



Why viruses build icosahedra: a quantitative framework for capsid architecture and nanocage design

Chang E

Submitted: May 10, 2026, Revised: version 1, June 19, 2026, version 2, July 10, 2026, version 3, August 13, 2026

Accepted: August 19, 2026

Abstract

Many viruses package their genomes within protein shells (capsids) exhibiting icosahedral symmetry and constructed from recurring pentameric and hexameric units. This study examines whether topological and geometric constraints can help explain the repeated emergence of this architecture across otherwise diverse viral lineages. Using Euler's polyhedron formula and angular-deficit analysis, the study shows that closure of a hexagon-dominated lattice requires twelve pentagonal units, while additional hexagonal units permit expansion without altering the underlying closure topology. Alternative tetrahedral and octahedral solutions can also satisfy shell closure but concentrate curvature at fewer sites and provide less efficient enclosure. To compare these architectural alternatives systematically, five representative capsid architectures were evaluated using five operationally defined criteria: enclosure efficiency, curvature organization, vertex coordination, polygon-class economy, and scalability. Under equal weighting, normalized architectural utility was 0.40 for tetrahedral, 0.50 for octahedral, 0.85 for all-pentamer, 0.95 for prolate pentamer–hexamer, and 1.00 for canonical pentamer–hexamer architectures. Robustness was evaluated by individually omitting criteria, altering criterion weights, and varying the less directly measurable scalability scores by ± 0.5 . The canonical pentamer–hexamer architecture generally remained highest-ranked, although it tied with or, under the strongest asymmetric perturbation, was modestly exceeded by its prolate variant; no other architectural class exceeded either pentamer–hexamer variant. These results support the pentamer–hexamer framework as a particularly favorable geometric solution within the assumptions of the model, combining distributed curvature, efficient enclosure, economical motif use, and architecture-preserving scalability. The analysis provides a testable geometric and organizational explanation for the recurrent use of pentamer–hexamer architectures in viral capsids, while not establishing evolutionary causation or accounting for molecular energetics and assembly kinetics. These principles may also inform the design of synthetic protein cages and other nanoscale delivery systems.

Keywords

Viral capsid, Icosahedral symmetry, Capsid assembly, Pentamer–hexamer architecture, Viral architecture, Self-assembly, Protein nanocages, Computational protein design, Virus evolution, Nanotechnology

Ethan Chang, Aragon High School, 900 Alameda de las Pulgas, San Mateo, CA 94402-3357, USA.
amoferronan@gmail.com

1. Introduction

1.1 Convergent evolution in capsid design

Many viruses, including DNA viruses, RNA viruses, and bacteriophages, independently assemble capsids with icosahedral symmetry despite major differences in genome type, host range, and evolutionary history (1,2). These capsids are typically built from repeating

pentameric and hexameric structural units (3,4). Pentamers occupy the twelve vertices of the icosahedron, whereas hexamers form most of the remaining surface (3–5) (Figures 1A–C). The repeated use of the same capsid proteins also minimizes the amount of genetic information needed to encode the capsid while facilitating efficient assembly.

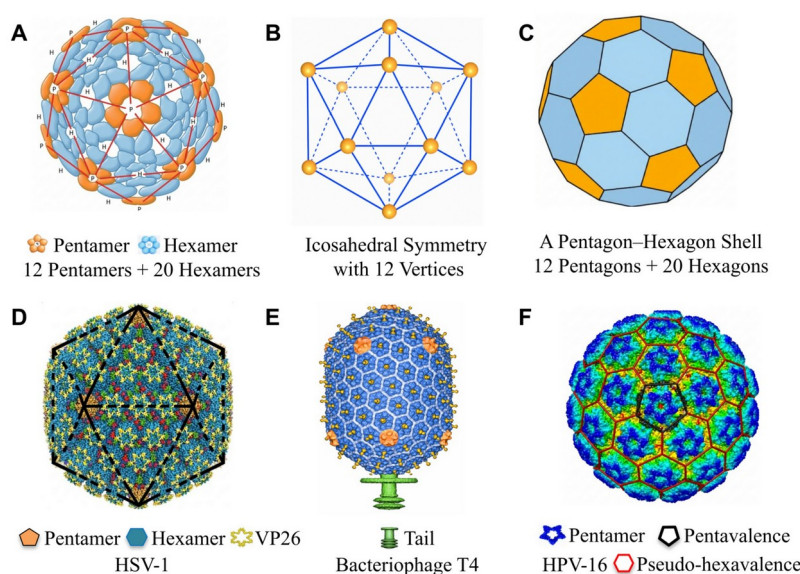


Figure 1. Convergent icosahedral architecture across diverse viral lineages. **(A)** Schematic of an icosahedral capsid assembled from 12 pentameric units (orange; P) and 20 hexameric units (blue; H). Red lines delineate the underlying 20 triangular faces of the icosahedral lattice. Modified from (5). **(B)** Idealized icosahedral geometry showing 12 vertices (orange spheres), each corresponding to a pentameric unit in an icosahedral capsid. Solid and dashed blue lines indicate edges on the near and far sides, respectively. **(C)** Idealized pentagon-hexagon shell with icosahedral symmetry, comprising 12 pentagons (orange) and 20 hexagons (blue). The pentagons occupy the 12 pentavalent positions, while the hexagons fill the intervening lattice. **(D)** Cryo-electron microscopy reconstruction of the Herpes Simplex Virus type 1 (HSV-1) capsid, showing the pentamer-hexamer organization. Twelve pentamers occupy the vertices of the capsid. VP26: small accessory proteins bound to hexamers, decorating the capsid without contributing to geometric organization. Modified from (6). **(E)** Cryo-electron microscopy reconstruction of the prolate capsid of bacteriophage T4. Despite prolate (elongated) geometry, twelve pentamers (orange) remain at vertices, while hexamers (blue) enable expansion. Accessory capsid proteins: yellow. Modified from (7). **(F)** Cryo-electron microscopy reconstruction of the human papillomavirus type 16 (HPV-16) capsid illustrating its all-pentamer architecture. Each blue structural unit is a pentamer assembled from five L1 proteins. The central pentamer (black pentagon) occupies a pentavalent site, whereas the surrounding pentamers occupy hexavalent lattice positions (red hexagons) despite maintaining five-subunit pentameric structures. The yellow network represents the underlying quasi-hexagonal lattice formed by neighboring pentamers, not a separate layer of hexameric proteins. The polygon overlays highlight the local lattice geometry and are not part of the original cryo-EM reconstruction. This organization illustrates how papillomavirus preserves the geometric framework of an icosahedral lattice using pentamers alone. Modified from (8).

Some viruses modify this overall design. Herpes simplex virus type 1 (HSV-1) exemplifies the canonical pentamer–hexamer icosahedral capsid structure (6) (Figure 1D), whereas bacteriophage T4 forms an elongated (prolate) capsid (7) (Figure 1E). Despite its elongated shape, the T4 virus still follows the same geometric principles that govern icosahedral capsid assembly. In contrast, human papillomavirus type 16 (HPV-16) adopts a unique all-pentamer capsid architecture (Figure 1F), in which pentamers are linked through flexible pentamer–pentamer interactions rather than being separated by hexamers. Although composed entirely of pentamers, the capsid retains icosahedral symmetry because most pentamers occupy positions geometrically equivalent to those of hexamers in canonical capsids (8). Despite this distinct organization, HPV-16 still conforms to the underlying principles of icosahedral geometry. Capsids vary widely in size, ranging from tens to hundreds of nanometers in diameter (2). By varying the number and arrangement of hexamers, the pentamer–hexamer capsids can expand or change shape without disrupting the twelve-pentamer topology needed for shell closure.

The recurrent emergence of icosahedral capsids built from pentamers and hexamers across unrelated viral lineages raises the question of which geometric and topological constraints favor this architecture. Structural studies, particularly those enabled by cryo-electron microscopy, have revealed both the remarkable diversity of viral capsids and the recurrence of common geometric principles across unrelated viral families (9). The widespread use of icosahedral capsids suggests that this geometry confers important advantages for genome

packaging and shell stability (10). Similar pentamer–hexamer architectures are found in viruses with diverse capsid proteins (11), suggesting that the framework represents a recurrent solution to common architectural constraints rather than reflecting a shared evolutionary ancestry. The following sections examine the geometric and topological principles that may underlie this convergence. These concepts are then integrated into a quantitative framework for comparing representative capsid architectures using common geometric and organizational criteria.

1.2 Geometric foundations of capsid assembly

1.2.1 *The hexagonal lattice and planar stability*

Hexagonal packing represents one of the most efficient local packing arrangements found in nature. It maximizes the number of nearest-neighbor contacts and minimizes unused space (12,13), while reducing assembly free energy and distributing mechanical stress evenly (14,15). In viral capsids, hexameric organization provides a stable and efficient interaction network that can be assembled from repeated protein subunits. These properties suggest that hexamers may represent more than a geometric convenience; they serve as an efficient structural framework for capsid organization.

Despite these advantages, a purely hexagonal lattice remains flat. In a hexagonal lattice, three hexagons with interior angles of 120° meet at each vertex to form a flat plane of 360° ($= 3 \times 120^\circ$). This configuration yields a flat, zero-curvature state (Figure 2A).

1.2.2 *Topological defects as curvature drivers*

This planar geometry creates a fundamental closure problem because a purely hexagonal

lattice cannot bend and close into a shell without the introduction of curvature. Curvature can be generated by defects that interrupt the planar lattice, analogous to removing a wedge from a flat sheet and reconnecting the exposed edges, thereby forcing the surface to bend (Figure 2B).

Once curvature is introduced, two additional geometric requirements arise: achieving complete shell closure without gaps and permitting variation in shell size and shape while preserving closure.

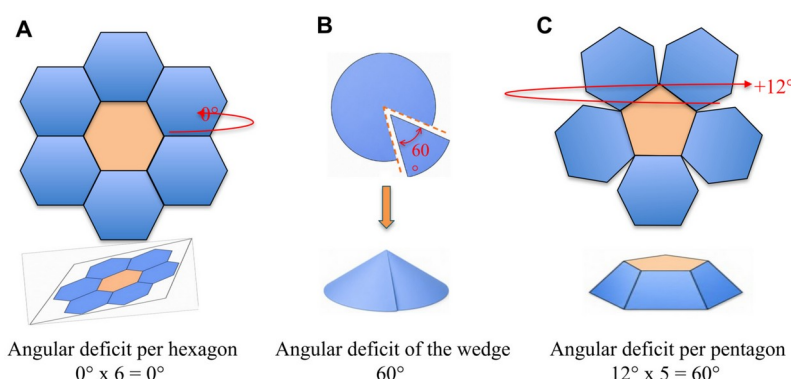


Figure 2. Angular deficit and curvature generation in polygonal lattices. (A) In a purely hexagonal lattice, three hexagons (interior angle 120° each) meet at a vertex, producing a total angle of 360° and zero angular deficit, resulting in a flat surface. (B) Removing a wedge (angular deficit) from a flat sheet allows the sheet to bend into a curved shape. (C) Substitution of a hexagon with a pentagon (interior angle 108°) reduces the total angle at the vertex to 348° , generating a $+12^\circ$ angular deficit per pentagon vertex (shown). This local curvature enables bending of the lattice. Each pentagon contributes a total of $+60^\circ$ curvature across its five vertices, driving global surface closure.

1.2.3 From viral architecture to synthetic design

Viral capsids function as nanoscale delivery structures that package genetic material and facilitate its transfer to host cells. Despite their relatively compact structural organization, viruses can achieve both structural integrity and evolutionary adaptability. Understanding these architectural principles has direct implications for nanotechnology. A major goal in this field is to design nanoscale containers for targeted tissue delivery of diagnostic and therapeutic cargoes, including vaccines.

Although the geometric principles of capsid assembly are well studied, they are often considered separately. Prior work in tiling

theory (16,17) and physical virology (13,15,18) offers a complementary perspective by treating capsids as symmetric elastic shells shaped by both geometric and energetic constraints. Similar architectural solutions are observed across evolutionarily distant viruses, suggesting that geometric constraints may influence capsid organization despite diverse evolutionary histories (10). The present study focuses primarily on geometry and examines the interdependence of topology, curvature, and enclosure in capsid assembly.

2. Theoretical framework for capsid architecture analysis

A geometric and topological framework based on polyhedral surfaces is used to examine viral

capsid organization (19). The capsid is considered as a closed shell assembled from polygonal subunits. The analysis uses an idealized geometric model: polygons are regular, edges have equal length, and three faces meet at each vertex. This model does not consider mechanical effects, such as elasticity, bending energy, or assembly kinetics. Under these conditions, global requirements for closed shells follow directly from Euler's polyhedron formula. In this idealized geometric model, pentagonal and hexagonal polygons represent the lattice positions occupied by pentameric and hexameric structural units, respectively, in canonical icosahedral capsids. The geometric terms *pentagon* and *hexagon* are therefore used when referring to the mathematical model, whereas *pentamer* and *hexamer* refer to the corresponding biological assemblies.

2.1 Euler's polyhedron formula

Topological constraints on closed shells can be derived using Euler's polyhedron formula (20), $V - E + F = 2$, where V denotes the number of vertices, E the number of edges, and F the number of faces. This relation depends only on how the faces and edges are connected, not on geometric detail. Euler's formula limits which polygonal assemblies can form closed shells.

2.2 Angular deficit and curvature

Local curvature is quantified using the concept of angular deficit (21). At each vertex, the interior angles of adjacent faces are summed and compared to 360° , which is the angle around a point in a flat plane. The angular deficit δ at a vertex is defined as $\delta = 360^\circ - \sum \theta_i$, where θ_i are the interior angles of the faces meeting at that vertex. A positive δ produces curvature that allows the surface to bend and close; $\delta = 0$

corresponds to a flat surface; and $\delta < 0$ produces concave, saddle-like curvature.

2.3 Structural determinants of capsid architecture

2.3.1 Topological necessity: Euler's law and the twelve-pentagon rule

In a hexagonal lattice, each hexagon has interior angles of 120° , and three hexagons meet at each vertex (Figure 2A). The total angle around each vertex is $3 \times 120^\circ = 360^\circ$, producing no angular deficit and thus no curvature. As a result, a lattice composed entirely of hexagons remains flat and cannot close into a shell.

Curvature can be introduced when hexagons are replaced by pentagons. When a pentagon replaces a hexagon in the lattice, the vertex transitions to a configuration of two hexagons and one pentagon (Figure 2C). The total angle around such a vertex then becomes $120^\circ + 120^\circ + 108^\circ = 348^\circ$ (a pentagon has interior angles of 108°). This produces an angular deficit of $360^\circ - 348^\circ = 12^\circ$ per vertex. Since each pentagon has five vertices, the total angular deficit is $5 \times 12^\circ = 60^\circ$ per pentagon. This positive angular deficit contributes to the curvature that enables the surface to bend.

Euler's polyhedron formula ($V - E + F = 2$) determines the total curvature required for a surface to bend and close into a shell. The sum of angular deficits (curvature) must equal 720° for any closed surface (20,21). Because each pentagon contributes 60° , closure requires $720^\circ \div 60^\circ = 12$ pentagons. Thus, within a lattice in which pentagons provide the positive-curvature defects in an otherwise hexagonal framework, exactly twelve pentagons are required for shell closure (3). This is observed in icosahedral viral

capsids: twelve pentameric units occupy the vertices of an otherwise hexagonal lattice (2,3). These requirements arise directly from geometry and topology, independent of biological detail.

Although this establishes the necessity of twelve pentagons for closure, it does not explain why pentagons, rather than other polygonal substitutions, yield the observed capsid architectures.

2.3.2 Polygonal substitutions and curvature distribution

Pentagons play a distinctive role in the modeled capsid architecture because alternative polygonal substitutions either oppose shell closure or concentrate positive curvature at fewer sites. Polygons with more than six sides ($n > 6$) generate negative angular deficits when inserted into a hexagonal lattice (Figure 3A).

These negative deficits produce saddle-like (hyperbolic) curvature and therefore oppose the net positive curvature required for shell closure within the present geometric framework. This geometric constraint may help account for the absence of heptagonal and octagonal structural units from known viral capsids organized on hexagon-dominant lattices.

Polygons with fewer than six sides introduce positive curvature. Each polygon type contributes a characteristic total angular deficit across its vertices: triangles contribute $+180^\circ$ per unit, squares $+120^\circ$, and pentagons $+60^\circ$ (Figures 3B–C; Table 1). Because a closed shell requires a total angular deficit of 720° , closure can be achieved using four triangles (tetrahedral symmetry), six squares (octahedral symmetry), or twelve pentagons (icosahedral symmetry) (Figure 4; Table 1).

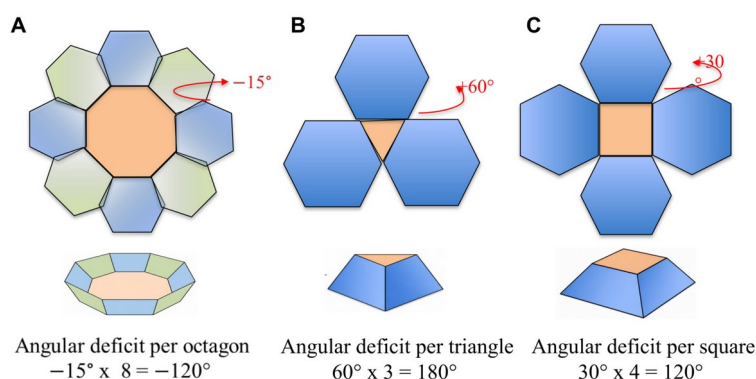


Figure 3. Curvature contributions of polygonal substitutions. (A) Polygons with more than six sides (e.g., octagons) introduce negative angular deficits, producing saddle-like (hyperbolic) curvature that opposes closure. (B and C) Polygons with fewer than six sides introduce positive curvature, with magnitude per polygon indicated: triangles ($+180^\circ$) and squares ($+120^\circ$).

A major difference among these geometrically permissible alternatives lies in the spatial distribution of curvature. Triangles and squares concentrate the required positive curvature at fewer sites—four and six sites, respectively—

because of their larger angular deficits per unit. This concentration of curvature produces sharper local bending and greater geometric distortion. Pentagons contribute smaller angular deficits, distributing the required total curvature

across twelve sites. This broader distribution reduces local curvature concentration and produces a smoother shell that more closely approximates a sphere.

The choice of polygon type therefore affects both the capacity for shell closure and the spatial distribution of curvature. Within the alternatives

considered here, pentagons provide the most spatially distributed curvature while satisfying the topological requirement for closure. This property distinguishes pentagon-based icosahedral geometry from the tetrahedral and octahedral alternatives and provides the basis for the comparison of enclosure efficiency developed in the following section.

Table 1. Geometric effects of polygon substitution in a hexagonal lattice

Permissible Polyhedral Symmetry	Angular Deficit per Unit	Vertex Number (Polygon Type)	Uniform Curvature Distribution	Normalized Surface Area ($S/V^{2/3}$ Value)
Tetrahedron	+180°	4 (triangle)	Poor	~1.49 (7.21)
Octahedron	+120°	6 (square)	Moderate	~1.18 (5.72)
Icosahedron	+60°	12 (pentagon)	High	~1.06 (5.15)
Sphere	—	—	Optimal	1.00 (4.84)

Angular deficit determines the positive curvature introduced into a hexagonal lattice and thereby influences shell closure. Increasing the number of curvature-generating vertices distributes curvature more uniformly, producing a closer approximation to a sphere. Relative surface area values are normalized to a sphere (sphere = 1.00), with lower values indicating greater enclosure efficiency.

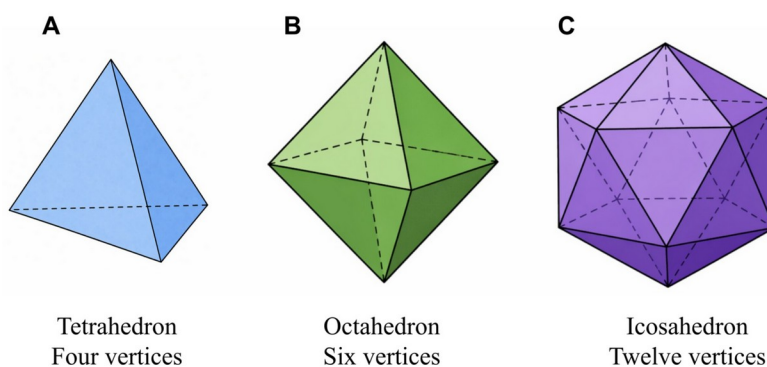


Figure 4. Polyhedral closure geometries from polygonal substitution. (A) Triangle-based curvature yields tetrahedral symmetry (4 units), (B) square-based curvature octahedral symmetry (6 units), and (C) pentagon-based curvature icosahedral symmetry (12 units). These geometries satisfy the requirement that total angular deficit equals 720° for surface closure. Increasing the number of curvature-generating vertices distributes curvature more uniformly, producing shells that better approximate a sphere.

2.3.3 Efficiency of icosahedral enclosure: A spherical approximation

For viruses, enclosure efficiency reflects the cost of forming a closed shell. It can be quantified using the dimensionless ratio $S/V^{2/3}$

(S : surface area; V : enclosed volume). Lower values indicate greater enclosure efficiency, which corresponds to less protein material required to enclose a given volume.

Among all shapes, the sphere minimizes surface area for a fixed volume and represents the theoretical optimum for enclosure efficiency (22). Polyhedral shells therefore approximate this optimum to varying degrees, depending on how evenly curvature is distributed across their surfaces.

For geometrically permissible polyhedral symmetries (Table 1), the normalized ratio $S/V^{2/3}$ (sphere = 1.00) decreases from tetrahedron (~1.49) to octahedron (~1.18) to icosahedron (~1.06), indicating that the icosahedron provides the most efficient enclosure among these polyhedral alternatives. Although a sphere represents the theoretical optimum, viral capsids are assembled from discrete, repeating protein subunits rather than from a continuously curved surface. A principal geometric limitation is therefore that spherical geometry requires continuous curvature, which is difficult to realize using discrete structural units. In addition, a continuously spherical architecture may provide less geometric flexibility for variation in capsid size and shape, as discussed below. The icosahedron

consequently provides a close polyhedral approximation to spherical enclosure while remaining compatible with construction from discrete, repeating protein subunits.

2.4 Design options for viral capsids with hexagonal framework

Euler's constraint further limits the space of permissible symmetries. Only five convex regular polyhedra are geometrically possible, namely tetrahedron, cube, octahedron, dodecahedron, and icosahedron (19). This leaves only a limited number of possible viral capsid geometries. Among these, cubes and dodecahedra are not compatible with closure in a hexagon-dominant lattice. In such a lattice, cube symmetry would require curvature to be concentrated as 90° per vertex ($720^\circ \div 8$ cube vertices); dodecahedral symmetry 36° per vertex ($720^\circ \div 20$ dodecahedron vertices). No regular polygon substitutions within a hexagonal framework can generate these curvatures. These geometries are therefore not observed among known capsids with hexagon-dominant lattices.

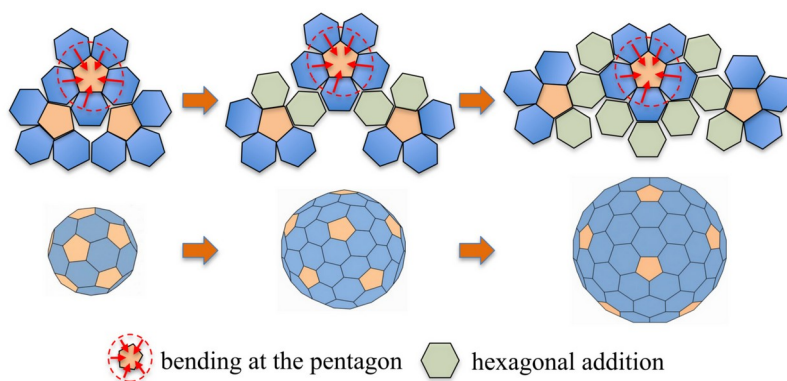


Figure 5. Scalable expansion of icosahedral capsids via hexagonal insertion. A fixed set of twelve pentagons (orange) is embedded within a lattice of hexagons (blue). Addition of hexagonal units (green) increases capsid size without altering global topology. This mechanism enables scalable enlargement and shape modulation while preserving the geometric requirement for twelve pentagonal units.

Tetrahedral and octahedral geometries are compatible with closure. However, these configurations concentrate curvature at fewer sites, producing sharper local distortion and lower enclosure efficiency (Table 1). In contrast, icosahedral geometry distributes curvature across a greater number of sites with minimal local distortion, providing a closer approximation to a sphere.

Beyond curvature distribution, the icosahedral geometry also provides a unique degree of scalability. In tetrahedral and octahedral geometries, shell closure and size are more tightly coupled because curvature is concentrated at relatively few locations (four and six sites, respectively). In pentamer–hexamer systems, a fixed set of twelve pentamers provides the curvature required for closure, while additional hexamers can be incorporated to increase shell size without altering the underlying topology (Figure 5). This closure-scale decoupling enables a common geometric design to support capsids of widely varying dimensions.

2.5 Conserved geometry and capsid variation

Icosahedral symmetry reflects an evolutionary convergence toward a common capsid geometry. This geometry enables shell closure, distributes curvature smoothly, and encloses volume efficiently. Within the pentamer–hexamer framework, pentamers provide the curvature required for shell closure, whereas hexamers permit expansion of shell size without altering the underlying closure topology. This principle is exemplified by the Caspar–Klug triangulation-number (T-number) system, in which progressively larger capsids are generated by adding increasing numbers of

hexameric positions while retaining the same twelve pentamers required for shell closure (3).

This scalability is complemented by quasi-equivalence (3,4), in which identical protein subunits adopt slightly different conformations to form both pentameric and hexameric assemblies. By allowing the same protein to occupy multiple structural roles, quasi-equivalence provides architectural economy. As a result, larger and more diverse capsids can be constructed from the same capsid proteins without requiring new structural components. Quasi-equivalence thus reduces genetic complexity while preserving structural versatility, enabling diverse capsid architectures to emerge from a limited set of recurring building blocks.

Importantly, this conserved pentamer–hexamer architecture is not restricted to a single type of capsid protein fold (a characteristic three-dimensional protein structure). Viruses with structurally distinct capsid protein folds, including the HK97, single-jelly-roll, and double-jelly-roll folds, can all implement the same underlying architectural framework despite differing molecular structures and assembly interfaces (11,23–26). These architectures are stabilized by different combinations of hydrophobic, electrostatic, and hydrogen-bonding interactions (23). Together, the observations demonstrate that the same geometric framework can be realized through diverse molecular interaction mechanisms.

Some viral lineages adopt alternative implementations of the underlying pentamer–hexamer framework. Human Papillomavirus 16 (HPV-16), for example, is built entirely from pentamers (4). In these capsids, twelve

pentamers occupy the vertices, whereas the remaining pentamers occupy positions surrounded by six neighboring units that are geometrically equivalent to those normally occupied by hexamers (Figure 1F) (8). This arrangement is possible because of flexible protein–protein interfaces. Tiling analyses (16,17) show that even these atypical structures still follow the same geometric requirements for icosahedral closure. These examples illustrate that alternative architectural solutions are possible, but they require additional structural adaptations to satisfy the same geometric requirements for shell closure. Such adaptations may involve tradeoffs in assembly simplicity, structural regularity, or scalability relative to the canonical pentamer–hexamer framework.

2.6 From geometric constraints to physical construction

Beyond geometry, capsid assembly is governed by physical forces and self-assembly processes. Assembly depends on a balance among protein–protein interactions, thermodynamic stability, and dynamic structural rearrangements that guide formation of the viral shell (23, 27–28). Viral capsid formation is a classic example of self-assembly, in which complex structures emerge spontaneously from local interactions among simpler components (29). Forming a curved, elastic shell requires bending energy, preventing real capsids from perfectly matching ideal geometric forms (13,15,18). Assembly becomes thermodynamically favorable when the free energy gained from subunit binding compensates for the bending energy required to deform the lattice into a closed shell (13,30). Kinetic factors, including nucleation, growth, and error correction during assembly, also influence the efficiency and fidelity of capsid assembly (23,27,28,30). Although these

physical factors do not alter the geometric requirements for shell closure, they determine which geometrically permissible structures can assemble efficiently and remain stable.

2.7 Quantitative comparison of representative capsid architectures

Because physical forces act upon an underlying geometric framework, geometry provides a useful basis for comparing alternative capsid architectures. To illustrate how geometric organization influences architectural performance, representative architectures were evaluated using common geometric and organizational criteria.

3. Framework design

The framework compares tetrahedral, octahedral, and icosahedral architectures, including the canonical pentamer–hexamer (e.g., HSV-1), all-pentamer (e.g., HPV-16), and prolate (e.g., bacteriophage T4) variants (Figures 1D–F and 4; Table 1). Five evaluation criteria—enclosure efficiency (E), curvature organization (C), vertex coordination (V), polygon-class economy (Pe), and scalability (S)—were scored on an ordinal scale from 0 to 2, with half-point increments permitted (Table 2).

The five criteria were scored using the operational anchors in Table 2. Enclosure efficiency was based on normalized $S/V^{2/3}$, curvature organization on angular deficit and the number of curvature-generating sites, vertex coordination on the number of polygons meeting at each vertex, and polygon-class economy on the number of distinct recurring polygon classes. Scalability was defined as the capacity to increase shell size while preserving the same structural framework. Uniform geometric rescaling alone was not considered

scalability. Preservation of the same structural shell-closure rules. Each architecture received framework required retention of the recurring the highest score whose full definition it polygon classes, local vertex relationships, and satisfied.

Table 2. Operational scoring definitions

Table 2A. Geometry

Score	Enclosure efficiency (E)	Curvature organization (C)	Vertex coordination (V)
0	$r \geq 1.35$	$\delta_{\max} \geq 180^\circ$; $N_c \leq 4$	≥ 5 polygons at each vertex
0.5	$1.25 \leq r < 1.35$	$120^\circ \leq \delta_{\max} < 180^\circ$; $4 < N_c \leq 6$	—
1	$1.15 \leq r < 1.25$	$60^\circ < \delta_{\max} < 120^\circ$; $6 < N_c < 12$	4 polygons at each vertex
1.5	$1.08 \leq r < 1.15$	$\delta_{\max} \leq 60^\circ$; $N_c \geq 12$ nonuniformly distributed curvature	—
2	$r < 1.08$; near-spherical	$\delta_{\max} \leq 60^\circ$; $N_c \geq 12$ uniformly distributed curvature	3 polygons at each vertex

Table 2B. Scalability

Score	Polygon-class economy (Pe)	Scalability (S)
0	≥ 4 distinct polygon classes	No architecture-preserving increase in shell size
0.5	—	—
1	3 distinct polygon classes	Repeatable size expansion through rearrangement or pseudo-hexavalent interactions; closure and expansion remain coupled
1.5	—	—
2	≤ 2 distinct polygon classes	Broad architecture-preserving expansion through addition of curvature-neutral expansion units while preserving a fixed curvature-generating closure set

Each row defines the meaning of a score for the criteria shown. Here, $r = (S/V^{2/3})/(36\pi)^{1/3}$, with $r=1$ for a sphere. $\delta_{\max}$ is the maximum integrated angular deficit at any curvature-generating site; N_c is the number of such sites. When multiple descriptors yielded different scores, the lower score was assigned. Half-point scores denote physically interpretable intermediate states; a dash indicates that no distinct intermediate category was defined.

3.1 Comparative evaluation

The representative architectures exhibit distinct strengths and tradeoffs across the five evaluation criteria (Table 3). Overall architectural utility (U) was calculated as the sum of the five equally weighted criterion scores divided by the maximum possible total score of 10, $U = (E + C + V + Pe + S) / 10$, where E, C, V, Pe, and S denote enclosure efficiency,

curvature organization, vertex coordination,

polygon-class economy, and scalability, respectively.

Tetrahedral and octahedral architectures receive the maximum polygon-class-economy score because each uses only one recurring polygon class. However, their small numbers of curvature-generating sites concentrate positive

curvature at four or six locations, producing greater local distortion and less efficient enclosure than icosahedral architectures (Table 1). The tetrahedral architecture has favorable three-polygon vertex coordination, whereas four polygons meet at each octahedral vertex. Neither architecture permits an increase in shell size while preserving the same structural framework. Uniform geometric rescaling alone does not satisfy the definition of scalability.

Table 3. Criterion scores and overall architectural utility for the evaluated capsid architectures

Architecture	Enclosure	Curvature	Vertex	Economy	Scalability	Utility
Tetrahedral	0	0	2	2	0	0.40
Octahedral	1	1	1	2	0	0.50
All-Pentamer	2	1.5	2	2	1	0.85
Pentamer–Hexamer	2	2	2	2	2	1.00
Prolate Pentamer–Hexamer	1.5	2	2	2	2	0.95

Scores were assigned using the operational definitions in Table 2; utility is the equally weighted normalized sum of the five criterion scores.

All-pentamer architectures achieve efficient icosahedral enclosure and favorable vertex coordination using a single polygon class.

However, they score lower in curvature organization and scalability than canonical pentamer–hexamer systems. As shell size increases, expansion requires increasingly complex quasi-equivalent arrangements among pentamers. Because the same pentameric units support both shell closure and shell expansion, the closure and expansion functions are less clearly separated than in pentamer–hexamer architectures.

Pentamer–hexamer architectures receive the maximum score for all five criteria. Twelve pentamers provide the positive curvature required for shell closure, while hexamers form a curvature-neutral lattice that permits shell enlargement through the addition of hexagons without changing the polygon classes, local vertex relationships, or closure rules. This division of structural roles combines efficient enclosure, organized curvature, favorable vertex coordination, polygon-class economy, and scalability. The prolate variant preserves most advantages of the pentamer–hexamer framework. It retains organized curvature, favorable vertex coordination, polygon-class economy, and scalability. Shell enlargement can occur through the addition of hexameric units while the underlying framework is preserved. However, axial elongation departs from near-spherical geometry and therefore reduces enclosure efficiency.

All five architectures use no more than two recurring polygon classes and consequently receive the maximum polygon-class-economy score. Although this criterion does not distinguish among the representative architectures, it was retained because economical use of a limited motif vocabulary remains relevant when evaluating the broader architectural space.

Under the equal-weight model, the pentamer–hexamer architecture achieved the highest utility score (1.00), followed by prolate pentamer–hexamer (0.95), all-pentamer (0.85), octahedral (0.50), and tetrahedral (0.40). This ranking highlights the architectural advantage of separating shell closure from shell expansion, enabling efficient enclosure and scalable enlargement while preserving organized curvature and a limited motif vocabulary.

3.2 Sensitivity analysis

To assess the robustness of the ranking to weighting and scoring assumptions, three sensitivity analyses were performed (Table 4). First, each criterion was omitted in turn, and utility was recalculated from the four remaining criteria. The pentamer–hexamer architecture remained the highest-scoring architecture in every analysis, although it tied with prolate pentamer–hexamer when enclosure efficiency

was omitted. Prolate pentamer–hexamer otherwise remained second, except when scalability was omitted, in which case it tied with all-pentamer. Removing polygon-class economy, which received the same score across all five architectures, did not change the ranking.

Second, the weight of each criterion was increased individually from 0.20 to 0.40, while the weights of the other four criteria were reduced proportionally to 0.15 each, preserving a total weight of 1.00. Pentamer–hexamer remained highest-scoring in all five tests, followed by prolate pentamer–hexamer and all-pentamer. The relative order of the two lowest-scoring architectures was also preserved except when vertex coordination was emphasized, under which tetrahedral ranked above octahedral. Thus, the principal ranking was robust to substantial variation in criterion weighting.

Table 4. Sensitivity of utility scores to alternative model specifications

Architecture	Baseline utility	Criterion omitted	Alternative weighting	Scalability score ± 0.5 ; weight 0.20	Scalability score ± 0.5 ; weight 0.40
Tetrahedral	0.40	0.25–0.50	0.30–0.55	0.40–0.45	0.30–0.40
Octahedral	0.50	0.38–0.63	0.38–0.63	0.50–0.55	0.38–0.48
All-Pentamer	0.85	0.81–0.94	0.76–0.89	0.80–0.90	0.66–0.86
Pentamer–Hexamer	1.00	1.00	1.00	0.95–1.00	0.90–1.00
Prolate Pentamer–Hexamer	0.95	0.94–1.00	0.90–0.96	0.90–0.95	0.86–0.96

Ranges represent utilities obtained by omitting each criterion individually, increasing each criterion weight from 0.20 to 0.40 while reducing the other four weights to 0.15 each, or varying scalability scores independently by ± 0.5 at scalability weights of 0.20 and 0.40. Utilities were normalized to the maximum possible score under each specification.

Third, because scalability was the only criterion not derived directly from a measured or counted quantity, it was classified at three ordinal levels (0, 1, and 2). To test uncertainty in these assignments, each score was varied independently by ± 0.5 within the 0–2 range. The

intermediate values were used only for sensitivity testing and did not represent separate operational categories. Utility was recalculated using scalability weights of 0.20 and 0.40, with the remaining weights adjusted to sum to 1. At 0.20, canonical pentamer–hexamer remained

highest except for a tie with its prolate variant when the canonical score alone, but not the prolate score, was reduced to 1.5. Prolate ranked modestly higher only when this asymmetric reduction was combined with the 0.40 weight. No other architectural class exceeded either pentamer–hexamer variant under the tested perturbations, supporting the robustness of the broader conclusion favoring pentamer–hexamer organization while not establishing a uniquely optimal geometry.

3.3 Framework scope and limitations

The comparative framework does not explicitly account for molecular energetics, elastic shell mechanics, assembly kinetics, genome–capsid interactions, or evolutionary history. It is also limited to relatively rigid, polyhedral protein capsids and therefore does not encompass pleomorphic enveloped viruses, whose morphology depends strongly on lipid membranes and virus–host interactions. The criteria consequently represent geometric and organizational proxies rather than direct measurements of assembly feasibility or evolutionary fitness. The framework treats the criteria as separable and combines them additively, although their biological effects may interact.

Applying the same criteria and assumptions to all architectures provides a consistent comparison, and the sensitivity analyses demonstrate robustness within the tested scoring and weighting variations. However, they do not establish evolutionary causation or capture every factor influencing capsid formation. This illustrative framework should therefore be viewed as a testable synthesis of the geometric and organizational principles developed in this

study rather than a definitive explanation of viral capsid evolution.

3.4 Beyond icosahedra: enveloped viruses with variable morphologies

Although the pentamer–hexamer architecture provides an effective solution for constructing rigid protein shells, it is not the only successful strategy for viral genome packaging. Other viral lineages employ fundamentally different architectural approaches. Enveloped viruses (e.g., influenza A virus and human immunodeficiency virus), for example, are surrounded by lipid membranes and can exhibit substantial variation in shape and size (31,32). This flexibility may facilitate processes such as membrane fusion, host-cell entry, and immune evasion (33,34). However, membrane-based architectures are generally more susceptible to environmental disruption and provide less rigid structural support than protein capsids. These contrasting strategies illustrate that successful viral architectures arise from tradeoffs among geometric, mechanical, and biological requirements rather than any single design principle.

3.5 Perspectives: Translating capsid geometry into nanocage design rules

The principal implication of this study for nanotechnology is not simply that viral capsids can serve as templates for synthetic nanocages, but that their architecture suggests a potentially transferable design rule: shell closure and shell expansion can be engineered as partially separable functions. In the pentamer–hexamer architecture examined here, a fixed set of curvature-generating pentamers satisfies the topological requirement for closure, whereas curvature-neutral hexamers permit enlargement of the shell without changing the underlying

closure rule. The quantitative comparison further suggests that this separation of functions can be achieved while retaining efficient enclosure, distributed curvature, favorable vertex coordination, and a limited vocabulary of structural motifs. This provides a more specific design hypothesis for synthetic nanotechnology than direct imitation of an icosahedral viral capsid.

One potential application is the development of modular protein nanocage families in which cage size can be varied without redesigning the entire assembly architecture. Current computational approaches have demonstrated increasingly sophisticated control over self-assembling protein architectures, including

highly stable icosahedral cages, multicomponent assemblies, and cages generated through programmed symmetry breaking (Figure 6) (35,38–40). The present analysis suggests an additional design strategy: computationally distinguish subunits or interfaces responsible primarily for generating curvature from those responsible primarily for extending curvature-neutral regions. If these functions can be encoded separately, a common curvature-generating module might be retained while the number or arrangement of expansion modules is altered to generate cages with different internal volumes. Such a system would provide a direct synthetic analogue of the closure–scale decoupling identified here for pentamer–hexamer capsids.

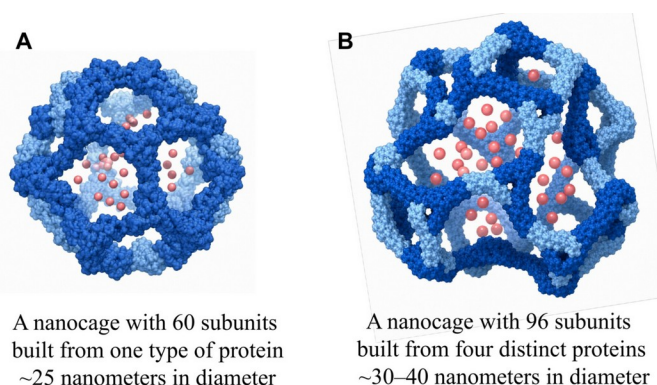


Figure 6. Design strategies for protein nanocage assembly. **(A)** Single-component protein nanocage formed from identical subunits, where assembly is governed by symmetry and repeated interactions (38). The nanocage encapsulates molecular cargo (light pink spheres). **(B)** Multi-component protein nanocage assembled from four distinct protein building blocks through symmetry breaking, enabling larger and more complex architectures (39,40). Both nanocage structures form closed, hollow shells.

The framework also suggests quantitative criteria that could be incorporated into computational screening of candidate nanocage architectures. Rather than evaluating predicted assemblies solely for structural stability, candidate designs could additionally be ranked according to enclosure efficiency, curvature distribution, vertex coordination, motif

economy, and architecture-preserving scalability. Modern structure-prediction and generative-design approaches, including AlphaFold- and RFDiffusion-related methods, make increasingly complex protein architectures computationally accessible (41,42), while topology-aware and reinforcement-learning approaches provide

routes for exploring alternative assembly geometries (43,44). The geometric criteria developed here could therefore function as pre-screening constraints or objective terms within such design workflows. For example, candidate cages could be penalized for excessive localized angular deficit, poor surface-area-to-volume efficiency, proliferation of distinct structural motifs, or inability to alter cage size without changing the fundamental assembly rules.

A particularly testable extension would be to design a series of synthetic cages in which the curvature-generating component is held constant while the number of curvature-neutral expansion units is systematically varied. The resulting structures could then be evaluated for successful closure, internal volume, surface-area-to-volume ratio, assembly yield, structural stability, and cargo capacity. Such experiments would directly test the manuscript's prediction that separating curvature generation from expansion provides a scalable architectural advantage. Importantly, the present geometric model does not establish that architectures receiving higher utility scores will necessarily assemble more efficiently or remain more stable experimentally; molecular energetics, elastic deformation, assembly kinetics, and protein–protein interactions remain outside the current framework. These variables would therefore provide the critical experimental tests linking the geometric predictions developed here to physical nanocage performance.

The analysis also suggests that the useful principle may extend beyond literal pentamers and hexamers. The important feature is the functional partition between units that generate the curvature required for closure and units that

expand the structure while contributing little additional curvature. Consequently, synthetic systems need not reproduce viral proteins or even exact viral polygonal arrangements to exploit the same architectural logic. Recent advances in computationally designed protein assemblies demonstrate that new interfaces, symmetries, and multicomponent cages can be engineered *de novo* (35,38–44). Exploring whether these systems can implement closure–expansion separation using nonviral building blocks would test whether the principle identified here represents a general rule for programmable nanoscale containers rather than a feature restricted to viral evolution.

Thus, a concrete nanotechnology trajectory emerging from this study is the development of geometry-guided, size-programmable nanocages: retain a defined curvature-generating closure module, systematically vary curvature-neutral expansion modules, and computationally optimize the resulting structures against measurable geometric and physical performance criteria (Figure 7). If experimentally validated, this strategy could enable families of related nanocages whose dimensions and cargo volumes are tunable without requiring complete redesign of the underlying assembly architecture. Such modularity could be valuable for encapsulation and delivery applications in which different cargos require different internal volumes while reproducible shell organization is retained (35–40). In this sense, the potentially transferable contribution of viral capsid geometry is not the icosahedral form itself, but the architectural principle of preserving fixed closure requirements while permitting programmable variation in shell expansion.

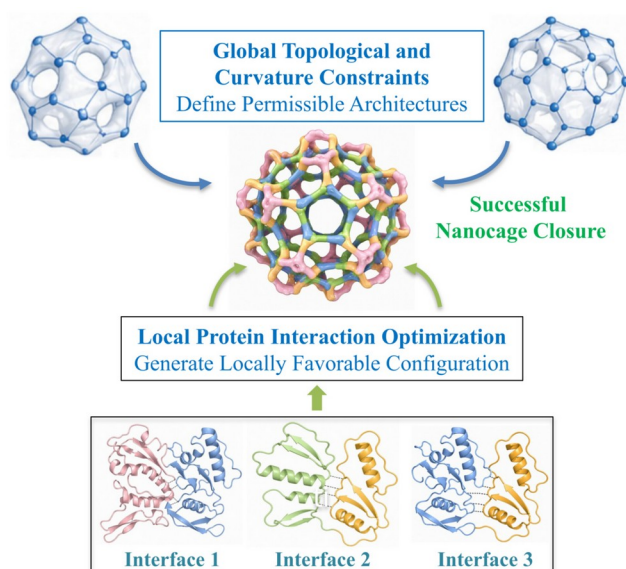


Figure 7. A design framework for nanocage assembly: integrating topological constraints with local interaction optimization. Global constraints define the allowable topology and curvature, limiting the structures that can form. Within these constraints, local protein–protein interactions determine how subunits assemble into a nanocage. Blue nodes indicate discrete curvature-inducing sites (angular deficits). Ribbon diagrams depict protein subunits and their interaction interfaces. Nanocage modified from (39,40).

4. Conclusion

This study examined why pentamer–hexamer organization recurs across otherwise diverse viral capsids by treating capsid architecture as a problem of simultaneously satisfying closure, curvature distribution, enclosure efficiency, structural economy, and scalability. The analysis identifies an important distinction between the roles of the two principal structural motifs: pentamers provide the positive curvature required for closure, whereas hexamers provide curvature-neutral surface expansion. The resulting architecture therefore does more than combine two geometrically compatible polygons; it partitions two potentially competing requirements—closing a finite shell and changing its size—between different structural elements.

This functional separation provides a less self-evident explanation for the versatility of pentamer–hexamer architectures. In architectures where the same structural units determine both closure and expansion, changing shell size can require changes in the arrangement responsible for closure. In the canonical pentamer–hexamer system, by contrast, the twelve-pentamer closure requirement can remain fixed while the hexameric component varies. Thus, architectural complexity need not increase proportionally with shell size. A relatively small motif vocabulary can generate a family of closed structures rather than a single optimized geometry. This suggests that the important property of the architecture is not simply icosahedral symmetry, but closure–scale decoupling.

The quantitative comparison supports this interpretation. Under the five-criterion framework, canonical pentamer–hexamer architecture achieved the highest baseline utility (1.00), followed closely by its prolate variant (0.95), with all-pentamer (0.85), octahedral (0.50), and tetrahedral (0.40) architectures ranking lower. Importantly, sensitivity analysis showed that the exact ordering of the canonical and prolate forms is not absolute: the two can become tied, and sufficiently strong asymmetric perturbation of scalability assumptions can modestly favor the prolate form. However, no other architectural class exceeded both pentamer–hexamer variants across the tested perturbations. The more robust conclusion is therefore not that one precise icosahedral geometry is universally optimal, but that pentamer–hexamer organization occupies a particularly favorable region of the architectural design space under the geometric and organizational criteria examined here.

This distinction also changes how evolutionary convergence should be interpreted. The recurrence of pentamer–hexamer capsids need not imply selection for one geometrically perfect shell. Instead, the architecture may provide a reusable solution in which the fundamental closure rule is conserved while size, elongation, protein fold, and molecular implementation remain free to vary. The existence of prolate pentamer–hexamer capsids is especially informative in this regard: departure from spherical geometry reduces enclosure efficiency but preserves many of the organizational advantages of the underlying framework. Consequently, geometric adaptability rather than geometric perfection may be the more important property of the architecture.

A further implication is that the pentamer–hexamer system may represent a near-minimal architectural vocabulary for combining constraint with variability. Twelve curvature-generating sites satisfy a fixed global topological requirement, while a second, curvature-neutral motif supplies an expandable structural field. This creates a system in which global form can be controlled through a small number of local assembly rules. The same logic suggests a testable principle for synthetic nanotechnology: instead of designing each nanocage size as an independent structure, one could attempt to preserve a curvature-generating closure module while varying curvature-neutral expansion modules to alter cage dimensions and cargo capacity. The potentially transferable feature of viral architecture is therefore not simply the icosahedral shape, but the separation of fixed closure rules from programmable expansion.

These conclusions remain conditional on the scope of the model. The framework evaluates geometric and organizational proxies and does not explicitly model molecular energetics, elastic mechanics, assembly kinetics, genome–capsid interactions, or evolutionary fitness. Accordingly, it does not establish why a particular architecture evolved or whether a high-scoring hypothetical architecture will assemble successfully. Rather, it identifies a specific architectural hypothesis that can now be tested experimentally and computationally: systems that separate curvature generation from curvature-neutral expansion should permit greater variation in shell dimensions while requiring fewer changes to their underlying assembly rules. Testing that prediction across natural capsids and engineered protein cages would determine whether closure–scale

decoupling is merely a useful description of viral geometry or a more general design principle for self-assembling nanoscale structures.

Acknowledgments

The author thanks the reviewers and editors for their thoughtful comments and suggestions, which prompted deeper consideration of the questions addressed, motivated additional literature review, and substantially improved the manuscript. The author also thanks Sophia Chen for carefully reading the manuscript and providing valuable editorial feedback, and Michael Chang for discussions of biological concepts.

5. References

1. Rossmann, M. G., & Johnson, J. E. (1989). Icosahedral RNA virus structure. *Annual Review of Biochemistry*, 58, 533–573. <https://doi.org/10.1146/annurev.bi.58.070189.002533>
2. Flint, S. J., Racaniello, V. R., Rall, G. F., & Skalka, A. M. (2015). *Principles of virology* (4th ed.). ASM Press.
3. Caspar, D. L. D., & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symposia on Quantitative Biology*, 27, 1–24. <https://doi.org/10.1101/SQB.1962.027.001.005>
4. Johnson, J. E., & Speir, J. A. (1997). Quasi-equivalent viruses: A paradigm for protein assemblies. *Journal of Molecular Biology*, 269(5), 665–675. <https://doi.org/10.1006/jmbi.1997.1068>
5. Willey, J. M., Sherwood, L. M., & Woolverton, C. J. (2017). *Prescott's microbiology*. McGraw-Hill.
6. Zhou, Z. H., Dougherty, M., Jakana, J., He, J., Rixon, F. J., & Chiu, W. (2000). Seeing the herpesvirus capsid at 8.5 Å resolution. *Science*, 288(5467), 877–880. <https://doi.org/10.1126/science.288.5467.877>
7. Fokine, A., & Rossmann, M. G. (2014). Molecular architecture of tailed double-stranded DNA phages. *Bacteriophage*, 4(2), e28281. <https://doi.org/10.4161/bact.28281>
8. Guan, J., Bywaters, S. M., Brendle, S. A., Ashley, R. E., Makhov, A. M., Conway, J. F., Christensen, N. D., & Hafenstein, S. (2017). Cryoelectron microscopy maps of human papillomavirus 16 reveal L2 densities and heparin binding site. *Structure*, 25(2), 253–263. <https://doi.org/10.1016/j.str.2016.12.001>

9. Baker, T. S., Olson, N. H., & Fuller, S. D. (1999). Adding the third dimension to virus life cycles: Three-dimensional reconstruction of icosahedral viruses from cryoelectron micrographs. *Microbiology and Molecular Biology Reviews*, 63(4), 862–922.
<https://doi.org/10.1128/MMBR.63.4.862-922.1999>
10. Bamford, D. H., Grimes, J. M., & Stuart, D. I. (2005). What does structure tell us about virus evolution? *Current Opinion in Structural Biology*, 15(6), 655–663.
<https://doi.org/10.1016/j.sbi.2005.10.012>
11. Abrescia, N. G. A., Bamford, D. H., Grimes, J. M., & Stuart, D. I. (2012). Structure unifies the viral universe. *Annual Review of Biochemistry*, 81, 795–822.
<https://doi.org/10.1146/annurev-biochem-060910-095130>
12. Hales, T. C. (2001). The honeycomb conjecture. *Discrete & Computational Geometry*, 25(1), 1–22. <https://doi.org/10.1007/s004540010071>
13. Zandi, R., Reguera, D., Bruinsma, R. F., Gelbart, W. M., & Rudnick, J. (2004). Origin of icosahedral symmetry in viruses. *Proceedings of the National Academy of Sciences of the United States of America*, 101(44), 15556–15560. <https://doi.org/10.1073/pnas.0405844101>
14. Zlotnick, A. (1994). To build a virus capsid: An equilibrium model of the self assembly of polyhedral protein complexes. *Journal of Molecular Biology*, 241(1), 59–67.
<https://doi.org/10.1006/jmbi.1994.1473>
15. Lidmar, J., Mirny, L., & Nelson, D. R. (2003). Virus shapes and buckling transitions in spherical shells. *Physical Review E*, 68, 051910. <https://doi.org/10.1103/PhysRevE.68.051910>
16. Twarock, R. (2004). A tiling approach to virus capsid assembly explaining a structural puzzle in virology. *Journal of Theoretical Biology*, 226(4), 477–482.
<https://doi.org/10.1016/j.jtbi.2003.10.006>
17. Keef, T., & Twarock, R. (2009). Affine extensions of the icosahedral group with applications to the three-dimensional organisation of simple viruses. *Journal of Mathematical Biology*, 59, 287–313. <https://doi.org/10.1007/s00285-008-0228-5>
18. Nelson, D. R. (2002). *Defects and geometry in condensed matter physics*. Cambridge University Press. <https://www.cambridge.org/us/universitypress/subjects/physics/condensed-matter-physics-nanoscience-and-mesoscopic-physics/defects-and-geometry-condensed-matter-physics?format=HB&isbn=9780521801591>
19. Coxeter, H. S. M. (1973). *Regular polytopes* (3rd ed.). Dover.

20. Euler, L. (1758). *Elementa doctrinae solidorum. Novi Commentarii Academiae Scientiarum Petropolitanae*, 4, 109–160.
21. Cromwell, P. R. (1997). *Polyhedra*. Cambridge University Press.
22. Osserman, R. (1978). The isoperimetric inequality. *Bulletin of the American Mathematical Society*, 84, 1182–1238. <https://doi.org/10.1090/S0002-9904-1978-14553-4>
23. Mateu, M. G. (2013). Assembly, stability and dynamics of virus capsids. *Archives of Biochemistry and Biophysics*, 531(1–2), 65–79. <https://doi.org/10.1016/j.abb.2012.10.015>
24. Nasir, A., & Caetano-Anollés, G. (2017). Identification of capsid/coat related protein folds and their utility for virus classification. *Frontiers in Microbiology*, 8, 380. <https://doi.org/10.3389/fmicb.2017.00380>
25. Krupovic, M., & Koonin, E. V. (2017). Multiple origins of viral capsid proteins from cellular ancestors. *Proceedings of the National Academy of Sciences of the United States of America*, 114(12), E2401–E2410. <https://doi.org/10.1073/pnas.1621061114>
26. Sevvana, M., Klose, T., & Rossmann, M. G. (2021). Principles of virus structure. In D. H. Bamford & M. Zuckerman (Eds.), *Encyclopedia of virology* (4th ed., Vol. 1, pp. 257–277). Academic Press. <https://doi.org/10.1016/B978-0-12-814515-9.00033-3>
27. Nguyen, H. D., Reddy, V. S., & Brooks, C. L., III. (2009). Invariant polymorphism in virus capsid assembly. *Journal of the American Chemical Society*, 131(7), 2606–2614. <https://doi.org/10.1021/ja807730x>
28. Hagan, M. F. (2014). Modeling viral capsid assembly. *Advances in Chemical Physics*, 155, 1–68. <https://doi.org/10.1002/9781118755815.ch01>
29. Whitesides, G. M., & Grzybowski, B. (2002). Self-assembly at all scales. *Science*, 295(5564), 2418–2421. <https://doi.org/10.1126/science.1070821>
30. Endres, D., & Zlotnick, A. (2002). Model-based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophysical Journal*, 83(2), 1217–1230. [https://doi.org/10.1016/S0006-3495\(02\)75245-4](https://doi.org/10.1016/S0006-3495(02)75245-4)
31. Harris, A., Cardone, G., Winkler, D. C., Heymann, J. B., Brecher, M., White, J. M., & Steven, A. C. (2006). Influenza virus pleomorphism characterized by cryoelectron tomography.

Proceedings of the National Academy of Sciences of the United States of America, 103(50), 19123–19127. <https://doi.org/10.1073/pnas.0607614103>

32. Briggs, J. A. G., Riches, J. D., Glass, B., Bartonova, V., Zanetti, G., & Kräusslich, H. G. (2009). Structure and assembly of immature HIV. *Proceedings of the National Academy of Sciences of the United States of America*, 106(27), 11090–11095. <https://doi.org/10.1073/pnas.0903535106>

33. Rey, F. A., & Lok, S. M. (2018). Common features of enveloped viruses and implications for immunogen design. *Cell*, 172(6), 1319–1334. <https://doi.org/10.1016/j.cell.2018.02.054>

34. Harrison, S. C. (2008). Viral membrane fusion. *Nature Structural & Molecular Biology*, 15, 690–698. <https://doi.org/10.1038/nsmb.1456>

35. King, N. P., Sheffler, W., Sawaya, M. R., Vollmar, B. S., Sumida, J. P., André, I., Gonen, T., Yeates, T. O., & Baker, D. (2012). Computational design of self-assembling protein nanomaterials with atomic-level accuracy. *Science*, 336(6085), 1171–1174. <https://doi.org/10.1126/science.1219364>

36. Edwardson, T. G. W., & Hilvert, D. (2019). Virus-inspired function in engineered protein cages. *Journal of the American Chemical Society*, 141(24), 9432–9443. <https://doi.org/10.1021/jacs.9b03705>

37. Butterfield, G. L., Lajoie, M. J., Gustafson, H. H., Sellers, D. L., Nattermann, U., Ellis, D., Bale, J. B., Ke, S., Lenz, G. H., Yehdego, A., Ravichandran, R., Pun, S. H., King, N. P., & Baker, D. (2017). Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature*, 552, 415–420. <https://doi.org/10.1038/nature25157>

38. Hsia, Y., Bale, J. B., Gonen, S., Shi, D., Sheffler, W., Fong, K. K., Nattermann, U., Xu, C., Huang, P.-S., Ravichandran, R., Yi, S., Davis, T. N., Gonen, T., King, N. P., & Baker, D. (2016). Design of a hyperstable 60-subunit protein icosahedron. *Nature*, 535, 136–139. <https://doi.org/10.1038/nature18010>

39. Lee, S., Kibler, R. D., Ahn, G., Hsia, Y., Borst, A. J., Philomin, A., Kennedy, M. A., Huang, B., Stoddard, B., & Baker, D. (2025). Four-component protein nanocages designed by programmed symmetry breaking. *Nature*, 638, 546–552. <https://doi.org/10.1038/s41586-024-07814-1>

40. Dowling, Q. M., Park, Y.-J., Fries, C. N., Gerstenmaier, N. C., Ols, S., Yang, E. C., Wargacki, A. J., Dosey, A., Hsia, Y., Ravichandran, R., Walkey, C. D., Burrell, A. L., Veessler,

D., Baker, D., & King, N. P. (2025). Hierarchical design of pseudosymmetric protein nanocages. *Nature*, 638, 553–561. <https://doi.org/10.1038/s41586-024-08360-6>

41. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>

42. Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borst, A. J., Ragotte, R. J., Milles, L. F., Wicky, B. I. M., Hanikel, N., Pellock, S. J., Courbet, A., Sheffler, W., Wang, J., Venkatesh, P., Sappington, I., Vázquez Torres, S., ... Baker, D. (2023). De novo design of protein structure and function with RFdiffusion. *Nature*, 620, 1089–1100. <https://doi.org/10.1038/s41586-023-06415-8>

43. Lutz, I. D., Wang, S., Norn, C., Courbet, A., Borst, A. J., Zhao, Y. T., Dosey, A., Cao, L., Xu, J., Leaf, E. M., Treichel, C., Litvicov, P., Li, Z., Goodson, A. D., Rivera-Sánchez, P., Bratovianu, A.-M., Baek, M., King, N. P., Ruohola-Baker, H., & Baker, D. (2023). Top-down design of protein architectures with reinforcement learning. *Science*, 380(6642), 266–273. <https://doi.org/10.1126/science.adf6591>

44. Albanese, K. I., Barbe, S., Tagami, S., Woolfson, D. N., & Schiex, T. (2025). Computational protein design. *Nature Reviews Methods Primers*, 5, 13. <https://doi.org/10.1038/s43586-025-00383-1>