Why large icosahedral viruses need scaffolding proteins

Siyu Li[a,1], Polly Roy[b], Alex Travesset[c,d], and Roya Zandi[a]

[a]Department of Physics and Astronomy, University of California, Riverside, CA 92521; [b]Department of Pathogen Molecular Biology, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, United Kingdom; [c]Department of Physics and Astronomy, Iowa State University, Ames, IA 50011-3160; and [d]Ames Laboratory, Iowa State University, Ames, IA 50011

While small single-stranded viral shells encapsidate their genome spontaneously, many large viruses, such as the herpes simplex virus or infectious bursal disease virus (IBDV), typically require a template, consisting of either scaffolding proteins or an inner core. Despite the proliferation of large viruses in nature, the mechanisms by which hundreds or thousands of proteins assemble to form structures with icosahedral order (IO) is completely unknown. Using continuum elasticity theory, we study the growth of large viral shells (capsids) and show that a nonspecific template not only selects the radius of the capsid, but also leads to the error-free assembly of protein subunits into capsids with universal IO. We prove that as a spherical cap grows, there is a deep potential well at the locations of disclinations that later in the assembly process will become the vertices of an icosahedron. Furthermore, we introduce a minimal model and simulate the assembly of a viral shell around a template under nonequilibrium conditions and find a perfect match between the results of continuum elasticity theory and the numerical simulations. Besides explaining available experimental results, we provide a number of predictions. Implications for other problems in spherical crystals are also discussed.

virus | self-assembly | scaffolding proteins | continuum elasticity theory

More than 50 y ago, Caspar and Klug (1) made the striking observation that the capsids of most spherical viruses display icosahedral order (IO), defined by 12 five-coordinated units (disclinations or pentamers) occupying the vertices of an icosahedron surrounded by hexameric units (Fig. 1). While many studies have shown that this universal IO is favored under mechanical equilibrium (2–4), the mechanism by which these shells grow, circumventing many possible activation barriers and leading to the perfect IO, remains mainly unknown.

Under many circumstances, small icosahedral capsids assemble spontaneously around their genetic material, often a single-stranded viral RNA (5–9). However, larger double-stranded (ds) RNA or DNA viruses require what we generically denote as the template: scaffolding proteins (SPs) or an inner core (10–15). The focus of this paper is on these large viruses that require a template for successful assembly.

The major difficulty in understanding the pathway toward IO is apparent from the results of the generalized Thomson problem, consisting of finding the minimum configuration for interacting M-point particles constrained to be on the surface of a sphere. Simulation studies show that the number of metastable states increases exponentially with M (16), and only with the help of sophisticated optimization algorithms at relatively small values of M (5, 17–19) is it possible to obtain IO ground states. These situations, typical of spherical crystals, become even more difficult when considering the assembly of large capsids, in which once protein subunits are attached and a few bonds are made, it becomes energetically impossible for them to rearrange: Should a single pentamer appear in an incorrect location, IO assembly would fail.

The combined effect of irreversibility and the inherent exponentially large number of metastable states typical of curved crystals puts many drastic constraints on IO growth. The complexity of the problem may be visualized by the various viral shells illustrated in Fig. 1, characterized by a structural index, the T number (1, 20–22) $T = h^2 + k^2 + hk$, with h and k arbitrary integers, such that the crystal includes $60T$ monomers or $10(T-1)$ hexamers and 12 pentamers (disclinations).

A possible mechanism to successfully self-assemble a desirable structure might consist of protein subunits with chemical specificity, very much like in DNA origami (23) where structures with complex symmetries are routinely assembled. In viruses, however, capsids are built either from one or from a few different types of proteins, so specificity cannot be the driving mechanism leading to IO (9, 18, 21, 24, 25). In this paper, we show that a "generic" template provides a robust path to self-assembly of large shells with IO. This is consistent with many experimental data in that regardless of amino acid sequences and folding structures of virus coat and/or scaffolding proteins, due to the "universal" topological and geometrical constraints, large spherical viruses need scaffolding proteins to adopt IO (Fig. 1). Although the focus of this paper is on virus assembly, the implication of our study goes far beyond that and extends to many other problems where curved crystals are involved, a point that we we further elaborate in *Conclusion* (26, 27).

The distinct feature of spherical crystals is that their global structure is constrained by topology. More concretely, if $s(\mathbf{x})$ is the disclination density, then

$$\int d^2\mathbf{x}\, s(\mathbf{x}) = 2\pi\chi, \qquad [1]$$

where χ is the Euler characteristic ($\chi = 2$ for a sphere). However, a capsid closes only at the end of the assembly, and thus

---

Significance

Despite a plethora of experimental data on the role of scaffolding proteins in the structure of viral shells, no theoretical/numerical explanations have been put forward to decipher the indisputable need of large shells for scaffolding proteins. This paper explains the underlying physical mechanisms for the formation of large viral shells and elucidates the "universal" role of scaffolding proteins in the formation of large spherical crystals.

PHYSICS
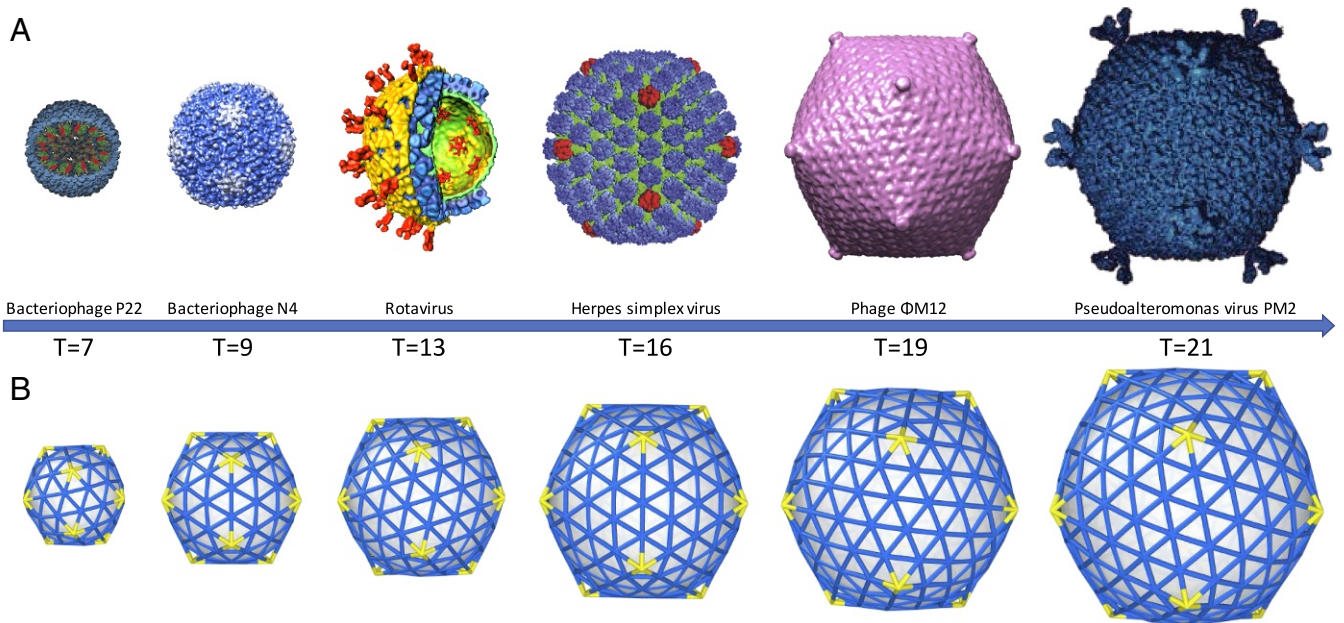
BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Fig. 1. (*A*) From *Left* to *Right*: bacteriophage P22 (28), bacteriophage N4 (29), rotavirus (30), herpes simplex virus (31), phage ΦM12 (32), and pseudoalteromonas virus PM2 (33). The triangulation number of each virus is shown below it. The scaffolding proteins and hydrogenases inside the capsid of bacteriophage P22 and the inner shell of rotavirus are shown. To form structures with IO, all viruses shown need scaffolding proteins as illustrated for bacteriophage P22. Only rotavirus requires a preformed scaffolding layer. Rotavirus belongs to the Reoviridae virus family, and these viruses all form $T = 13$ and have multishell structures. Bacteriophage P22 reprinted with permission from ref. 28, Springer Nature: Nature Chemistry, copyright (1995). Rotavirus reprinted from ref. 30. Copyright (2003), with permission from Elsevier. Herpes simplex virus from ref. 31. Reprinted with permission from AAAS. Pseudoalteromonas virus republished with permission of Microbiology Society, from ref. 33; permission conveyed through Copyright Clearance Center, Inc. Bacteriophage N4 and phage ΦM12 are rendered in UCSF Chimera, based on Electron Microscopy Data Bank accession nos. EMD-1472 and EMD-5718. (*B*) Capsids obtained in the simulations from *Left* to *Right*: $T = 7$, $T = 9$, $T = 13$, $T = 16$, $T = 19$, and $T = 21$.

Eq. **1** does not really restrict the number of disclinations during the growth process, as pentamers or other disclinations may be created or destroyed at the boundaries. For a complete shell, the easiest way to fulfill Eq. **1** is with 12 $q = +\frac{\pi}{3}$ disclinations, and this is the case we follow hereon.

A minimal model for spherical crystals consists of a free energy

$$F_c = \int d^2\mathbf{x}\left[\mu u_{\alpha\beta}^2 + \frac{\lambda}{2}(u_{\alpha\alpha})^2\right] + \frac{\kappa}{2}\int d^2\mathbf{x}\left(H(\mathbf{x}) - H_0\right)^2$$
$$\equiv F_c^l + F_c^b, \quad\quad [2]$$

where $u_{\alpha\beta}$ is the strain tensor. The coefficients μ, λ are the Lame coefficients, which depend on the microscopic underlying interactions. Here $H(\mathbf{x})$ is the extrinsic curvature of the template, H_0 the spontaneous curvature, and κ the bending rigidity. By integrating the phonon degrees of freedom, we can recast the term F_c^l in Eq. **2** as a nonlocal theory of interacting disclinations, with free energy (34)

$$F_c^l = \frac{K_0}{2}\int d^2\mathbf{x}\, d^2\mathbf{y}\left[(K(\mathbf{x}) - s(\mathbf{x}))\, G(\mathbf{x}, \mathbf{y})(K(\mathbf{y}) - s(\mathbf{y}))\right],$$
$$[3]$$

where $K(\mathbf{x})$ is the Gaussian curvature and K_0 is the Young modulus. The disclination density $s(\mathbf{x}) = \sum_{i=1}^{12} q_i\delta(\mathbf{x} - \mathbf{x}_i)$ has as variables the positions of 12 disclinations, each of charge $q_i = \frac{\pi}{3}$. The function $G(\mathbf{x}, \mathbf{y})$ is the inverse of the Laplacian square (34). All previous studies for the model in Eq. **3** have been done for curved crystals without a boundary. In this paper, we provide the necessary formalism to include the presence of a boundary.

A discrete version of Eq. **2** is given by (26, 27, 34, 35)

$$F_d = E_s + E_b = \sum_i \frac{1}{2}k_s(b_i - b_0)^2 + \sum_{i,j} k_b[1 - \cos(\theta_{ij} - \theta_0)]$$
$$[4]$$

with θ_0 a preferred angle, related to the spontaneous curvature H_0. The stretching energy sums over all bonds i with b_0 the equilibrium bond length and the bending energy is between all neighboring trimers indexed with ij. We further assume that there is an attractive force between the trimers and the preformed scaffolding layer (inner core) (see Fig. 4), which, consistent with our minimal model, involves a simple Lennard-Jones potential, $E_{LJ} = \sum_i 4\epsilon[(\frac{\sigma}{r_i})^{12} - 2(\frac{\sigma}{r_i})^6]$ with ϵ the depth of the potential and σ the position of minimum energy corresponding to optimal distance between the center of the core and subunits. In *Methods*, we associate dynamics to these models, which corresponds to following a local minimum energy pathway.

Methods

Discrete Model. The growth of the shells is based on the following assumptions (6, 9, 19, 24): At each step of growth, a new trimer is added to the location in the boundary which makes the maximum number of bonds with the neighboring subunits. This is consistent with the fact that protein–protein attractive interaction is weak and a subunit can associate and dissociate until it sits in a position that forms a few bounds with neighboring proteins. These interactions eventually become strong for the subunits to dissociate and trimer attachment becomes irreversible (5). The attractive interactions between subunits, whose strength depends on electrostatic and hydrophobic forces, are implicit in the model. Note that pH and salt can modify the strength of protein–protein and protein–template interactions and thus the growth pathway.

www.manaraa.com

A crucial step in the assembly process is the formation of pentamers, which occurs only if the local energy is lowered, as illustrated in Fig. 2. After the addition of each subunit or the formation of a pentamer, using the HOOMD package (36, 37), we allow the triangular lattice to relax and to find its minimum energy configuration (20).

The proposed mechanism follows a sequential pathway where trimers (T) attach to the growing capsid ($C \to C'$) according to the reaction

$$T + C \leftrightarrows TC$$
$$TC \to C' \qquad [5]$$

with characteristic rates k_D, k_D' and k_r. The rate $k_D = 2\pi D_T R_T$ is diffusion limited, with R_T the trimer radius and D_T its diffusion coefficient, so that the reaction speed is linear in trimer concentration $v_{TC} = k_D[T]$, k_D' is the detachment rate as the trimer searches for the local minimum, and k_r is the irreversible rate of attachment of the trimer to the capsid. The combined reaction rate is therefore $k_T = \frac{k_r k_D}{k_D' + k_r}$. Once the second reaction in Eq. 5 takes place, there is no possibility for correcting mistakes: If a pentamer forms in the incorrect location, IO is frustrated. With some additional assumptions about the dependence of k_r on the coordination of the growing capsid, it is possible to derive overall rates for capsid formation.

Two important parameters arise in discussing spherical crystals with the model in Eq. 4. One is the Foppl von-Karman (FvK) number (26)

$$\gamma = \frac{b_0^2 k_s}{k_b}, \qquad [6]$$

which measures the ratio of stretching to bending moduli. When the FvK number is large, the protein subunits optimize stretching and bend away from their preferred radius of curvature, showing some degree of faceting, which is the case of large viruses (Fig. 1). For the case of template-driven self-assembly, we introduce a parameter

$$\eta = \frac{k_b}{\epsilon}, \qquad [7]$$

which measures the relative strength of the bending rigidity to the attraction of the trimers to the template. For small η, the proteins follow the core curvature during growth all of the time, regardless of the protein's spontaneous curvature. For large η, the shell detaches from the core and follows its preferred curvature. In this paper we are mostly interested in the regime $\eta \approx \mathcal{O}(1)$ and $\gamma \gg 1$, where the template, rather than the spontaneous curvature, dictates the size of the capsid.

Continuum Model. We now consider the model given in Eq. 3 on a spherical cap with an aperture angle θ_m, so that its geodesic radius is $R_m = \theta_m R$ (see Fig. 4B). The Lame term (F_c^l) in Eq. 3 can then be written as

$$F_c^l = \frac{1}{2K_0} \int d^2\mathbf{x} \sqrt{g} \, (\Delta\chi)^2, \qquad [8]$$

where $g_{\mu\nu}$ is the metric defining the surface and the Laplacian is $\Delta = -\frac{1}{\sqrt{g}} \partial_\mu g^{\mu\nu} \partial_\nu$, with χ the Airy stress function that satisfies

$$\frac{1}{K_0} \Delta^2 \chi(\mathbf{x}) = s(\mathbf{x}) - K(\mathbf{x}). \qquad [9]$$

In *SI Appendix, section S1*, we provide the detailed calculations. We note that approximate solutions of Eq. 9 are available under the assumption that the Laplacian is computed with a flat metric (38), which immediately leads to $\int d^2\mathbf{x} K(\mathbf{x}) = \int \frac{d^2\mathbf{x}}{R^2} = \frac{A}{R^2} = \pi \neq 2\chi\pi = 4\pi$, directly violating the topological constraint in Eq. 1. Therefore previous results (39) are limited to small curvatures or aperture angles ($\theta_m \ll \pi$). The generalization of Eq. 3 to include boundaries proceeds by defining the stress tensor by the expression $\sigma^{\alpha\beta} = g^{\alpha\beta} \Delta\chi(\mathbf{x}) - g^{\alpha\mu} g^{\beta\nu} \nabla_\mu \nabla_\nu \chi(\mathbf{x})$. We now include a stress-free condition $\sigma_{\alpha\beta} n^\beta = 0$ at the boundary, where n^α is the normal to the boundary. For a spherical cap (see Fig. 4B), we use the metric $ds^2 = g_{\mu\nu} dx^\mu dx^\nu = dr^2 + R^2 \sin^2(r/R) d\phi^2$. Note that following the simulation outcomes, we ignore boundary fluctuations. This is mainly because of the strength of protein–protein interactions and line tension implicit in the growth model and is consistent with the simulation results.

With the above definitions, the topological constraint in Eq. 1 is satisfied exactly for a sphere. The free energy in Eq. 3 then becomes

$$F_c^l(\theta_m, \mathbf{x}_i) = E_0(\theta_m) + \sum_{i=1}^N E_{0d}(\mathbf{x}_i, \theta_m) + \sum_{i=1}^N \sum_{j=1}^N \hat{E}_{dd}(\mathbf{x}_i, \mathbf{x}_j, \theta_m) \qquad [10]$$

with E_0 the free energy of the hexamers and E_{0d} the interplay between Gaussian curvature and pentamers, and $\hat{E}_{dd}$ describes disclination (pentamer) interactions. It is convenient to separate this last term as

$$F_c^{dd} = \sum_{i=1}^N \sum_{j=1}^N \hat{E}_{dd}(\mathbf{x}_i, \mathbf{x}_j, \theta_m) = \sum_{i=1}^N E_{self}(\mathbf{x}_i) + \sum_{i=1}^N \sum_{j>i}^N E_{dd}(\mathbf{x}_i, \mathbf{x}_j), \qquad [11]$$

where $E_{self}(\mathbf{x}_i)$ is the disclination self-energy, which depends on the location of a pentamer relative to the boundary.

Results

Consistent with the assumptions describing the dynamics of growth noted in the previous section, we consider the spherical cap (see Fig. 4B) with an aperture angle, which monotonically varies from $\theta_m = 0$ to $\theta_m = \pi$ (sequential growth) as a function of time $\theta_m(t)$. Then, for each value of θ_m we calculate the free energy in Eq. 10 and compare it to the one with an additional new defect (local condition). Once the latter one is favorable, the new defect is added.

For small values of θ_m, the cap grows defect-free. In Fig. 3 we plot the energy of a spherical cap for $\theta_m = 0.7$. The dotted line in Fig. 3 shows the disclination self-energy E_{self}, the dashed line is the Gaussian curvature–disclination interactions E_{0d}, and the solid line is the sum of both energies as a function of the location of disclination in the cap. The diamond in Fig. 3 corresponds to the minimum of energy and indicates the location of the first (and only) disclination appearing in the cap, around $r \sim 0.66R$. This value is very close to the geodesic distance following from the local "screening" of the Gaussian curvature $\int d^2\mathbf{x} s(\mathbf{x}) = \int d^2\mathbf{x} K(\mathbf{x}) \to \frac{\pi}{3} = 2\pi(1 - \cos(\theta_m))$ such that $r = \arccos(5/6)R = 0.59R$. Somewhat counterintuitively, the first disclination does not appear at the center of the cap, which is the result of the competition between the disclination self-energy whose minimum is at the boundary and the Gaussian curvature–disclination interaction E_{0d} with its minimum energy occurring at the cap center; see *SI Appendix, Fig. S2*, where the contour plots of the different elastic free energies as a function of the location of the first disclination, r are shown. As the shell grows, the appearance of a new disclination becomes
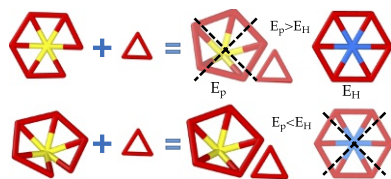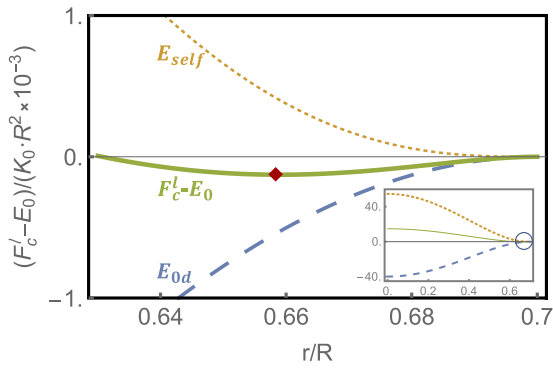

Fig. 2. Dynamics of formation of a hexamer vs. a pentamer. Five trimers are attached at a vertex with an opening angle close to $\pi/3$ at the top and much smaller than $\pi/3$ at the bottom. If the energy per subunit of formation of a pentamer E_P is higher than that of a hexamer E_H, then a hexamer forms (*Top*); otherwise, a pentamer assembles (*Bottom*).

www.manaraa.com

Fig. 3. The 1D energy plot for the first disclination. The dotted line corresponds to disclination self-energy E_{self} (Eq. **11**), the dashed line corresponds to the Gaussian curvature–disclination interactions E_{0d}, and the solid line is the result of the addition of both energies $F_c^l - E_0$ (Eq. **10**) as a function of the location of disclination in the shell for $\theta_m = 0.7$. The energy goes through a minimum for $r = 0.66R$. *Inset* graph shows the zoom-out energy plot where the circle region corresponds to the main graph.

energetically favorable; i.e., a new energy valley for the formation of a new disclination emerges, as illustrated in Fig. 4*B*, where we show the contour plots of total elastic energies for spherical caps with $\theta_m = 0.8$ through $\theta_m = \pi$. The bigger ball in each plot indicates the position of the latest energy well, which is where the addition of the next disclination takes place. Remarkably, both in the continuum model and in simulations, during the growth process, the disclinations always appear in the positions that eventually become the vertices of an icosahedron.

Results from the discrete model in Eq. **4** are shown in Fig. 4*A*. Here again, the disclinations universally appear at the vertices of an icosahedron, in complete agreement with the analytical calculation. The simulations were performed for all values between $T = 7$ and $T = 21$ and in all cases the IO was achieved without a single error. The size of the core in Fig. 4*A* is commensurate with $T = 13$ structures. We note that for these simulations the protein's spontaneous radius $1/H_0$ is much smaller than the core radius, R_c ($R_c H_0 \gg 1$), a point that is discussed in more detail below. *SI Appendix*, Movie S1 illustrates the growth of a $T = 21$ structure, which includes 420 triangles.
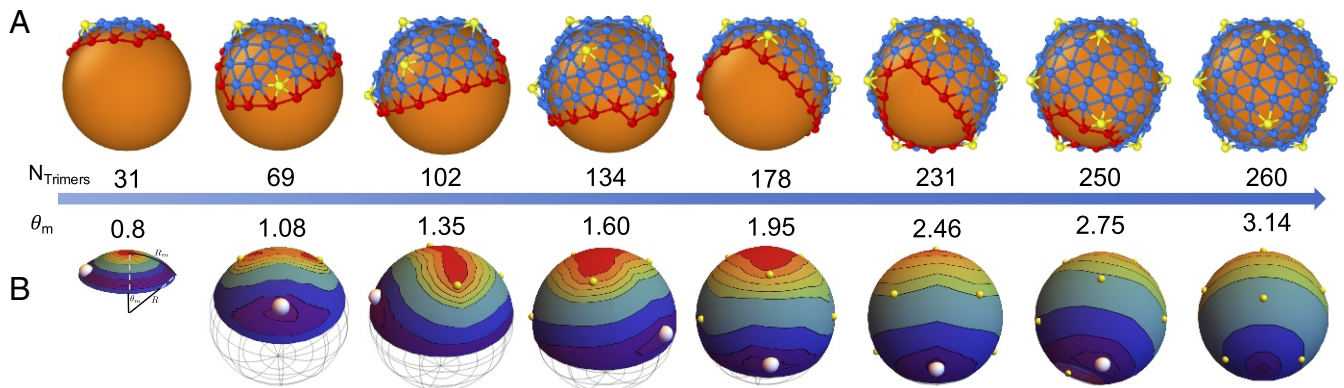
Fig. 5 shows the stretching energy vs. N (number of subunits assembled) as a $T = 13$ shell grows for six different values of FvK number $\gamma > 1$. We note that for large spontaneous radius of curvature and small γ when bending rigidity is dominant, no large icosahedral shell assembles successfully. Rather interestingly, there are conspicuous differences in the dynamics as a function of the FvK parameter γ.

For small values of $\gamma = 2$ (thick black line in Fig. 5) the shell elastic energy grows almost linearly as a function of N but does not show IO. This takes place for higher γ values. The arrows in Fig. 5 indicate a drop in the elastic energy associated with the appearance of pentamers; see *SI Appendix, section S3* for more details. At the beginning of the growth, the shells with different values of γ might follow different pathways and thus the number of hexamers might vary before the first few pentamers form. However, as the shell grows, the pentamers appear precisely at the same place, independently of γ. Note that the bending energy of the shells always grows linearly as a function of number of subunits for any γ (*SI Appendix*, Fig. S3).

Discussion

Our results show that for large shells ($T > 4$) successful assembly into IO requires a nonspecific attractive interaction between protein subunits and a template. This interaction is implicit in the continuum model and is included in the simulations as a generic attractive Lennard-Jones potential. Furthermore, we find that the location of pentamers is completely controlled by the stretching energy as is the case in the continuum elasticity theory.

In the absence of the template, small spherical crystals ($T = 1$ and $T = 3$) assemble spontaneously, for almost any FvK parameter γ. However, as we increase the spontaneous radius of curvature, the final structure depends on the value of γ. For small γ, large spherical shells without any specific symmetry form, and at large $\gamma > 5$ curved hexagonal sheets, which eventually assemble into tubular or conical structures, are obtained. Thus, our results predict that large shells with IO cannot grow without a template.

A template can have a significant impact on the structure and symmetry of the shell. While a weak subunit–core attractive interaction has a minimal role in the shell shape, a very strong subunit–core interaction will override the mechanical properties of proteins. The subunits sit tightly on the template

Fig. 4. The snapshots of a $T = 13$ growth in discrete simulation (*A*) and continuum theory (*B*). The caps in *A* correspond to the simulation growth with triangles representing the trimers. The yellow vertices belong to pentamers, blue ones to hexamers, and red ones to the cap edge. The gold core mimics the preformed scaffolding layer or inner core. The caps in *B* denote the energy contour plots for the newest disclinations that appear in the purple energy well, with geodesic shell size $R_m = R\theta_m$. The red region has the highest energy and the purple region the lowest one. There is a yellow ball in the position of each disclination. The largest ball corresponds to a newly formed disclination.

to form a sphere with no specific symmetry. We were able to observe large shells with IO only for $\eta \sim 1$ but at high γ. In this regime, for pentamers to overcome the core attraction and form in the "correct" position, they must assume a symmetric shape and buckle up (Fig. 1*B*). Indeed, a strong bending energy is needed to overcome the shell adsorption. We find that without decreasing γ (increasing k_b) but with increasing spontaneous curvature, the bending energy associated with the deviation from the preferred curvature of subunits adsorbed to the core becomes strong enough to make the pentamers buckle and assume a smooth shape. Quite interestingly, we find that this is the strategy that Nature has taken to form large shells with IO.

The role of the inner core or the preformed scaffold layer presented above is very similar to the role of SPs, which assemble at the same time as the capsid proteins (CPs); i.e., the template grows simultaneously with the capsid (Fig. 6). In fact, one can think of the inner core as a permanent "inner scaffold" (15). For example, bacteriophage P22 has a $T = 7$ structure, but in the absence of scaffolding (Fig. 1, P22) often a smaller $T = 4$ forms. Similarly, herpes simplex virus makes a $T = 16$ structure but without the SPs, a $T = 7$ assembles. More relevant to the present study is the case of infectious bursal disease virus (IBDV), a dsRNA virus that in the presence of SPs forms a $T = 13$ capsid but in the absence the subunits assemble to a $T = 1$ capsid. This is exactly the condition for formation of the $T = 13$ structure in Fig. 4 where the preferred curvature between subunits is such that in the absence of scaffold they form a $T = 1$ structure. The Reoviridae virus family also forms $T = 13$ but it has multishell structures, which act as inner cores. For instance, in this family bluetongue virus is a double-capsid particle, with an outer (necessary for infection) and an inner capsid (encloses RNA genome). The inner capsid, termed the "core," has two protein layers. The surface layer (or shell) is a $T = 13$ capsid that assembles around the inner shell, a $T = 2$ structure (an inner core). Interestingly, it has been suggested that there is an evolutionary connection between SPs of IBDV and the inner capsid of bluetongue virus (15).

Conclusions

Our model establishes that successful self-assembly of components into a spherical capsid with IO requires a template that determines the radius of the final structure. This template is very nonspecific, and in its absence, protein subunits assemble into either smaller capsids or structures without IO.
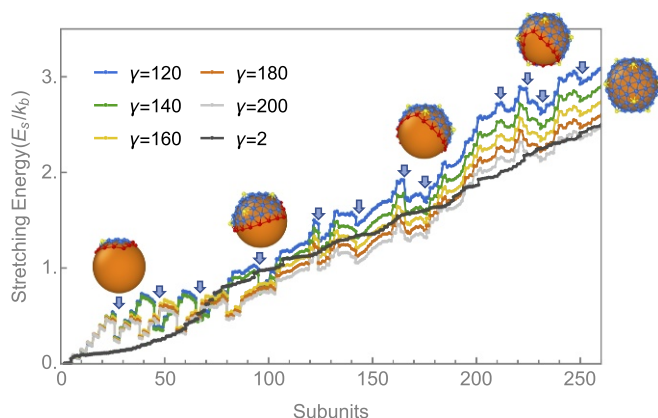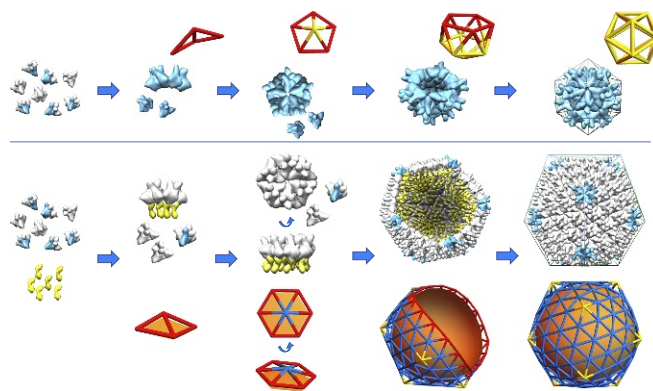

Fig. 6. The role of SPs in the formation of the $T = 13$ capsid of IBDV. Without SPs, the CPs (blue and white subunits) of IBDV form a $T = 1$ structure (*Top*). In the presence of SPs (yellow subunits), they form a $T = 13$ structure (*Bottom*). The results of our simulations are also illustrated next to each intermediate step. Note that SPs (yellow subunits) do not assemble without the CPs but probably experience some conformational changes during the assembly. However, our focus here is solely on the impact of scaffolding on the CPs, resulting in a change in the capsid T number. For preformed SPs, as in the case of bluetongue virus, the core is spherical and there is no indication of any changes in the size of the spherical template during the assembly.

Even though the focus of this study was on the impact of the preformed scaffolding layer, based on the experimental observations we conclude that the SPs, which assemble simultaneously with CPs (Fig. 6), play basically the same role as the inner core in the assembly of large icosahedral shells. Fig. 6 shows that in the absence of SPs, CPs of IBDV form a $T = 1$ structure but when the same IBDV proteins coassemble with SPs (yellow units), a $T = 13$ forms. Fig. 6 also shows the pathway of formation of $T = 1$ and $T = 13$ structures obtained in our simulations. We emphasize that the mechanical properties of subunits are the same for both shells, and the difference in structures arises from the substrate or SPs.

The contribution of the SPs is twofold. The CPs of many viruses including bluetongue virus noted above do not assemble in the absence of SPs. On the one hand, it appears that SPs lower the energy barrier and help capsid subunits to aggregate. On the other hand, by forcing the CPs to assemble into a structure larger than their spontaneous radius of curvature, they contribute to preserving IO.

Examples of the role of templates in the formation of spherical crystals are not limited to viruses, but include crystallization of metals on nanoparticles (40), solid domains on vesicles (41, 42), filament bundles (38), and colloidal assemblies at water–oil interfaces (43). Nevertheless, it has been shown (44) that sufficiently rigid crystals grow as almost flat sheets free of defects, unable to assemble with IO. This regime, however, seems not to be accessible to viral capsids, as the hydrophobic interaction between monomers forces close-packing structures that are incompatible with grain boundaries.

This study sheds light at a fundamental scale on the role of the mechanical properties of building blocks and scaffolding proteins. The proposed mechanism is consistent with available experiments on viruses involving either scaffolding proteins or inner capsids. Further experiments will be necessary to validate many predictions of our described mechanism.

Fig. 5. The stretching energy of a $T = 13$ shell as a function of the number of trimeric subunits. For small FvK numbers ($\gamma = 2$, black line), there is no significant drop in energy as a pentamer forms. However, for large FvK numbers ($\gamma \gg 1$), the formation of pentamers drastically lowers the energy of the elastic shell.

Li et al.

PHYSICS

BIOPHYSICS AND
COMPUTATIONAL BIOLOGY

1. Caspar DL, Klug A (1962) Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp Quant Biol* 27:1–24.
2. Bruinsma RF (2006) Physics of RNA and viral assembly. *Euro Phys J E* 19:303–310.
3. Fejer S, Chakrabarti D, Wales D (2010) Emergent complexity from simple anisotropic building blocks: Shells, tubes, and spirals. *Nano Lett* 4:219–228.
4. Rapaport DC (2004) Self-assembly of polyhedral shells: A molecular dynamics study. *Phys Rev E* 70:051905.
5. Perlmutter JD, Qiao C, Hagan MF (2013) Viral genome structures are optimal for capsid assembly. *Elife* 2:e00632.
6. Comas-Garcia M, Cadena-Nava RD, Rao A, Knobler CM, Gelbart WM (2012) In vitro quantification of the relative packaging efficiencies of single-stranded RNA molecules by viral capsid protein. *J Virol* 86:12271–12282.
7. Sikkema FD, et al. (2007) Monodisperse polymer-virus hybrid nanoparticles. *Org Biomol Chem* 5:54–57.
8. Sun J, et al. (2007) Core-controlled polymorphism in virus-like particles. *Proc Natl Acad Sci USA* 104:1354–1359.
9. Ning J, et al. (2016) In vitro protease cleavage and computer simulations reveal the HIV-1 capsid maturation pathway. *Nat Commun* 7:13689.
10. Thuman-Commike PA, Greene B, Malinski JA, King J, Chiu W (1998) Role of the scaffolding protein in p22 procapsid size determination suggested by t = 4 and t = 7 procapsid structures. *Biophys J* 74:559–568.
11. Earnshaw W, King J (1978) Structure of phage p22 coat protein aggregates formed in the absence of the scaffolding protein. *J Mol Biol* 126:721–747.
12. Aksyuk AA, et al. (2015) Subassemblies and asymmetry in assembly of herpes simplex virus procapsid. *MBio* 6:e01525-15.
13. Saad A, Zhou ZH, Jakana J, Chiu W, Rixon FJ (1999) Roles of triplex and scaffolding proteins in herpes simplex virus type 1 capsid formation suggested by structures of recombinant particles. *J Virol* 73:6821–6830.
14. Saugar I, et al. (2010) Electrostatic interactions between capsid and scaffolding proteins mediate the structural polymorphism of a double-stranded RNA virus. *J Biol Chem* 285:3643–3650.
15. Coulibaly F, et al. (2005) The birnavirus crystal structure reveals structural relationships among icosahedral viruses. *Cell* 120:761–772.
16. Erber T, Hockney GM (1991) Equilibrium configurations of n equal charges on a sphere. *J Phys A Math Gen* 24:L1369–L1377.
17. Morris JR, Deaven DM, Ho KM (1996) Genetic-algorithm energy minimization for point charges on a sphere. *Phys Rev B* 53:R1740–R1743.
18. Zandi R, Reguera D, Bruinsma RF, Gelbart WM, Rudnick J (2004) Origin of icosahedral symmetry in viruses. *Proc Natl Acad Sci USA* 101:15556–15560.
19. Hagan MF, Zandi R (2016) Recent advances in coarse-grained modeling of virus assembly. *Curr Opin Virol* 18:36–43.
20. Wagner J, Zandi R (2015) The robust assembly of small symmetric nanoshells. *Biophys J* 109:956–965.
21. Chen T, Zhang Z, Glotzer SC (2007) A precise packing sequence for self-assembled convex structures. *Proc Natl Acad Sci USA* 104:717–722.
22. Vernizzi G, Olvera de la Cruz M (2007) Faceting ionic shells into icosahedra via electrostatics. *Proc Natl Acad Sci USA* 104:18382–18386.
23. Rothemund PW (2006) Folding DNA to create nanoscale shapes and patterns. *Nature* 440:297–302.
24. Yu Z, et al. (2013) Unclosed HIV-1 capsids suggest a curled sheet model of assembly. *J Mol Biol* 425:112–123.
25. Nguyen T, Bruinsma R, Gelbart W (2006) Continuum theory of retroviral capsids. *Phys Rev Lett* 96:078102.
26. Lidmar J, Mirny L, Nelson DR (2003) Virus shapes and buckling transitions in spherical shells. *Phys Rev E* 68:051910.
27. Vernizzi G, Sknepnek R, De La Cruz MO (2011) Platonic and Archimedean geometries in multicomponent elastic membranes. *Proc Natl Acad Sci USA* 108:4292–4296.
28. Jordan PC, et al. (2016) Self-assembling biomolecular catalysts for hydrogen production. *Nat Chem* 8:179–185.
29. Choi KH, et al. (2008) Insight into DNA and protein transport in double-stranded DNA viruses: The structure of bacteriophage n4. *J Mol Biol* 378:726–736.
30. Pesavento JB, Estes MK, Prasad B (2003) II, 1. structural organization of the genome in rotavirus. *Viral Gastroenteritis, Perspectives in Medical Virology*, eds Desselberger U, Gray J (Elsevier Science, Amsterdam), Vol 9, pp 115–127.
31. Zhou ZH, et al. (2000) Seeing the herpesvirus capsid at 8.5 Å. *Science* 288:877–880.
32. Stroupe ME, Brewer TE, Sousa DR, Jones KM (2014) The structure of Sinorhizobium meliloti phage ϕm12, which has a novel t = 19l triangulation number and is the founder of a new group of t4-superfamily phages. *Virology* 450:205–212.
33. Oksanen HM, et al. (2017) ICTV virus taxonomy profile: Corticoviridae. *J Gen Virol* 98:888–889.
34. Bowick MJ, Nelson DR, Travesset A (2000) Interacting topological defects on frozen topographies. *Phys Rev B* 62:8738–8751.
35. Seung H, Nelson DR (1988) Defects in flexible membranes with crystalline order. *Phys Rev A* 38:1005–1018.
36. Anderson JA, Lorenz CD, Travesset A (2008) General purpose molecular dynamics simulations fully implemented on graphics processing units. *J Comput Phys* 227:5342–5359.
37. Glaser J, et al. (2015) Strong scaling of general-purpose molecular dynamics simulations on GPUs. *Computer Phys Commun* 192:97–107.
38. Grason GM (2012) Defects in crystalline packings of twisted filament bundles. I. Continuum theory of disclinations. *Phys Rev E* 85:031603.
39. Castelnovo M (2017) Viral self-assembly pathway and mechanical stress relaxation. *Phys Rev E* 95:052405.
40. Bao H, Peukert W, Taylor RK (2011) One-pot colloidal synthesis of plasmonic patchy particles. *Adv Mater* 23:2644–2649.
41. Korlach J, Schwille P, Webb WW, Feigenson GW (1999) Characterization of lipid bilayer phases by confocal microscopy and fluorescence correlation spectroscopy. *Proc Natl Acad Sci USA* 96:8461–8466.
42. Guttman S, et al. (2016) How faceted liquid droplets grow tails. *Proc Natl Acad Sci USA* 113:493–496.
43. Bausch AR, et al. (2003) Grain boundary scars and spherical crystallography. *Science* 299:1716–1718.
44. Meng G, Paulose J, Nelson DR, Manoharan VN (2014) Elastic instability of a crystal growing on a curved surface. *Science* 343:634–637.