# Emergent Complexity from Simple Anisotropic Building Blocks: Shells, Tubes, and Spirals

**Szilard N. Fejer, Dwaipayan Chakrabarti, and David J. Wales***

University Chemical Laboratories, Lensfield Road, Cambridge CB2 1EW, United Kingdom

ARTICLE

**ABSTRACT** We describe a remarkably simple, generic, coarse-grained model involving anisotropic interactions, and characterize the global minima for clusters as a function of various parameters. Appropriate choices for the anisotropic interactions can reproduce a wide variety of complex morphologies as global minima, including spheroidal shells, tubular, helical and even head—tail morphologies, elucidating the physical principles that drive the assembly of these mesoscopic structures. Our model captures several experimental observations, such as the existence of competing morphologies, capsid polymorphism, and the effect of scaffolding proteins on capsid assembly.

**KEYWORDS:** virus capsid · self-assembly · mesoscopic · chiral tubes · helices · spirals · shells

The structure of a biological macromolecule generally plays a key role in determining function and is usually achieved through self-assembly of molecular building blocks. Although the majority of viral capsids, assembled from multiple copies of protein monomers or oligomers, are spheroidal with icosahedral symmetry,[1] a rich variety of nonspherical morphologies are also known.[2,3] There are essentially two types of curved surfaces abundant in nature: spheroidal and tubular. Spheroidal virus capsids are often icosahedral in symmetry.[1] Larger assemblies usually have a lower symmetry, for example alfalfa mosaic virus capsids have an oblate ellipsoidal shape,[4] while HIV virus capsids are conical.[2] The number of virus capsid proteins forming icosahedral structures can be rationalized using the concept of quasi-equivalence.[5] Several models used to represent virus capsids[6−9] support geometries corresponding to spherical shells or infinite tubes, while recent molecular dynamics simulations have also found nonicosahedral capsids as byproducts of icosahedral shell formation.[10] The existence of nonicosahedral capsids in nature implies that there must be certain conditions, either kinetic or thermodynamic, that favor these structures over icosahedra. Protein units belonging to different viruses can self-assemble into their "preprogrammed" capsid structure under very similar conditions (concentration, temperature, pH, ionic strength, *etc.*) in a living cell. Nonspherical viruses are often polymorphic, and this property is attributed to the greater flexibility of their capsids.[4] Although it is known that the shape of the building blocks directly influences the overall shape of the assembled structure, the way in which it does so is not understood well enough to allow for efficient bottom-up design of nanostructures.[11] Virus capsid proteins are typical building blocks for which small changes in particle shape and interaction anisotropies introduced by evolution can drive assembly into topologically different shells, while most evolutionary changes do not affect the overall structure of the capsid.

In this contribution we present a simple, yet generic, unconstrained model, based on two-body anisotropic interactions, which contains all the features necessary for self-assembly into a wide variety of shapes found in nature. We show that the main factor determining the shape of such macromolecular assemblies is the directionality and strength of the anisotropic interaction between the building blocks. By identifying and analyzing the most favorable structures for clusters of anisotropic bodies, we also show how the directionality of the interactions influences the overall mesoscopic assembly. Our main objective is to identify universal aspects of the assembly of virus capsid structures. In our model each building block consists of an ellipsoidal core and two "polarized" repulsive Lennard-Jones (LJ) sites within a rigid-body framework.[12] The building blocks represent

*Address correspondence to dw34@cam.ac.uk.

Received for review October 4, 2009 and accepted December 24, 2009.

**Published online January 7, 2010.**
**10.1021/nn9013565**

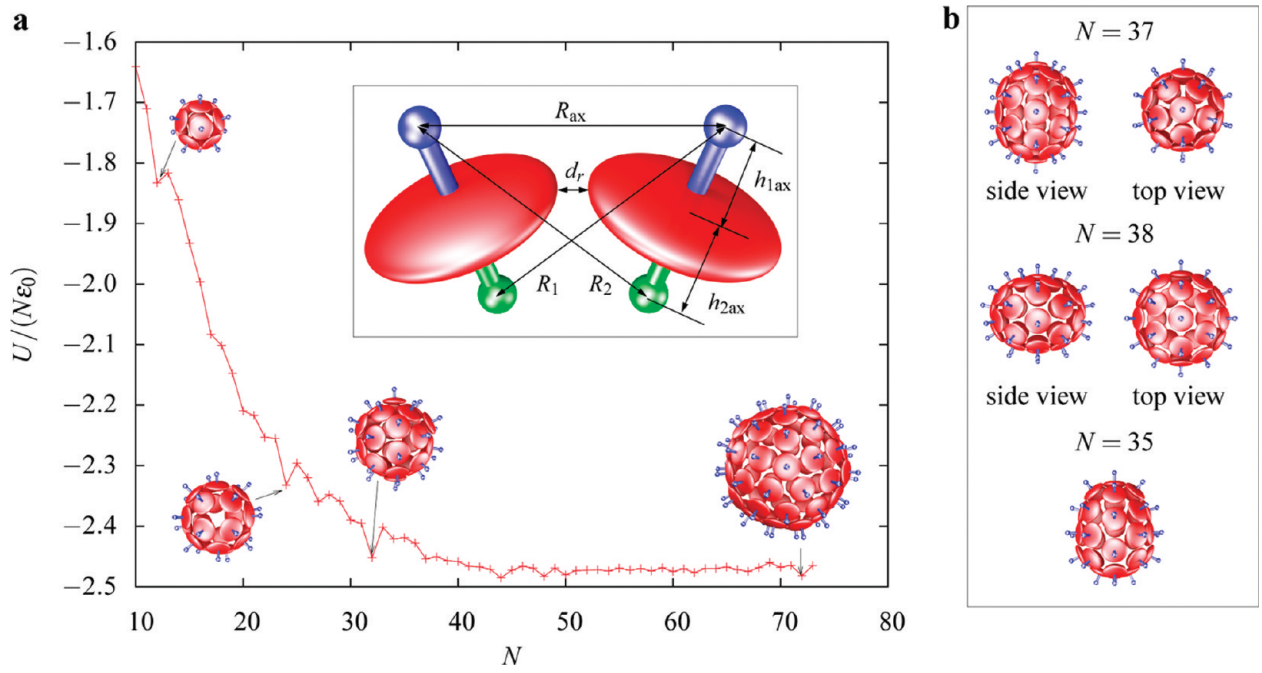


**Figure 1. Global minima for the simplest parametrization of the model: (a) Energy/particle ($U/N$) in units of $\varepsilon_0$ of the global minima as a function of cluster size. Inset: definition of interactions between two anisotropic building blocks with two repulsive sites on either side of the ellipsoid. Primary apex sites are colored in blue. (b) Conical, prolate, and oblate-shaped global minima. Secondary apex sites are only displayed in the dimer structure.**

protein pentamers or hexamers. The preferred geometry of two building blocks interacting with the simplest parametrization considered is shown in the inset of Figure 1a, together with the definition of the intersite interactions.

A key result in the present work is the identification of characteristics in the pair potential that systematically favor particular morphologies. We find a wide range of parameter space supporting spheroidal assemblies analogous to virus capsid shells. Clusters with global minima corresponding to capsid triangulation numbers[5] $T = 1, 3, 4$, and 7 have been identified, as well as octahedral shells, together with nonspherical structures of lower symmetry (Figure 1).

In an experimental system, solvent properties such as pH and ionic strength change the surface charges on the protein subunits, thereby directly affecting the strength and directionality of protein−protein interactions. A phase diagram constructed using these quantities is conceptually analogous to changing appropriate parameters in our model, in a systematic way. We find that such modifications of the anisotropy in the interaction between building blocks are sufficient to drive the assembly of clusters into structures such as tubes and spirals.

## RESULTS AND DISCUSSION

**"Magic Number" Clusters.** In our model the curvature of the assembling shell can be controlled by changing the repulsive character of the LJ sites. Here we present results for the simplest parametrization, using axially symmetric oblate ellipsoids with an aspect ratio of 0.3 to mimic the shape anisotropy of the protein pentamers (and hexamers) that often act as building blocks in capsid assembly.[17,18] The semiaxis lengths of the ellipsoids are therefore fixed to 0.15, 0.5, and 0.5. The repulsive LJ sites are placed on opposite sides along the shortest semiaxis, at a distance of 0.5 from the center (inset of Figure 1a). All distance units here are absolute, which allows us to change the range of the potential explicitly. We have located the likely global minima for clusters containing $2 \leq N \leq 73$ particles for the above parametrization, and Figure 1a shows the energy per particle as a function of cluster size. The diagram has some features in common with the model of Bruinsma *et al.*,[9,19,20] where each particle is confined to the surface of a sphere. However, during capsid assembly, incomplete shells might relax to distorted nonspherical configurations that are lower in energy than spherical geometries. Cluster sizes that are incommensurate with a spherical topology would therefore exhibit artificially high energies when confined to a sphere. Our model shows that the clusters do indeed relax to nonspherical geometries when the number of particles is insufficient to give a complete icosahedral or octahedral configuration. The number of particles in the vicinity of a partially formed capsid in a real system is determined by the (fluctuating) local concentration of building blocks, so the probability of forming capsids with a lower (or higher) number of particles than the ideal value could be rather high.

The interaction energy that we have defined is in fact an effective free energy, since it has been shown that virus capsid self-assembly is usually an

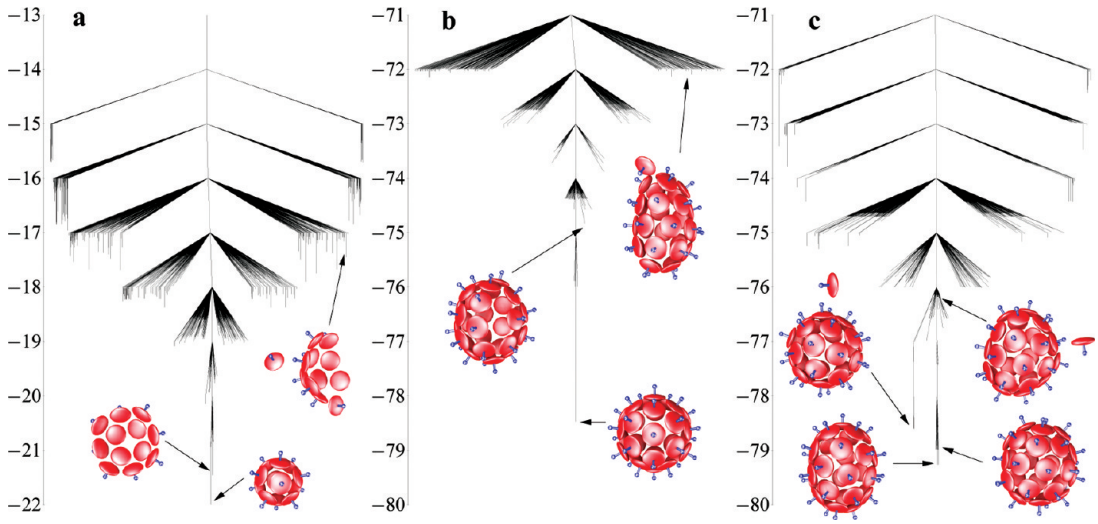


**Figure 2. Disconnectivity graphs constructed for different sized shells: (a) $N = 12$; (b) $N = 32$; (c) $N = 33$. The structures of low-energy minima are also depicted.**

entropy-driven process, in which there is a net increase in entropy during the burial of hydrophobic surfaces involved in capsomer association. This effect is due to the entropy gain of bound water molecules released on association, which overcomes the entropy loss associated with protein aggregation.[21] The average association free energy determined experimentally is typically between $-3$ and $-4$ kcal/mol,[21,22] and this value is generally used to parametrize coarse-grained models based on geometrical considerations.[10,23] Using a value of $-3.5$ kcal/mol, the unit ($\varepsilon_0$) of the absolute energy used in Figure 1 would correspond to 1.5 kcal/mol.

For the parametrization considered here, the following "magic number" clusters have significantly lower energies per particle than their neighboring sizes: $N = 12$, 24, 27, 32, 44, 48, and 72. Structures $N = 12$, 32, and 72 are icosahedral and correspond to capsid triangulation numbers $T = 1$, 3, and 7 in the Caspar−Klug notation.[5] The latter two structures provide a good example of how capsids containing only pentameric units in icosahedral shells adjust to being in different environments:[17] our building blocks overcome the size mismatch between pentameric and hexameric environments by adopting a slightly larger distance from the center for 5-fold coordinated sites. In the results presented in Figure 1 all particles have the same size, and we refer to capsomers in pentameric or hexameric positions based on the number of nearest neighbors. The global minimum for $N = 24$ is an octahedral structure, observed *in vitro* for polyoma virus capsids.[17] To our knowledge, the only other model that supports the $N = 24$ octahedral capsid geometry as a low-energy minimum, without restricting the building blocks to the surface of a sphere, is our multisite rigid-body model based on pentagonal pyramids as building blocks.[14]

Interestingly, the energy gap between the global and second-lowest minima for the "magic number" cluster sizes $N = 27$, 44, and 48 is between 1 and 2 orders of magnitude smaller than for $N = 12$, 24, 32, and 72, suggesting that the former shells are more flexible, and hence capsids with such sizes are likely polymorphic. We also find relatively small gaps between low-energy minima for the sizes adjacent to highly symmetric structures (*e.g.*, clusters $N = 31$ and 33 would be more polymorphic than $N = 32$). We suggest that the flexibility of capsids assembled from an incorrect number of particles makes it easier for new building blocks to incorporate into the shell, therefore facilitating capsid growth.

Using the discrete path sampling technique,[24,25] we mapped out the potential energy landscape for cluster sizes $N = 12$, 32, and 33 in order to gain insight into the size-dependent kinetics of assembly. The three disconnectivity graphs[26−28] constructed are shown in Figure 2. For $N = 12$, the graph corresponds to the characteristic "palm tree" motif associated with efficient self-assembly over a wide range of temperature.[27] The two lowest-energy minima are the icosahedral structure and a partial shell reminiscent of half the octahedral "snub-cube" structure, which is the global minimum for 24 particles. The fastest path between the $T = 1$ shell and the partially formed shell involves two low-energy transition states, suggesting that for this parametrization, the growth of the $T = 1$ capsid is not kinetically hindered. The disconnectivity graph for $N = 32$ also has a funnel-like shape, with the $T = 3$ icosahedral structure well separated in energy from the other minima on the landscape. However, for $N = 33$, there are two distinct sets of low-energy structures separated by a large barrier. The prolate global minimum can interconvert with other low-energy minima belonging to the same basin *via* relatively small barriers, and is therefore quite flexible. The set of low-energy minima that is separated by a large barrier from the global minimum contains structures with a capsomer loosely bound to the icosahe-

dral $N = 32$ shell. Therefore, on the landscape for 33 particles, the assembling shell can easily be trapped in a $T = 3$ icosahedral + 1 capsomer configuration.

On the basis of our results, the proposed mechanism for efficient formation of highly symmetric shells from misassembled structures is an overall downhill process: incomplete, but closed shells require little activation energy to incorporate more capsomers, until a particularly stable structure (containing a magic number of particles) is formed. From that point onward, shell growth is disfavored, since incorporating new capsomers has a significantly larger activation energy in this case. The flexibility of shells containing particle numbers that are incommensurate with highly symmetric configurations enhances the self-assembling character of the capsid assembly process, and drives it toward larger assemblies. As a direct consequence of the multiple funnel-like shape of the overall energy landscape for our model systems, once an incorrect shell is assembled, it will grow until it reaches the bottom of the catchment basin for the next shell with a magic number of particles. We emphasize that by incorrectly assembled shell we mean closed shells that are the lowest-energy configurations for a certain number of particles, and the flexibility of these structures enhances further shell growth. This result contrasts with previous models that produce more rigid assembling shells, because they employ more strictly defined orientational interactions for the building blocks.[6,10,23] Once a correctly formed shell emerges, however, it is stable for a wide range of temperature. Incorporating extra particles in a stable shell is kinetically hindered, as shown in Figure 2c. The ratio of capsid sizes present in a real system is dependent on the temperature and concentration of particles, factors that control the relative abundance of nucleation and shell growth events.[10,23]

The $N = 27$ cluster is roughly spherical, with $D_3$ point group symmetry. Polymorphic spheroidal capsids assembled from 27 units (12 pentamers and 15 hexamers) have been found experimentally.[4,29] A highly symmetrical structure with 27 particles was previously found to be a magic number as well in a model based on cone-shaped building blocks.[30] We have observed that decreasing the anisotropy of the building blocks results in global minima of higher symmetry than for the parametrization considered in Figure 1, corresponding to the symmetries of the Thomson problem for like charges confined to the surface of a sphere[31,32] for the same number of particles. More details about the observed properties, which suggest a high degree of polymorphism for these cluster sizes, can be found in the Supporting Information. We also observe global minima with oblate and prolate, cone-like shapes, and even biaxial geometries for cluster sizes that are incommensurate with a completely symmetric capsid; examples are shown in Figure 1b.

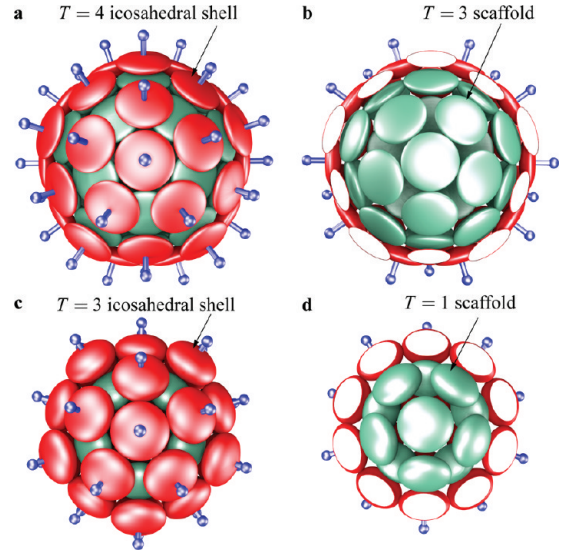


**Figure 3. Two-layer assemblies: (a) Global minimum of a cluster containing 32 particles of type A (green, inner layer) and 42 particles of type B (red, outer shell). (c) Global minimum for a 44-particle system (12 particles of type A, colored in green and 32 particles of type B, colored in red). The outer layer here is a $T = 3$ shell. Half of the outer shells are removed in structures b and d to reveal the icosahedral $T = 3$, and $T = 1$ inner layers, respectively. The primary apex sites of the inner shell are not displayed for clarity.**

**Scaffolding.** It is important to note that the global minimum for the triangulation number $T = 4$ ($N = 42$) is not icosahedral, but has a lower $D_{5h}$ symmetry for the parametrization considered in Figure 1. This result agrees with the findings of Zandi *et al.*[33] and other studies.[30,32] However, the icosahedral structure for our model is a low-energy local minimum for various parametrizations. We find that the global minimum is icosahedral if two different sized (12 "smaller pentameric" and 30 "hexameric") building blocks are used, and the repulsive interaction is longer range ($r^{-6}$ instead of $r^{-12}$).

In nature scaffolding plays a major role in the correct formation of certain virus capsids.[34,35] The existence of a scaffolding protein might guide the assembly of capsid proteins into a well-defined geometry, possibly selecting configurations that would not otherwise be kinetically accessible. To probe how scaffolding affects the assembly of $T = 4$ capsids for our model, we have designed a binary system using the same building block geometry as for Figure 1. The difference between the two types of particles is that the repulsive sites of particle type A do not interact with the repulsive sites of particle type B ($\varepsilon_{rep}^{AB} = 0$). We have located the global minimum of a 74-particle binary cluster parametrized in this way (starting from 32 type A and 42 type B particles in a random configuration), and found that it has icosahedral symmetry (Figure 3a,b), containing an inner $T = 3$ shell, which acts as a scaffold for the correctly assembled outer $T = 4$ shell. Moreover,

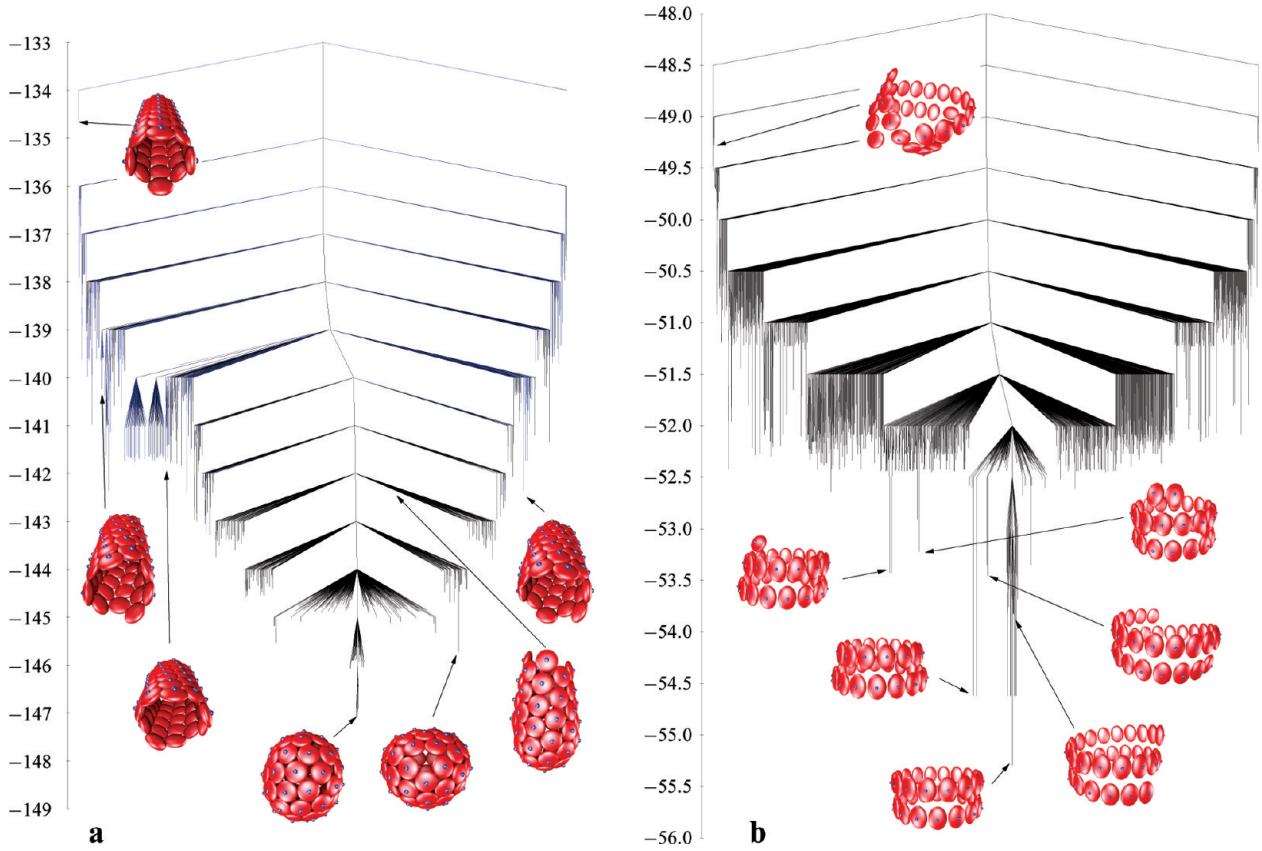


**Figure 4. (a) Disconnectivity graph for $N = 60$ particles with an alternative parametrization that supports low-energy tubular structures as illustrated. The building blocks have been scaled up in size by a factor of 1.5, and the secondary apex sites are not displayed for clarity. Nodes corresponding to likely tubular structures are colored blue. The tubular structures were identified on the basis of the moments of inertia calculated from the center of mass coordinates of the rigid bodies: $|I_y - I_z| < 200$ and $I_x/I_z < 0.8$. (b) Disconnectivity graph for $N = 32$ particles with a parametrization supporting low-energy spiral and stacked disk morphologies, as illustrated.**

we did not find a local minimum that would have an outer shell with $D_{5h}$ symmetry for this binary system.

The size mismatch between a $T = 3$ and $T = 1$ shell composed of the same subunits is larger than that between a $T = 4$ and $T = 3$ shell. To compensate for this effect, efficient packing of ellipsoids can be achieved if their anisotropy is decreased. Figure 3c shows the global minimum for a binary system composed of 12 type A and 32 type B particles, with the ellipsoids having an aspect ratio of 0.6. This system forms an inner $T = 1$ shell that has a $T = 3$ shell wrapped around it in register, with the 5-fold rotation axes of the two shells aligned. The aspect ratio was chosen to be similar to the height/width ratio of retrovirus CA capsid protein pentamers, and the assembled structure is in excellent agreement with recent experimental results[36] showing that double-layered icosahedral structures arise spontaneously during *in vitro* assembly of the protein units. We therefore postulate that the anisotropy of the building blocks (pentamers or hexamers) determines primarily whether double-layered shell structures can form. We did not find icosahedral global minima containing an outer $T = 3$ shell and an inner $T = 1$ scaffold for the aspect ratio used in the original parametrization (0.3). Furthermore, an icosahedral structure for the 74-particle system with less anisotropic ellipsoids (aspect ratio 0.6) is a high-energy local minimum, suggesting that side-by-side contacts between such ellipsoids are lost in the $T = 4$ and larger shells when the anisotropy is decreased. Since close contacts between protein units are actually holding the capsids together, it is likely that multilayered spherical capsid shells form when the contact between protein units in different shells is competitive with side-by-side contacts within the shell. The anisotropy of the building block therefore determines the stability of a particular two-layered structure, and favors only shells of a certain size, as seen in Figure 3.

**Polymorphism.** By changing the interaction anisotropy of the ellipsoids, together with the curvature, we find a large parameter space supporting low-energy minima that are open tubes (Figure 4a). These structures are always higher in energy than spheroidal clusters, but have larger catchment basins (nodes colored blue in the disconnectivity graph), with entropy favoring tubular structures. Retroviruses are typical viruses that have competing morphologies coexisting in solution.[36−38] We find that for these low-energy tubular structures to appear, the main necessary condition is that the interaction along the longer semiaxes of the

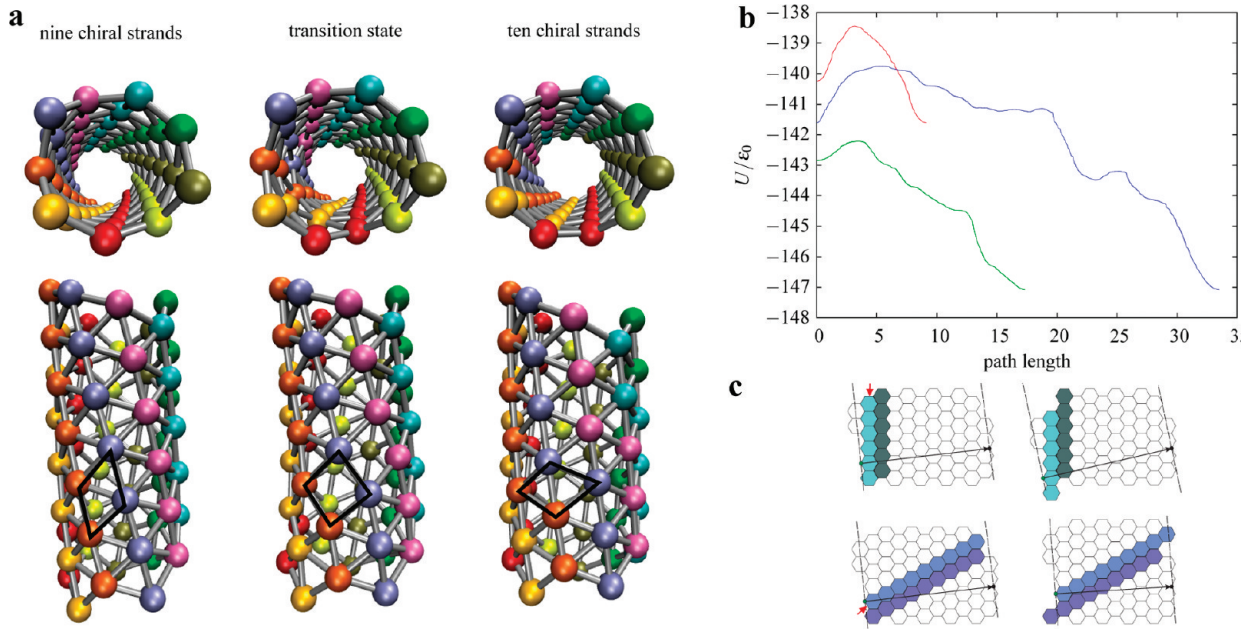


**Figure 5. (a) Structures of the stationary points involved in the pathway shown by a red curve in the b panel. Only the centers of mass of each building block are shown, and each strand in the chiral tube is colored differently for the starting structure. The four building blocks involved in breaking and making new contacts during a diamond−square−diamond type rearrangement are shown with a black polygon. (b) Blue line: fastest pathway from a chiral tubular structure with 10 strands to the spheroidal global minimum for the 60-particle cluster supporting tubular assemblies: (red line) single transition state pathway between two tubular structures (9 and 10 vertical strands); (green line) single transition state pathway from an open conical structure to the global minimum. (c) Schematic view of mechanisms that change the chiral pitch of the tubular structure. A Voronoi construction is used to represent each building block with a hexagon, and the tube is rolled out onto a flat sheet. The direction of the tube axis is shown by the long dashed lines, and the tube is remade after the rearrangement by rolling the sheet up having the green edge overlapping with the black edge, as shown by the arrows. The bonds are broken and reformed between the differently colored strands: (top part) "vertical" sliding mechanism changing the chiral pitch of the tube; (bottom part) diagonal sliding mechanism, changing the number of strands in the tube.**

ellipsoids must be significantly stronger than the "stacking" interaction (along the shortest semiaxis).

Figure 4a shows a disconnectivity graph constructed for a 60-particle cluster that supports low-energy tubular structures, but has a spheroidal global minimum (column "Tubular" in Supporting Information, Table 1). The graph can effectively be split into two distinct regions. The low-energy region around the global minimum does not contain any tubular structures and has a funnel-like topology with only one kinetic trap that corresponds to a low-energy oblate structure having a mirror plane. The higher energy region contains mostly minima identified as being tubular in character (nodes colored in blue in Figure 4a). This region of the graph is frustrated, with low-lying minima separated by large barriers, and several side funnels containing chiral tubular structures of different diameters. We also find structures that have varying curvature and are closed at one end in the main funnel of the spheroidal global minimum. These open conical structures, although lower in energy than tubular assemblies, are intermediate in character, because they can rearrange into more stable, closed structures *via* small barriers. Open conical assemblies can be regarded as precursors to the formation of chiral conical structures, as observed for HIV capsids.[38] In this case, the number of particles (60) is not enough to close the shell while maintaining a varying curvature. The energies on the graph have been rescaled by $\varepsilon = 4.5$ to allow for a quantitative comparison of the barriers connecting the different funnels with those in Figure 2. When rescaling the energies, we assume that the average energy per capsomer in the lowest energy closed shell for this new parametrization is the same as for the simplest parameters (Figure 1), around $-2.5\varepsilon_0$.

HIV CA proteins, when assembling *in vitro*, are known to form chiral tubes of different diameters,[38] with CA hexamers acting as building blocks that pack together in a way that causes hexamers to interdigitate in the tube. Our observation that chiral tubes exist as low-energy minima for a simple model system with axially symmetrical building blocks shows that the fine structure of the HIV CA protein hexamers is not necessary for the chirality to appear. The observed interdigitation of the CA hexameric units might further stabilize the already formed chiral structure, or perhaps fine-tune the chiral pitch.

There are two types of polymorphism encoded in their underlying potential energy landscapes. The structural polymorphism of closed shells containing the same number of particles is dynamic, meaning that these shells can interconvert *via* low barriers (see Figure 2c). On the other hand, the polymorphism observed between spheroidal and tubular structures is the result of assembly processes kinetically trapped into the local minima corresponding to either a tubular or spheroidal

structure. Experimental systems exhibiting dynamic polymorphism include influenza viruses, which have flexible shells. The polymorphic structures observed for retroviral capsid proteins[36,38] are an example of kinetically trapped polymorphism. These structures cannot interconvert *in vitro* because of the high energy barriers separating them. The model system shown in Figure 4a is an interesting example of an energy landscape associated with kinetically trapped polymorphism.

**Cooperative rearrangements between tubular structures.** For the system composed of 60 building blocks, we have located low-lying tubular minima with different chiralities that are separated by relatively large barriers from the global minimum. This region of the landscape is therefore frustrated, and even the interconversion of different tubular structures entails a high activation barrier. Interestingly, pathways between tubes of different chiral pitch and radius were found to be relatively simple, in most cases involving only one transition state. However, the geometrical rearrangements involved are highly cooperative, and can be best described as a succession of diamond−square−diamond rearrangements[39] (see Figure 5).

The mechanisms are illustrated on schematic Voronoi representations of a 60-particle chiral tube rolled out onto a sheet, with the original tubular axis and the particles that move clearly indicated (Figure 5c). During the rearrangement a flat representation of the sheet is broken into two along a line, and the particles belonging to one side slide in plane by one unit, and reform the contact. We have identified two rearrangements, differing only in the direction along which the sheets slide. The "vertical" sliding mechanism (top part of Figure 5c) does not change the number of strands. After the rearrangement 10 strands still exist in the tubular structure, but the angle they make with respect to the tube axis is greater than for the original structure. In contrast, the "diagonal" sliding mechanism depicted in the bottom part of Figure 5c changes the number of strands and does not affect the chiral angle as much as the vertical sliding rearrangement. In our model system, the activation energy for such a rearrangement is around $3\varepsilon_0$, which would correspond to 10.5 kcal/mol in a protein shell. However, the barrier depends on the relative stability of the two tubes. The chiral structure with nine strands shown in Figure 5a needs only $1.5\varepsilon_0$ to convert to the less strained tube with 10 strands. We expect that the activation energy for such rearrangements will increase with the size of the tube, since more contacts would need to be broken in a cooperative fashion. We therefore predict that such sliding mechanisms are likely to occur in short retroviral tubes acting as nuclei during the assembly of rodlike capsids. It is probable that more strained tubes use similar mechanisms to relax into tubes with larger radii during the nucleation phase, thereby reducing the number of tubes with different diameters observed experimentally.[38]

The blue line in Figure 5b shows the energy profile of a pathway with the largest contribution to the overall steady-state rate constant ignoring recrossings,[25] for the interconversion between a tubular structure and the global minimum on the potential energy landscape shown in Figure 4a. The first step of the pathway has a high activation barrier, and results in the disruption of the tubular structure. From this step onward, the path is basically downhill toward the formation of the spheroidal global minimum, with subsequent steps requiring 10−20 times less activation energy than the first, rate-determining step.

In the present model, it is the overall strength of the interaction between building blocks that determines the relative stability of icosahedral (and octahedral) shells and nonspheroidal (oblate, prolate, conical) capsules. A lower interaction strength means higher flexibility. The interaction strength between capsid protein pentamers and hexamers depends on the nature of the contacts, and is system dependent. Since the value for the average interaction energy between two protein units in a capsid is rather small, even a small change in the protein structure can change the interaction strength significantly. When using shorter-range interactions between the ellipsoids, close-packing is more favorable, and the rigidity of the shells is therefore enhanced (data not shown).

**Rings and Spirals.** The shell structures reported above were found using uniaxial ellipsoids for the core of our building blocks. However, it is possible to change the shape of the ellipsoid or the interaction strength along the three main axes. Having a stronger interaction between two ellipsoids in one direction, without any repulsive sites, results in strands as global minima (data not shown). The addition of repulsive sites along an orthogonal semiaxis forces the strands to become curved, and they organize themselves into rings for a small number of particles. The radius of the ring is defined by the curvature that arises for the lowest-energy configuration of two building blocks. Interestingly, for this simple model, single or multiple stranded spirals emerge as global minima for certain cluster sizes, this being a very efficient way for the system to maintain a maximum number of strong side-by-side contacts.

We find that for the parametrization considered here, the morphology of the global minimum for small clusters (whether it is stacked rings or spirals) depends on the number of particles. For example, for $N = 32$ particles, the global minimum is two stacked rings of 16 particles each (Figure 4b), while in a cluster of size $N = 27$, the global minimum is a spiral. The tobacco mosaic virus (TMV) is a well-known example of a self-assembled helical structure that can be transformed into stacked rings by changing solvent conditions,[40,41] which influences the strength of the anisotropic con-

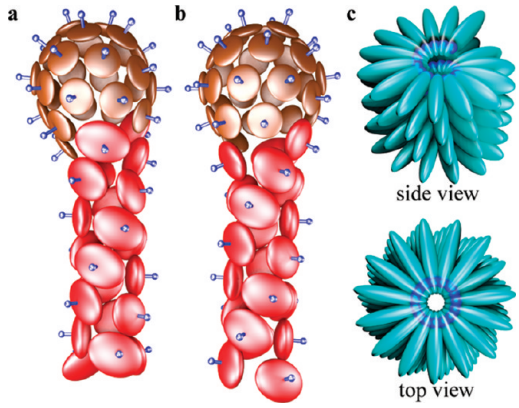


**Figure 6. More complex structures: (a) Global minimum and (b) second-lowest minimum for a binary system containing 29 particles that favor spheroidal shells and 21 particles that assemble into helical tubes when separated; (c) spiral assembly resembling the structure of tobacco mosaic virus, created with a patchy LJ site + ellipsoid model.**

tacts among building blocks. All low-lying minima separated by large barriers in the disconnectivity graph shown in Figure 4b are spirals or stacked disks of different diameters, suggesting that once a spiral structure is formed, it is kinetically unfavorable to change morphology into stacked rings, even if the stacked ring is the global minimum for a certain cluster size, and vice versa. Experimental results on the assembly mechanism of TMV[41] suggest that nucleation has a very important role, with small helical aggregates acting as nuclei. Our model fully supports this observation and gives insight into the kinetic reason why helical aggregates and stacked disks of different sizes are slow to interconvert *in vitro* under physiological temperatures, provided that solvent conditions such as pH or ionic strength are kept constant.

**More Complex Structures.** Using the original three-site model, we have designed a system containing two types of particles that self-assemble into two different morphologies when separated. Particles of type A prefer spheroidal structures, while particles of type B prefer helical tubes. We have set the repulsive interaction between the particles A and B to be slightly stronger than between two particles of type A or of type B, to account for the different shape of the core ellipsoids. The favored morphologies of such assemblies were found to have distinct "head" and "tail" regions, as shown for the two lowest-energy minima characterized for a binary system with 29 + 21 particles in Figure 6a,b. The symmetry of the head region is broken by the attachment of the helical tail, but in some structures certain symmetry elements are retained for the head (see also Supporting Information, movie 4). The observed spontaneous attachment of the tail to a spherical unit suggests that such fusing processes could have played a significant role in the evolution of head−tail viruses, providing the first computational support for the theory of modular evolution in bacteriophages.[42]

The preferred configuration of two building blocks plays a major role in determining the structure of a large cluster. Rings and spirals result when the preferred orientation is along one axis of the ellipsoid, forming a V-shape. The geometry of the building block in the example presented in Figure 4b is different from the shape of the coat proteins in tobacco mosaic virus, yet the observed low-energy structures are very similar in morphology (single-stranded spirals and stacked disk aggregates).

We have also attempted to design a model that supports helical assemblies from particles that are similar in shape to those in TMV. We considered a V-shaped dimer, with one LJ site buried in an ellipsoid, at roughly the same distance from the center of the building block to where the RNA is located in the native structure of TMV. The ellipsoids therefore become "sticky" around the location of the LJ site. The shape of the ellipsoid controls the anisotropy of the stickiness: by increasing the semiaxis length in one direction perpendicular to the longest axis, the ellipsoids tend to bind along the shorter semiaxis. The strength of the LJ interaction controls the effective curvature of the dimer. The structure presented in Figure 6c is a low-energy minimum and has very similar geometrical properties to the tobacco mosaic virus, with 15.8 subunits per turn (TMV has 16.33) and an inner central channel with a radius that is 12% of the value for the outer shell (in TMV this value is 20%). Other low-energy minima identified are stacked rings and tubular structures containing spirals and stacked rings simultaneously, as with the previous model discussed above. These observations further reinforce our conclusion that the preferred geometry of two building blocks determines primarily the shape of a large cluster, not the potential employed.

**General Remarks.** The building blocks of our model can act as both pentameric and hexameric units, therefore we did not consider assembly from protein monomers explicitly, except for our TMV model. Van Workum and Douglas[7] developed a model that incorporates explicitly long-range electrostatic and short-range dispersion forces contributing to the aggregation of protein monomers. They have observed the independent formation of pentameric and hexameric units from rigid triangular building blocks formed by three point dipoles and van der Waals spheres. Such building blocks with discrete 3-fold rotational symmetry were found to support assembly into $T = 1$ icosahedral shells and, under certain conditions, tubular structures could also be grown during their simulation. Their work suggests that multipolar interactions are important in the assembly of nanoscale biological aggregates. Our approach is complementary to the model of Van Workum and Douglas and deals with considerably larger building blocks and their assemblies, using an effective potential dependent on particle shape and interaction anisotropies. We have shown in this work that long-range "soft" in-

teractions are crucial in enhancing the self-assembling properties of closed shells and tubular structures.

## CONCLUSION

In summary, we have created remarkably simple anisotropic building blocks that support assembly of a wide range of mesoscopic structures found in nature, including spheroidal shells, tubular, helical, and even head−tail structures. We have found that nonspherical (oblate, prolate, conical) shells spontaneously appear whenever the number of particles is insufficient or incommensurate with a highly symmetric shell, this being the most efficient way to maximize favorable contacts while minimizing interparticle repulsion. The anisotropy of the building blocks determines in a systematic way the preferred geometry of the dimer, which in turn largely defines the morphology of larger clusters. Our results also suggest that the shape anisotropy of the building blocks is a crucial factor in size selection of the two-layered virus capsid structures observed experimentally. We have also identified a simple cooperative sliding mechanism for changing the diameter and chiral pitch of tubular assemblies, and predict that similar rearrangements are possible to occur in small aggregates acting as nuclei for tube growth. The flexibility of our coarse-grained model makes it possible to study important generic features of viral self-assembly, such as the qualitative effect of a scaffolding protein. The emergent behavior observed shows that efficient self-assembly of a wide variety of nanoscale structures can be obtained for suitable anisotropic interactions. We have demonstrated with this simple model that small changes in the interaction strength and anisotropy can give rise to structures with completely different, but predictable, morphologies. The shape and interaction anisotropies of the model can be tuned based on experimental or computational data, paving the way for further exploration of the dynamics of capsid assembly and interconversion between competing morphologies. The structural trends derived from the model systems considered here can also be used as a starting point for the rational design of nanoscale assemblies with a particular shape.

## METHODS SUMMARY

The interaction energy between the ellipsoids of different bodies employed in the present work is based on their closest approach ($d$, in Figure 1a), as defined by the Paramonov−Yarilaki potential.[13] The repulsive LJ sites are placed on a main axis of the ellipsoid and are located on opposite sides of the plane defined by the other two axes (inset of Figure 1a). There is no interaction between the repulsive LJ sites and the ellipsoidal core. The primary LJ sites interact with sites of the same type and also with the secondary LJ sites in other bodies, while the secondary LJ sites do not interact with each other. We have shown recently that this polarized definition of repulsive sites gives rise to dimer configurations that are similar to the orientation of amphiphilic molecules, which greatly enhances the self-assembling behavior of model pyramids into icosahedral shells.[14] The ellipsoid at the core can be either uniaxial (prolate or oblate) or biaxial. We have focused our investigation on a region of the parameter space where the lowest energy configuration of two building blocks has a side-by-side contact angle lower than 180° for the uniaxial ellipsoids, and in some cases also a preferred orientation along one axis of the biaxial ellipsoids. To reliably locate the global minima for a selected cluster size, basin-hopping global optimization[15,16] was performed starting from random configurations. Further details about the model, our global optimization approach, and the parameters used can be found in the Supporting Information.

*Acknowledgment.* S.N.F. and D.C. gratefully acknowledge support from the Gates Cambridge Trust and Oppenheimer Trust, respectively.

*Supporting Information Available:* Details of the potential used, supplementary discussion, a table of parameters used, four supplementary movies and four supplementary figures (deviation of cluster energies from the best fit, energy difference plot of the global minima and second lowest minima found as a function of cluster size, a graph showing a low barrier and high barrier pathway for incorporating an extra capsomer in an incomplete and icosahedral $T = 3$ shell, respectively, and a figure explaining the analogy of the tubular structures found with carbon nanotubes). This material is available free of charge *via* the Internet at http://pubs.acs.org.

## REFERENCES AND NOTES

1. Flint, S. *Principles of Virology: Molecular Biology, Pathogenesis, and Control*; ASM Press: Washington, D.C., 2000.
2. Benjamin, J.; Ganser-Pornillos, B. K.; Tivol, W. F.; Sundquist, W. I.; Jensen, G. J. Three-Dimensional Structure of HIV-1 Virus-like Particles by Electron Cryotomography. *J. Mol. Biol.* **2005**, *346*, 577–588.
3. Butan, C.; Winkler, D. C.; Heymann, J. B.; Craven, R. C.; Steven, A. C. RSV Capsid Polymorphism Correlates with Polymerization Efficiency and Envelope Glycoprotein Content: Implications that Nucleation Controls Morphogenesis. *J. Mol. Biol.* **2008**, *376*, 1168–1181.
4. Cusack, S.; Oostergetel, G. T.; Krijgsman, P. C. J.; Mellema, J. E. Structure of the Top a-t Component of Alfalfa Mosaic Virus—a Non-icosahedral Virion. *J. Mol. Biol.* **1983**, *171*, 139–155.
5. Caspar, D. L.; Klug, A. Physical Principles in the Construction of Regular Viruses. *Cold Spring Harb. Symp. Quant. Biol.* **1962**, *27*, 1–24.
6. Wilber, A.; Doye, J.; Louis, A.; Noya, E.; Miller, M.; Wong, P. Reversible Self-Assembly of Patchy Particles into Monodisperse Icosahedral Clusters. *J. Chem. Phys.* **2007**, *127*, 085106.
7. VanWorkum, K.; Douglas, J. F. Symmetry, Equivalence, and Molecular Self-Assembly. *Phys. Rev. E* **2006**, *73*, 1–17.
8. Hagan, M. F.; Chandler, D. Dynamic Pathways for Viral Capsid Assembly. *Biophys. J.* **2006**, *91*, 42–54.
9. Guerin, T.; Bruinsma, R. Theory of Conformational Transitions of Viral Shells. *Phys. Rev. E* **2007**, *76*, 061911.
10. Nguyen, H. D.; Reddy, V. S.; Brooks III, C. L. Invariant Polymorphism in Virus Capsid Assembly. *J. Am. Chem. Soc.* **2009**, *131*, 2606–2614.
11. Glotzer, S. C.; Solomon, M. J. Anisotropy of Building Blocks and Their Assembly into Complex Structures. *Nat. Mater.* **2007**, *6*, 557–562.
12. Chakrabarti, D.; Wales, D. J. Simulations of Rigid Bodies in an Angle-Axis Framework. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1970–1976.
13. Paramonov, L.; Yaliraki, S. N. The Directional Contact Distance of Two Ellipsoids: Coarse-Grained Potentials for Anisotropic Interactions. *J. Chem. Phys.* **2005**, *123*, 194111.

14. Fejer, S. N.; James, T. R.; Hernández-Rojas, J.; Wales, D. J. Energy Landscapes for Shells Assembled from Pentagonal and Hexagonal Pyramids. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2098–2104.
15. Li, Z.; Scheraga, H. A. Monte Carlo-Minimization Approach to the Multiple-Minima Problem in Protein Folding. *Proc. Natl. Acad. Sci U.S.A.* **1987**, *84*, 6611.
16. Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101*, 5111.
17. Salunke, D. M.; Caspar, D. L.; Garcea, R. L. Polymorphism in the Assembly of Polyomavirus Capsid Protein VP1. *Biophys. J.* **1989**, *56*, 887–900.
18. Xie, Z.; Hendrix, R. Assembly *in Vitro* of Bacteriophage HK97 Proheads. *J. Med. Biol.* **1995**, *253*, 74–85.
19. Nguyen, T. T.; Bruinsma, R. F.; Gelbart, W. M. Continuum Theory of Retroviral Capsids. *Phys. Rev. Lett.* **2006**, *96*, 1–4.
20. Bruinsma, R. F.; Gelbart, W. M.; Reguera, D.; Rudnick, J.; Zandi, R. Viral Self-Assembly as a Thermodynamic Process. *Phys. Rev. Lett.* **2003**, *90*, 248101.
21. Ceres, P.; Zlotnick, A. Weak Protein−Protein Interactions Are Sufficient To Drive Assembly of Hepatitis B Virus Capsids. *Biochemistry* **2002**, *41*, 11525–11531.
22. Johnson, J. M.; Tang, J.; Nyame, Y.; Willits, D.; Young, M. J.; Zlotnick, A. Regulating Self-Assembly of Spherical Oligomers. *Nano Lett.* **2005**, *5*, 765–770.
23. Nguyen, H. D.; Reddy, V. S.; Brooks III, C. L. Deciphering the Kinetic Mechanism of Spontaneous Self-Assembly of Icosahedral Capsids. *Nano Lett.* **2007**, *7*, 338–344.
24. Wales, D. J. Discrete Path Sampling. *Mol. Phys.* **2002**, *100*, 3285–3306.
25. Evans, D. A.; Wales, D. J. Folding of the GB1 Hairpin Peptide from Discrete Path Sampling. *J. Chem. Phys.* **2004**, *121*, 1080–1090.
26. Becker, O. M.; Karplus, M. The Topology of Multidimensional Potential Energy Surfaces: Theory and Application to Peptide Structure and Kinetics. *J. Chem. Phys.* **1997**, *106*, 1495–1517.
27. Wales, D. J. *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*; Cambridge University Press: Cambridge, U.K., 2003.
28. Wales, D. J.; Miller, M. A.; Walsh, T. R. Archetypal Energy Landscapes. *Nature* **1998**, *394*, 758.
29. Lister, R.; Saksena, K. Some Properties of Tulare Apple Mosaic and Ilar Viruses Suggesting Grouping with Tobacco Streak Virus. *Virology* **1976**, *70*, 440–450.
30. Chen, T.; Zhang, Z.; Glotzer, S. C. A Precise Packing Sequence for Self-Assembled Convex Structures. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 717–722.
31. Thomson, J. J. On the Structure of the Atom: An Investigation of the Stability and Periods of Oscillation of a Number of Corpuscles Arranged at Equal Intervals Around the Circumference of a Circle; with Applications of the Results to the Theory of Atomic Structure. *Philos. Mag.* **1904**, *7*, 237.
32. Edmundson, J. R. The Distribution of Point Charges on the Surface of a Sphere. *Acta Cryst. A* **1992**, *48*, 60–69.
33. Zandi, R.; Reguera, D.; Bruinsma, R. F.; Gelbart, W. M.; Rudnick, J. Origin of Icosahedral Symmetry in Viruses. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15556–15560.
34. Thuman-Commike, P. A.; Greene, B.; Malinski, J. A.; Burbea, M.; McGough, A.; Chiu, W.; Prevelige, P. E., Jr. Mechanism of Scaffolding-Directed Virus Assembly Suggested by Comparison of Scaffolding-Containing and Scaffolding-Lacking P22 Procapsids. *Biophys. J.* **1999**, *76*, 3267–3277.
35. Dokland, T. Scaffolding Proteins and Their Role in Viral Assembly. *Cell. Mol. Life Sci.* **1999**, *56*, 580–603.
36. Cardone, G.; Purdy, J. G.; Cheng, N.; Craven, R. C.; Steven, A. C. Visualization of a Missing Link in Retrovirus Capsid Assembly. *Nature* **2009**, *457*, 694–698.
37. Kingston, R. L.; Fitzon-Ostendorp, T.; Eisenmesser, E. Z.; Schatz, G. W.; Vogt, V. M.; Post, C. B.; Rossmann, M. G. Structure and Self-Association of the Rous Sarcoma Virus Capsid Protein. *Structure* **2000**, *8*, 617–628.
38. Li, S.; Hill, C. P.; Sundquist, W. I.; Finch, J. T. Image Reconstructions of Helical Assemblies of the HIV-1 CA Protein. *Nature* **2000**, *407*, 409–413.
39. Lipscomb, W. N. Framework Rearrangement in Boranes and Carboranes. *Science* **1966**, *153*, 373–378.
40. Durham, A.; Finch, J.; Klug, A. States of Aggregation of Tobacco Mosaic Virus Protein. *Nat. New Biol.* **1971**, *229*, 37.
41. Klug, A. The Tobacco Mosaic Virus Particle: Structure and Assembly. *Phil. Trans. R. Soc., B* **1999**, *354*, 531–535.
42. Botstein, D. A Theory of Modular Evolution for Bacteriophages. *Ann. N.Y. Acad. Sci.* **1980**, *354*, 484–491.