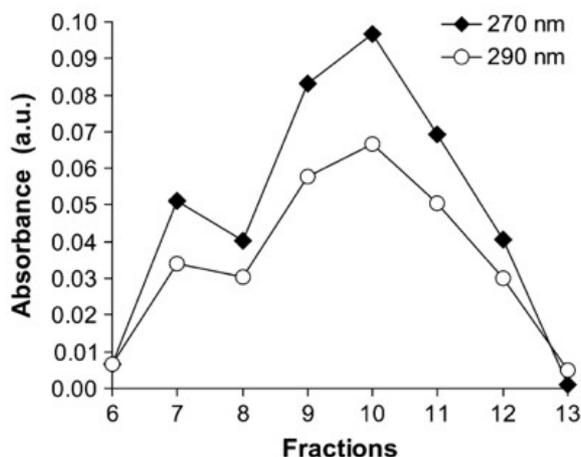

Figure 4: Separation of the products of the 2-MDa PSS plus CP assembly reaction on a 10–40% sucrose gradient. Comigration of species absorbing strongly at both 270 and 290 nm was found for fractions 7–13. The absorbances peaked at fraction 10. [10]

VLP size distributions

The normalized VLP capsid size distributions are shown in figure 5 (each is based on measurements of more than 100 particles). Figure 5 f shows the bimodal distribution obtained by combining the data from Reactions 1–5. The widths of the VLP size distributions, as measured by the standard deviations of the Gaussian fits, are essentially identical for all five packaging reactions, and they are comparable to the width we measure for wt CCMV, 1.70 nm. Although some of the dispersion in capsid sizes is related to variations in experimental conditions such as drying, a dispersion in size even in the absence of such artifacts is expected because capsids in solution are dynamic objects that have breathing modes.

The fact that the sizes of hundreds of VLP capsids from five independent packaging reactions converge to the two discrete values of 22 and 27 nm strongly implies that they are determined by the inherent icosahedral symmetry of the protein shells. Indeed, it has been established that the CCMV CP forms icosahedral structures with triangulation numbers T=1, 3, 4, and 7. Without an image reconstruction analysis based on cryo-EM or x-ray diffraction, the assignment of a T number to a particular species is difficult. Absolute measurements of capsid diameters by TEM have sometimes been employed but are questionable because it is well established that the drying of the sample results in shrinkage that cannot easily be quantified. If we assume, however, that the staining and sample drying on the TEM grid affects all the VLPs to the same degree, we can deduce the T numbers from the ratios of the capsid diameters associated with the peaks in figure 4. The number of CPs in a capsid is 60T. If the unit area per CP is constant, the total capsid surface area, which is proportional to the square of the capsid diameter, scales linearly with T:

$$R_{ij} = \frac{D_{Ti}}{D_{Tj}} = \sqrt{\frac{T_i}{T_j}} \qquad (5)$$

Here D_{Ti} is the diameter of capsid with T=i. Experimental support for the validity of equation 5 can be found in the literature. For example, the hepatitis B virus can assemble into particles of two sizes, known from cryo-EM to be T = 3 and T = 4 structures; the ratio $R_{3,4}$ of their diameters is 300 A˚:260 A˚= 1.15, which is in excellent agreement with $\sqrt{4/3} = 1.155$. Similarly, the assembly of empty capsids of CCMV can produce T = 2 and T = 3 particles, with diameters of 280 and 230 A˚; these $R_{3,2} = 1.22$, in good agreement with $\sqrt{3/2} = 1:225$. Furthermore, T = 4 and T = 7 capsids have been identified in cryo-EM studies of P22 procapsids; the ratio R7,4 of their average diameters, 1.33, can be compared with $\sqrt{7/4} = 1:323$.
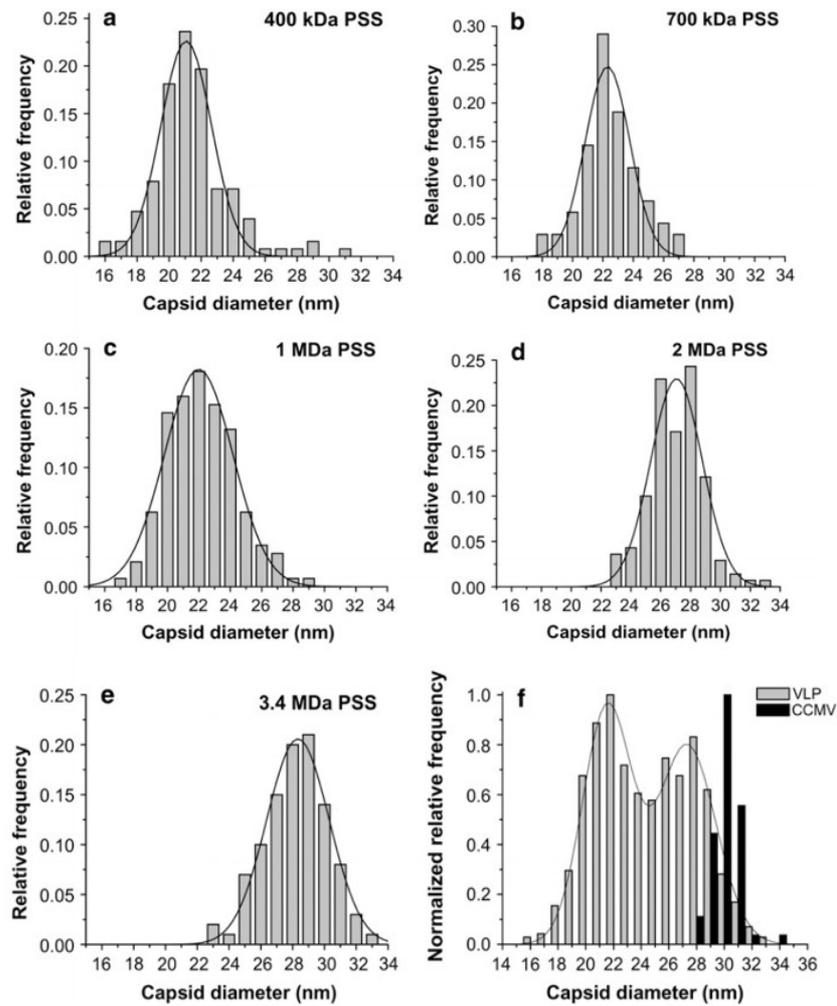
Figure 5: Normalized VLP capsid size distribution histograms for Reactions 1–5. For each reaction, one dominant capsid size was found: (a) VLP 400 kDa, (b) VLP 700 kDa, (c) VLP 1 MDa, (d) VLP 2 MDa, (e) VLP 3.4 MDa; and (f) a combined histogram of capsids from all five reactions. A fit of the histogram to two Gaussians indicates that the capsid sizes converge to two values: 22 nm and 27 nm. These values agree well with the sizes of T ¼ 2 and T ¼ 3 capsids formed by CCMV CPs. [10]

Finally, CCMV capsids of three sizes, established by cryo-EM to be T=1, 2, and 3, were shown to be assembled in vitro from a CP that lacked most of the N-terminal domain; their diameters of 290, 250, and 180 A˚ give $R_{3,2}$=1.2, $R_{3,1}$=1.6, and $R_{2,1}$=1.4, in good agreement with values of 1.2, 1.7 and 1.4, calculated from Eq. 1. These data lend confidence to the use of Eq. 5 to identify the T numbers of the two capsid types for the VLPs that we have studied. The values of 21.5 and 27.3 nm for the two diameters give R=1.27, corresponding to the expected ratio for T=2 and 3, i.e., R= $\sqrt{3/2}$ = 1:225, and differing significantly from the ratios that would be obtained for other reasonable choices of T numbers.

A jump in T number from T=2 to T=3 occurs between molecular masses of 1 MDa and 2 Mda. This switch of the capsid size to a larger T number strongly suggests that for each type of capsid there exists a maximum capacity for efficient packaging. As the packaged cargo size increases, the size of the capsid increases correspondingly. The discreteness of this capsid-size transition results from the inherent icosahedral symmetry property of the CPs. This preference of the CP for discrete sizes is mirrored in the widths of the size distributions. It is significant that the VLPs have essentially

identical widths despite the variations in the polydispersities of the different PSS polymers. It is notable, as well, that the widths are quite similar to that of wt CCMV (figure 5 f). The compressible PSS is able to accommodate to the capsid diameters preferred by the protein, but only to a limited degree, which is why larger VLPs are found for higher molecular masses of polymer.

Genome packaging

The capsid of double-stranded (ds) DNA phages easily self-assembles - sometimes in the presence of scaffolding proteins but in the absence of nucleic acids. Subsequently, a packaging motor internalizes its genome [11, 12], as is also the case for dsRNA bacteriophages [13]. The self-assembly and packaging process can be reproduced *in vitro* using extracts from infected hosts as shown for instance for phage λ [14] and phage Φ29 [15]. Infectious Φ29 virions can also be constructed from cloned gene products and synthetic nucleic acids and infectious phage λ virions have been successfully produced from purified components. During the packaging process, the genome, which can be many microns long, needs to be compressed to fit inside a capsid whose diameter is hundreds of times smaller. Electrostatic repulsion between the nucleic acid strands will hinder the compaction of the genome significantly. Moreover, extensive bending energies need to be overcome as the persistence length of dsDNA and dsRNA is comparable to the diameter of the capsid. In general, however, electrostatic repulsion dominates over bending energy and will give the primary contribution to the internal force built-up. A third factor influencing the packaging is the reduction of entropy of the system. As a result of these energetically and entropically unfavourable changes, it is likely that high forces are involved in the packaging of dsDNA or dsRNA inside phage capsids. The energy needed for this force generation comes from ATP consumption and approximately one ATP molecule is hydrolyzed for every two base-pairs that are packaged.

Conclusions

Influence of RNA length

It was shown that there exist a correlation between length of the encapsidated genome and the optimal radius of a capsid which encloses it. In general, if the coat proteins do not exhibit a preference for a particular radius of curvature, the longer genomes require larger capsids. But, if it is allowed by the radius of gyration (measurement of the compactness of protein structures – radius of gyration of an branched polymer is smaller than that of an linear chain of equal degree of polymerization) of genome, which is dependent on the secondary and tertiary structures of RNA and its charge density, long genomes can be encapsulated by a small capsid. The experiments also show that this correlation can be changed by the ratio of protein to genome concentrations - this is caused by entropy effect which are not yet analyzed in the literature. If, however, capsid proteins have a preference to come together at such an angle as to achieve certain curvature or size, deviation from this curvature cannot be at a cost of preventing the capsids from forming other T numbers in addition to their native ones.
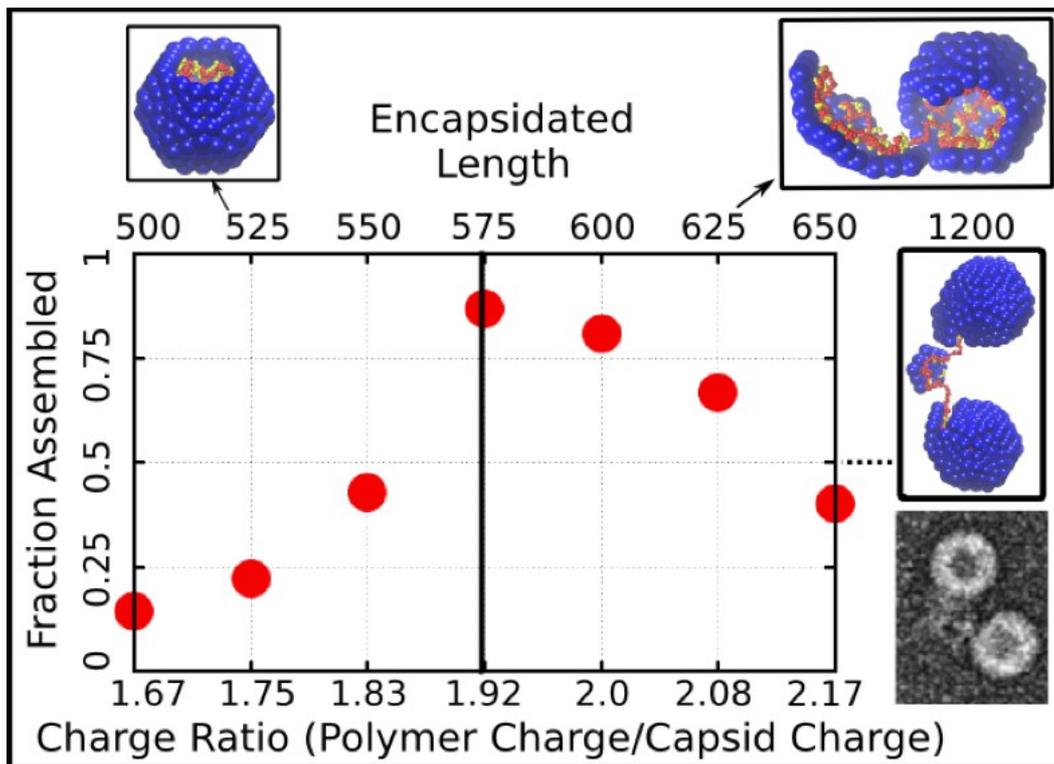
Figure 6: Fraction of trajectories leading to a complete capsid as a function of polymer length (top axis) or charge ratio (bottom axis). The dashed line indicates the thermodynamic optimum charge ratio or length. Snapshots of typical outcomes above and below the optimal length are shown. On the right we can see a comparison between a typical assembly outcome for polymer length 1200 and an EM image of CCMV proteins assembled around an RNA which is twice the CCMV genome length. Beads are colored as in figure 1. [16]

Simulation show that the presence of a polymer is essential for capsid assembly under the simulated conditions, since the subunit-subunit interactions are too weak for formation of empty capsids, which is consistent with most known ssRNA virus proteins. It was also shown in the simulations that the length of a polymer is restricted by the type of capsid it enters. If too long, the polymer will not be incorporated whole when the assembly of capsid is near completion. Instead, part of it will be outside the capsid, or if it will be long enough, multiple capsids will assemble around the same polymer (figure 6). For polymers whose length is well below the thermodynamic optimum length, encapsulation occurs before the assembly of the capsid is complete which substantially slows down the addition of remaining subunits.

One of the important results shown by simulations is that, when looking at the single stranded RNA, increasing the fraction of nucleotides that are base-paired in relation to the unpaired ones (up to the biological fraction of 50%) increases the thermodynamic optimum length. That increase can be as large as 200-250 nucleotides for a small capsid such as T =1, which indicates that base-pairing can contribute significantly to the amount of polymer that can be packaged. Simple models have shown that spontaneous overcharging takes place during the assembly of viral shells, and that the optimal thermodynamic polyelectrolyte length is closely related with the length for which the highest yield of complete viral particles can be obtained by the process of dynamical assembly.

Despite the fact that, for example, for the T=7 capsid the interactions between the coat proteins were expected to be stronger than those in a T=4 structure, the interactions between proteins and

genome can easily bring down the energy of a T=4 structure significantly below that of a T=7 structure. The example of this can be seen in figure 7, for the case of the shorter genome. This is one example of how the genome size can have an effect on regulating the size of a capsid, if coat proteins are flexible enough to form different T structures, such as CCMV.
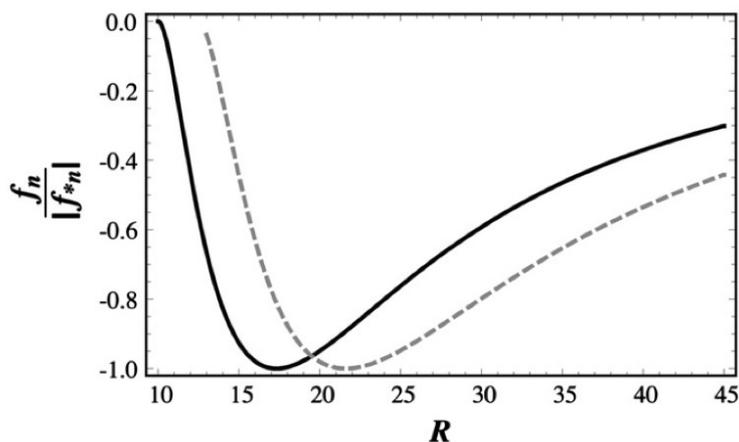

Figure 7: Free energy per protein subunit fn scaled to its minimum value f*n versus the capsid radius R in nanometers. The values of the combination of parameters (nMn/gb) for the corresponding optimal radius, were chosen such as to obtain the radii corresponding to typical T ¼ 4 and T ¼ 7 viruses. We set the molecular weight corresponding to the dashed curve for the T ¼ 4 structure 1.6 times higher than that of the solid curve for the T ¼ 7 structure, while keeping all the other parameters constant. [17]

Influence of RNA Branching

It is clear, however, that capsid size or even capsid shape are not determined by the genome length alone. Because of the secondary and tertiary structures of the nucleic acids, what is also important for the successful capsid assembly is the three-dimensional shape and size of the RNA. Recently, some studies have indicated that the secondary structure of viral single stranded RNAs has to be different from the structure of random or ribosomal RNAs – in viral RNAs, the widest distance between two points on a chain is still smaller by as much as one third than in the random or ribosomal ones. The reason for this may be evolutionary pressure which forced the viral RNAs to be more compact in order to prevent accidental encapsulation of non-viral RNAs – and this provides a possible answer as to why virus capsid "favors" its own RNA more than the other ones present in the host cell.

When the capsid size is largely determined by the properties of the coat proteins, the level of branching in the genome doesn't have a big effect on the optimal number of monomers that are encapsulated. Despite of that, the level of branching of the genome does seem to have a significant effect on the free energy of encapsulation and consequently on the critical concentration necessary for the encapsulation. This is the reason that a simple change in the order of genes can have large effect on the encapsulation of RNA – it changes the critical concentration, as it is exponentially dependent on the free energy.

Inherent propensity to form branch points can be quantified by the fugacity, fb - the chain is linear if fb = 0, and becomes more branched as fb increases. Some RNAs are encapsulated more efficiently than others. With larger inherent propensity to form branch points, the larger is the optimal chain

length that can be accommodated in the capsid. Free energies associated with branched polymers have deeper minima than those for linear polymers for a set of salt concentrations. This effect explains why some RNAs are encapsulated more efficiently than other RNAs or other linear polyelectrolytes. In the process of assembly of viroidal particles inside the cell, viral RNA is in competition with cellular RNA to fill the virus capsids. Inherently branched secondary structure of viral RNA allows it to maximize the amount of encapsulated genome and make assembly more efficient, and in that was it can out-compete cellular RNA.

Experiments in witch RNA was modeled as a simple polyelectrolyte chain showed that branching is also important for efficiency of polymer adsorption on the inner surface of viral shell - the branched polymer is adsorbed more densely onto the surface than the linear chain. Viral RNAs posses inherent branching in secondary structure, which allows viruses to maximize the amount of genetic information stored in them. It also makes assembly of the capsid more efficient, which is why viral RNAs have an advantage over cellular nucleic acids during the replication of virus in its host's cell. Hydrogen bonding between mutually complementary nucleotides along the backbone promotes intra-chain base paring, which leads to a structure of the RNA molecule that is highly branched and that furthermore promotes its compaction in free solution.

Influence of capsid protein concentration

The question is, do genome concentrations and stoichiometry or ratio of the proteins have any impact on the size of the spherical viral shells that form in the solution? The concentration of coat proteins in respect to genome concentration is, in addition to genome size, important for determining the size of the capsid, as shown by the experiments. For a fixed concentration of protein and at low concentrations of RNA the dominant structures are T=7 structures. If we, however, increase the concentration of RNAs, the T=4 will increasingly prevail over others. This tells us that even if for a certain RNA length T=7 structures are energetically more favorable than the structures of T=4, the concentration of free coat proteins in a solution can be low enough that T=4 structures become entropically more favorable compared to T=7 structures. From the mass action equations we can also see that in the opposite case, when the RNA is of such length as to prefer T=4 structure, T=7 structures will nevertheless be formed if the protein concentration is high and the RNA concentration low.

Also important for the virus assembly, and related to the RNA three-dimensional structure, is the gene order in RNA. For example, changing the order of genes in viral RNA can be enough to suppress encapsidation. This can be explained by the fact that the effect of changing the sequence of nucleic acids is modification of it's level of branching – usually, the viruses tend to increase it. This in turn effects free energy and with that the critical concentration needed for encapsulation.

Energy-depended packaging

Virus capsid simultaneously need to be both sufficiently stable in order to provide protection for the genetic material inside them in the extra-cellular environment, and unstable enough to release nucleotides when in the host cell. Assembly around RNAs is predominately driven by electrostatic interactions between RNA phosphate groups and basic amino acids, often located in flexible tails known as arginine rich motifs. There is a correlation between the net charge of these protein motifs

and the genome length for many ssRNA viruses, with a charge ratio of negative charge on NAs to positive charge on proteins typically of order 2:1 (i.e., viruses are 'overcharged').

Packing of genomes into viral capsids can be done in two ways – with and without the use of energy. When it is energy-depended, electrostatic forces play a vital role – viral particles are formed due to interactions between negatively charged viral RNA and positively charged viral capsid protein residues.

Experiments have shown that the protein subunits of many simple RNA viruses can, in certain solution conditions, form a capsid not only around their own RNAs but also around heterologous and non-viral RNAs and artificial linear polyanions. It is thus experimentally confirmed that the interactions between the capsid protein subunits, which are positively charged, and negatively charged genome are largely nonspecific, electrostatic in origin and provide the main driving force for the assembly of viruses. Because the PSS is able to compress, it can adjust to the diameters of capsids that are preferred by the capsid proteins. However, because this only applies to a limited degree, for higher molecular masses of polymer, there are larger VLPs.

Viral capsid optimal thickness is set in such a way that the free energy loss by the elastic compression is exactly compensated by the overall free energy gain from encapsulation. If CPs had a sufficiently strong preference to form capsids with a particular radius of curvature, there would only be one VLP size independent of the overall and relative concentrations of CP and PSS and of the PSS molecular mass. Experiments performed on an anionic polymer poly(styrene sulfonate) (PSS) show that two basic criteria can be set for the formation of a stable capsid. First, for a certain PSS size, the capsid needs to be large enough that any changes in polymer energy and entropy which are related to confinement are not too costly. Second, the capsid needs to be small enough that the PSS is able to neutralize the effects of amino acid residues, which are positively charged, on the inside of the capsid surface without the amino acid being absorption onto that surface (which would lead to a loss of configurational entropy).

For some viruses, for example CCMV, empty capsids can not form in physiological conditions without the PSS or ssRNA. Reason for this is that for T=2 and T=3 empty capsids the energy minima that would correspond to them are higher then chemical potential of unaggregated capsid protein subunits in solution. It's also important to notice that the free energies which are associated with branched polymers have deeper minima than those associated with linear polymers, given the same conditions (same salt concentrations). This provides explanation as to why some RNAs have a higher efficiency of encapsulation compared to other RNAs or other linear polyelectrolytes.

References

Gonca Erdemci-Tandogan, Jef Wagner, Paul van der Schoot, Rudolf Podgornik, and Roya Zandi (2013) RNA topology remoulds electrostatic stabilization of viruses

[1] Michel, J. P., Ivanovska, I. L., Gibbons, M. M., Klug, W. S., Knobler, C. M., Wuite, G. J. L. and 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. 103, 6184–6189

[2] Ivanovska, I. L., de Pablo, P. J., Ibarra, B., Sgalari, G., MacKintosh, F.C., Carrascosa, J. L., Schmidt, C. F. and Wuite, G. J. (2004) Bacteriophage capsids: Tough nanoshells with complex elastic properties; Proc. Natl. Acad. Sci. 101, 7600–7605

[3] Kuznetsov, Y. G. and McPherson, A. (2006) Atomic force microscopy investigation of Turnip Yellow Mosaic Virus capsid disruption and RNA extrusion; Virology 352, 329–337

[4] Anselmetti, D., Dreier, M., Lüthi, R., Richmond, T., Meyer, E., Frommer, J. and Güntherodt, H.-J. (1994) Biological materials studied with dynamic force microscopy; J. Vac. Sci. Technol. B 12, 1500–1503

[5] Lyubchenko, Y. L., Oden, P. I., Lampner, D., Lindsay, S. M. and Dunker, K. A. (1993) Atomic force microscopy of DNA and bacteriophage in air, water and propanol: the role of adhesion forces; Nucleic Acids Res. 21, 1117–1123

[6] Carrasco, C., Carreira, A., Schaap, I. A. T., Serena, P. A., Gómez-Herrero, J., Mateu, M. G. and Pablo de, P. J. (2006) DNA-mediated anisotropic mechanical reinforcement of a virus. Proc; Natl. Acad. Sci. 103, 13706–13711

[7] Plomp, M., Rice, M. K., Wagner, E. K., McPherson, A. and Malkin, A. J. (2002) Rapid visualization at high resolution of pathogens by atomic force microscopy: Structural studies of herpes simplex virus-1; Am. J. Pathol. 160, 1959–1966

[8] Kuznetsov, Y. G., Malkin, A. J., Lucas, R.W., Plomp, M. and McPherson, A. (2001) Imaging of viruses by atomic force microscopy; J. Gen. Virol. 82, 2025–2034

[9] W. H. Roos, R. Bruinsma and G. J. L. Wuite (2010) Physical virology; Nature physics

[10] Yufang Hu, Roya Zandi, Adriana Anavitarte, Charles M. Knobler and William M. Gelbart (2008) Packaging of a Polymer by a Viral Capsid: The Interplay between Polymer; Biophysical Journal Volume 94, 1428–1436

[11] Earnshaw, W. C. and Casjens, S. R. (1980) DNA packaging by the double-stranded DNA bacteriophages; Cell 21, 319–331

[12] Bazinet, C. and King, J. (1985) The DNA translocating vertex of dsDNA bacteriophage; Annu. Rev. Microbiol. 39, 109–129

[13] Black, L.W. (1989) DNA Packing in dsDNA bacteriophages; Annu. Rev. Microbiol. 43, 267–292

[14]Kaiser, D. and Masuda, T. (1973) In vitro assembly of bacteriophage lambda heads; Proc. Natl. Acad. Sci. 70, 260–264

[15] Hohn, B. and Hohn, T. (1974) Activity of empty, headlike particles for packaging of DNA of bacteriophage λ in vitro; Proc. Natl. Acad. Sci. 71, 2372–2376

[16] Jason D Perlmutter, Cong Qiao, and Michael F Hagan (2013) Viral genome structures are optimal for capsid assembly Length and Capsid Size

[17] Roya Zandi and Paul van der Schoot (2009) Size Regulation of ss-RNA Viruses; Biophysical Journal Volume 96, 9–20