Limits of economy and fidelity for programmable assembly of size-controlled triply-periodic polyhedra

Carlos M. Duque[a,b,c], **Douglas M. Hall**[d], **Botond Tyukodi**[e], **Michael F. Hagan**[f], **Christian D. Santangelo**[c,g], **and Gregory M. Grason**[d]

[a]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; [b]Center for Systems Biology Dresden (CSBD), 01307 Dresden, Germany; [c]Department of Physics, University of Massachusetts, Amherst, MA 01003; [d]Department of Polymer Science and Engineering, University of Massachusetts, Amherst, MA 01003; [e]Department of Physics, Babes-Bolyai University, 400084 Cluj-Napoca, Romania; [f]Martin A. Fisher School of Physics, Brandeis University, Waltham MA 02453; [g]Department of Physics, Syracuse University, Syracuse, NY 13210

This manuscript was compiled on April 22, 2024

We propose and investigate an extension of the Caspar-Klug symmetry principles for viral capsid assembly to the programmable assembly of size-controlled triply-periodic polyhedra, discrete variants of the Primitive, Diamond, and Gyroid cubic minimal surfaces. Inspired by a recent class of programmable DNA origami colloids, we demonstrate that the economy of design in these crystalline assemblies – in terms of the growth of the number of distinct particle species required with the increased size-scale (e.g. periodicity) – is comparable to viral shells. We further test the role of geometric specificity in these assemblies via dynamical assembly simulations, which show that conditions for simultaneously efficient and high-fidelity assembly require an intermediate degree of flexibility of local angles and lengths in programmed assembly. Off-target misassembly occurs via incorporation of a variant of disclination defects, generalized to the case of hyperbolic crystals. The possibility of these topological defects is a direct consequence of the very same symmetry principles that underlie the economical design, exposing a basic tradeoff between design economy and fidelity of programmable, size controlled assembly.

self-assembly | programmable materials | addressable assembly | triply-periodic polyhedra | self-closing assembly

In Nature, self-assembly underlies the creation of functional materials. From photonic nanostructures (1, 2) to extracellular media (3, 4) to nanoencapsulation (5–7), robust and dynamic control over the precise structure and size of these assemblies is essential to their adaptive properties. Living systems have evolved pathways to direct multi-protein assembly towards morphologies with a tunable size scale (8), from finite diameter shells and tubules (9) to periodically-modulated material composites (10). Indeed, length scale is a fundamental but crucial element of structure control for functional materials. Yet, achieving structures with well-controlled lengths, particularly when the target size is much larger than the constituent size, poses a basic challenge. Notably, the ability to target triply-periodic (i.e. crystalline) architectures with specific symmetries and periodicities that can be scaled to larger than the subunit sizes is fundamental to control of desirable material functions, such as photonic bandgap behavior. In particular, hybrid structures related to the Diamond and Gyroid surfaces are well-known to exhibit prominent bandgaps, and gyroid-like structures have evolved in diverse species of birds, beetles and butterflies as a means of producing structural coloration (11–14). In these assemblies, targeted wavelength-selective properties are achieved by control over periodicity at the scale of 100s of nm, at least an order of magnitude larger than the protein building blocks themselves.

These examples from nature have inspired recent efforts to design "programmable" building blocks to realize synthetic analogs of hierarchically-organized biological materials (15–19). This strategy targets two key aspects of the protein building blocks of functional biological assemblies. First, new approaches to designing building block *geometry* from the molecular to the colloidal scale have enabled the fabrication of shapes of staggering complexity (20–23). Second, encoding multiple species of subunits with specific interactions to favor a particular network of contacts allows the assignment of a specific "address" to every assembled subunit (15, 17, 24–27). Yet, while such *addressable assembly* offers one simple and generic approach to this problem, where the number of interacting elements must grow with the target size, this paradigm suffers from a corresponding explosion in the complexity of multi-species mixtures as the target grows arbitrarily large(28–30). This is notably the case for programmable *crystalline* assemblies, where the unit cell dimensions typically scale with the size of the programmable building blocks themselves (16, 31–37), so far limiting applications, for example, to the use of nanometric DNA based assembly units to target photonic

Significance Statement

As first suggested by Caspar and Klug, many viruses assemble icosahedral shells (capsids) because the high symmetry of the icosahedron enables economical assembly – enclosing a large volume with relatively few distinct protein subunit types. We generalize this design principle to triply-periodic polyhedra, mesoporous structures approximating cubic minimal surfaces. We demonstrate their programmable assembly from a minimal number of distinct subunits forming arbitrarily large unit cells of tunable, defined size. However, while high symmetry points enable economy in these target structures, they can be seeds of mis-assembly. This design strategy, and the fundamental tradeoff between economy and fidelity, lays the groundwork for deploying rapidly advancing nanotechnology approaches to programmable assembly to achieve size-controlled architectures with tunable functional properties.

All authors designed and performed research, analyzed data and wrote the paper.

The authors declare no competing interests.

[2]To whom correspondence should be addressed. E-mail: grason@umass.edu

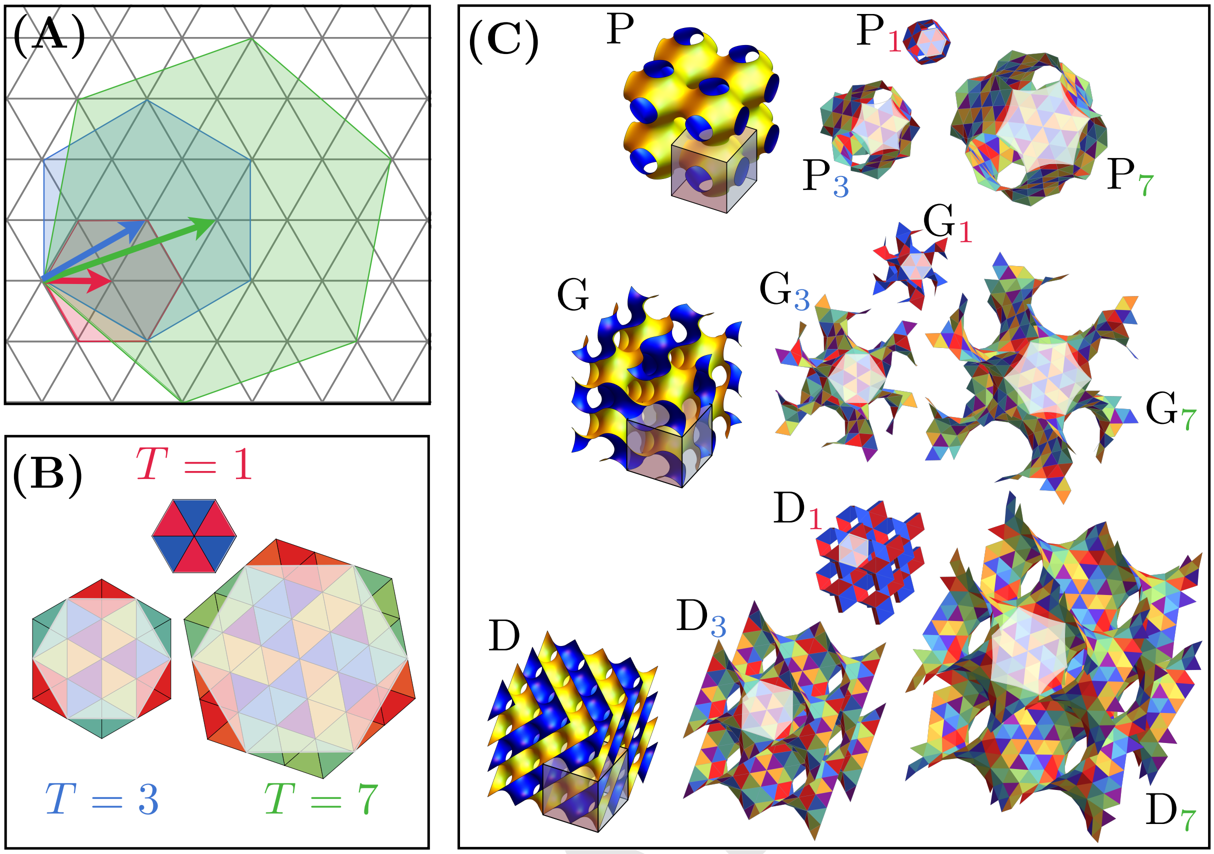


Fig. 1. Design of size-controlled triply-periodic minimal surface assembly from programmable triangular particles. **(A)** A pair of integers (h, k) is used to define a triangulation vector, $\mathbf{L}$, which connects a pair of vertices on a triangular lattice, defining the fundamental hexagonal patch. The hexagons in red, blue, and green are built with $\mathbf{L}$ vectors corresponding to the triangulation numbers $T = 1$, 3 and 7. **(B)** Underlying triangular faces associated with each fundamental hexagon. Triangles with the same colors represent the same subunit type. **(C)** From top to bottom: $2 \times 2 \times 2$ subsets of primitive (P), gyroid (G), and diamond (D) TPMS. In each case, transparent boxes are used to highlight a single cubic unit-cell. Each cubic unit-cell is triangulated with the corresponding fundamental hexagon shown in panel **(B)**. The shaded areas on each triangulated unit cell highlight the fundamental hexagon used in the construction.

structures with properties in the optical range (38).

DNA nanotechnology, in particular, has multiple strategies to implement addressable assembly, and has been exploited to target and realize assemblies with a precisely defined, complex structures (29, 30, 39). Recent works leverage the unprecedented combination of control over geometry and interaction specificity possible through DNA origami to realize a class of quasi-spherical shells and cylindrical tubes (40–44). Crucially, their target diameters are regulated by programming their curvature, achieving finite sizes that are much larger than the subunits. The design strategy for size-controlled shells (42) takes advantage of symmetry-based principles proposed for icosohedral viral shells (45–48), the celebrated *Caspar-Klug* (CK) construction (49). The CK rules provide a rational means to determine the minimal number of inequivalent subunits (*i.e.* conformations of capsid proteins) needed to form closed shells of arbitrarily large diameter, an economy of design presumably favored by selection pressures in viral evolution. In this context, CK rules might be considered as one class of solutions to the generic problem of miminizing the *complexity* of specifically interacting subunit mixtures needed to achieve size-programmed assembly (50, 51).

In this article we propose and explore the extension of the symmetry-based principles of CK to an entirely distinct class of programmable assemblies, triply-periodic polyhedra (52, 53), shown schematically in Fig. 1. Like tubules and shells, these are 2D surface-like assemblies of triangular subunits, in our case inspired by DNA origami assemblies (42–44). In contrast to shell-like assemblies, our proposed design rules employ negative, rather than positive, Gaussian curvature. In particular, we target the design of a related class of triply-periodic minimal surfaces (TPMS), the Gyroid (G), Diamond (D) and Primitive (P) surfaces (54, 55), which all have cubic symmetry and negative Gaussian curvature. In a conceptually related work, Tanaka and coworkers demonstrated the ability to form "bicontinuous" structures related to P and G surfaces via a model of polygonal nanoplate assembly (53). However, that design scheme requires distinct polygonal shapes, distinct symmetries (i.e. P vs. G), and moreover, unit cell dimensions are controllable only by changing the size of the plate-like particle.

We show here that the high symmetries of the P, G and D structures facilitate a similar economy of design as CK capsids, arising from the commensurability of their crystallographic space-groups with their decomposition into triangular building blocks. Like CK designs of closed shells, we show how the combination of *interaction specificity*, encoded in specifically binding edge types, and *geometric specificity*, encoded by the target edge lengths of and dihedral angles between subunits, allows the programming of unit cell sizes that are, essentially, arbitrarily large compared to the subunit size. We analyze the economy of design in triply-periodic polyhedral assemblies in terms of the scaling of the number of distinct subunits needed to achieve a given target periodicity in the crystal, and show that these "inverted" structures achieve similarly optimal scaling with increasing target size in comparison to icosohedral shells.

While the economy of design deriving from symmetry guarantees a unique target ground state, it does not guarantee that it correctly assembles, and if so, at reasonable rates or under physical conditions realizable in experiments. Notably, recent experiments on self-closed tubular assemblies of DNA particles show that subunit flexibility, specifically dihedral bending, gives rise to significant off-target assembly into structures of undesired diameter (44). To understand how geometric specificity, in the form of angular and length flexibility of subunits, limits the ability to achieve controllable crystal dimensions through economically programmed assembly, we study grand canonical Monte Carlo simulations of a physical model of triangular subunit assembly. We consider the effects of variable elastic stiffness, and show that the range of subunit flexibility is restricted by the simultaneous requirements for high-fidelity and rapid assembly, which are respectively favored by low or high flexibility. Notably, the failure mechanism leading to misassembly for sufficiently flexible subunits, in the form of generalized disclination in hyperbolic crystalline structures, can be traced directly to the very same symmetry-based design that guarantees economical assembly.

Design economy of triply-periodic polyhedra

The economy of the CK construction for quasi-spherical assemblies stems from the fact that spherical shells can be decomposed into regular triangular "subunits", corresponding to geometric structures known as *deltahedra*. In the original CK construction, subunits were themselves triplets of proteins that form the viral capsid (49). Here, we follow the design of Sigl *et al* (42) and consider the triangular subunit as a single, self-assembling unit, shown schematically in Fig. 2A. The most economical shell designs are closed tilings of equilateral triangles, *i.e.* the Platonic solids: tetrahedron (T), octahedron (O), and icosahedron (I). Among these, the icosahedron possesses the most point symmetries, and correspondingly the largest number of equivalent triangular facets (20). Structures composed of more than this number of triangles necessarily break the 3-fold symmetry of the equilateral subunit, and so increase the number of *inequivalent* triangular elements needed to form them.

CK argued that subtriangulations of the original triangular net that preserve icosahedral symmetry lead to the fewest symmetry-inequivalent positions on the closed shell, and thus require the fewest distinct subunits. Such subtriangulations are constructed from triangular subregions of a planar triangular lattice (46, 47), and are parameterized by the lattice translation vector between vertices $\mathbf{L} = h\mathbf{a}_1 + k\mathbf{a}_2$, where (h, k) are a pair of integers and $\mathbf{a}_1$, $\mathbf{a}_2$ are basis vectors of the triangular lattice. Then $T \equiv |\mathbf{L}|^2 = h^2 + k^2 + h\ k$ is the number of subtriangles per base triangle, resulting in a structure with $20T$ subunits. However, the commensurability of subtriangulation and icoshedral symmetry implies that complete shells can be assembled from fewer distinct triangle types, N_T, which is equal to $\lceil T/3 \rceil$ for deltahedral shells (i.e. with equilateral base faces) *. A design objective to maximize the target size, D, of an assembly for a minimal number of distinct subunit types, N_T, suggests a measure of *economy*,

$$e(D) \equiv D/N_T. \quad [1]$$

As shown in Fig. 2C, this measure decreases with target size as $e(D) \sim 1/D$ for deltahedral assemblies – assemblies made from equilateral triangles – where we use the maximal medial thickness (56) as the standard measure of size D †. This scaling can be understood from CK theory, since the triangulation

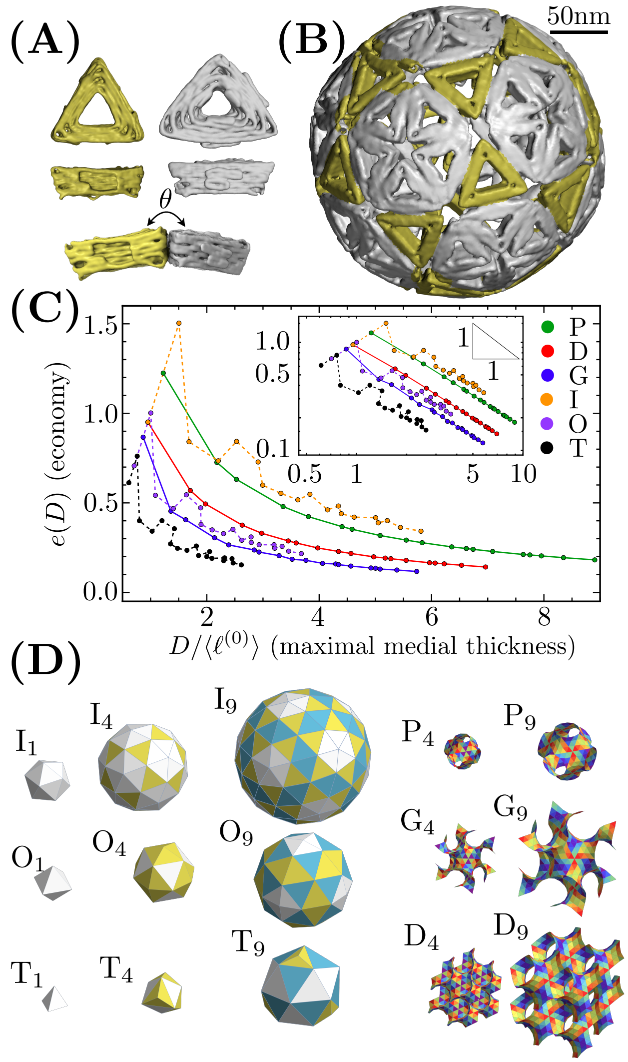


Fig. 2. Economy of programmable assembly of shells and TMPS via triangular particles. **(A)** Cryo-eM reconstruction of the two different triangular DNA origami subunit types used in the self-assembly of $T = 4$ icosahedral shell shown in **(B)** adapted from Sigl *et al.* (42). The angle θ highlights the preferred dihedral angle between the subunits. **(C)** Measure of *economy*, $e(D)$ for polyhedral crystals (solid lines) and spherical shells (dashed-lines) as a function of the maximal medial thickness, D (in units of mean edge length). Inset: Log-log plot of the same economy measures. **(D)** Triangulated shells of icosahedal, octahedral, and tetrahedral symmetries with triangulation number T = 1, 4, and 9 built with $N_T = \lceil T/3 \rceil$ distinct subunits (left). $T = 4$ and 9 polyhedral crystals built with $N_T = T$ distinct subunits (right).

*There are T inequivalent internal edges in a deltahedral tiling which are distributed into groupings of three (i.e. closed triangles). The minimal number of distinct triplets is $\lceil T/3 \rceil$.

†The *medial thickness* of a bounding surface at a particular point is the radius of largest sphere enclosed by the surface tangent to that point (56). Hence, the *maximal medial thickness*,

number is proportional to the surface area and thus $T \propto D^2$. Notably, icosahedral shells maximize this measure of economy among deltahedral shells, which is also consistent with the CK logic. The non-monotonic behavior for shell economies stems from the fact that $N_T = \lceil T/3 \rceil$ for T, O, and I. This leads to "magic numbers" of especially high $e(D)$ values when T is an integer multiple of 3 (e.g. $T = 3, 9, 12, 21 \ldots$), corresponding to shell geometries in which the 3-fold axis lies at a vertex of the triangular net (i.e. as opposed to the general case where the 3-fold axis falls at the center of a triangular subunit). Since the $\bar{3}$ symmetry at the center of the hexagonal base of $\mathrm{P_T}$, $\mathrm{D_T}$, and $\mathrm{G_T}$ is not compatible with the center of a triangular subunit (and, instead, only at 6-coordinated vertices), such magic numbers do not exist for triply-periodic polyhedra.

This design principle can be extended to triply-periodic triangular assemblies associated with the P, D, and G cubic minimal surfaces, by decomposing these structures into basic hexagonal elements, essentially following what has been dubbed "hexagulation" in studies of dense particle packings on these surfaces (57). As developed by Sadoc and Charvolin (58) and elaborated by others (59–61), high-symmetry tilings of P, G, and D can be derived from the $\{6, 4\}$ tiling of the hyperbolic plane projected onto triply-periodic tesselations of $\mathbb{E}^3$ (Euclidean three space), provided that point symmetries at the center, vertices and edges of the hexagonal patches are preserved in the space group embedding. However, while CK triangulations are constructed from triangular tilings meeting at 5-fold vertices, $\{6, 4\}$ tilings are constructed from hexagonal tiles that meet in 4-fold vertices (58), which is possible on the hyperbolic plane, as illustrated in Supporting Fig. S1 [‡]. When projected into cubic tesselations of $\mathbb{E}^3$, vertices of a projected skew hexagonal base "tile" are constrained to lie on specific Wyckoff positions of the corresponding crystallographic space group of the structure. See for example, Fig. 3A for G, where the central point of the shaded hexagon sits at the 16a position of $Ia\bar{3}d$, a point of 3-fold roto-inversion symmetry, while the six outer vertices sit at 24d, points of 4-fold roto-inversion (62). Like the triangular base elements of the CK construction, the hexagonal base tile of the triply-periodic P, G, and D polyhedra can be subtriangulated in a way that preserves the point symmetries of the $\{6, 4\}$ tiling (see Methods and examples shown in Supporting Fig. S2). The corresponding triply-periodic triangulations are similarly triangulated by a vector $\mathbf{L}$ that connects a 4-coordinated vertex to center of the hexagon, and thus the tilings are classified by the T number (Fig. 1A).

We use these triply-periodic triangulations of $\mathbb{E}^3$ as the basis for triangular subunit designs that target the assembly of TPMS whose vertex positions map onto the corresponding minimal surface for arbitrary subtriangulation. We denote these polyhedral target structures as P_T, D_T and G_T according to the respective cubic surface and T number. The geometric and topological data of the embedded structure – edge length, connectivity and dihedral angles – are then used to design target values for the subunit shape and interactions. In our construction, the mean edge length of a basic triangular element is fixed for all structures, while the projection subtriangulation can introduce variation in edge lengths. Fig. 3 shows an example for the G_3 structure in terms of 9 mutually interacting subunit edges. The interaction matrix in Fig. 3 includes both edge specificity and geometric data about the edge lengths and the dihedral angles formed when particular subunits meet. Notably, the coincident symmetries of the subtriangulation and the $Ia\bar{3}d$ space group allow for the large units cells ($96T$ triangular particles per cubic repeat of G) to be constructed from only $N_T = T$ distinct subunits. Supporting Figs. S3-6 show corresponding examples of binding specificity and geometry of edges for P, G and D for $T = 1, 3$ and 7

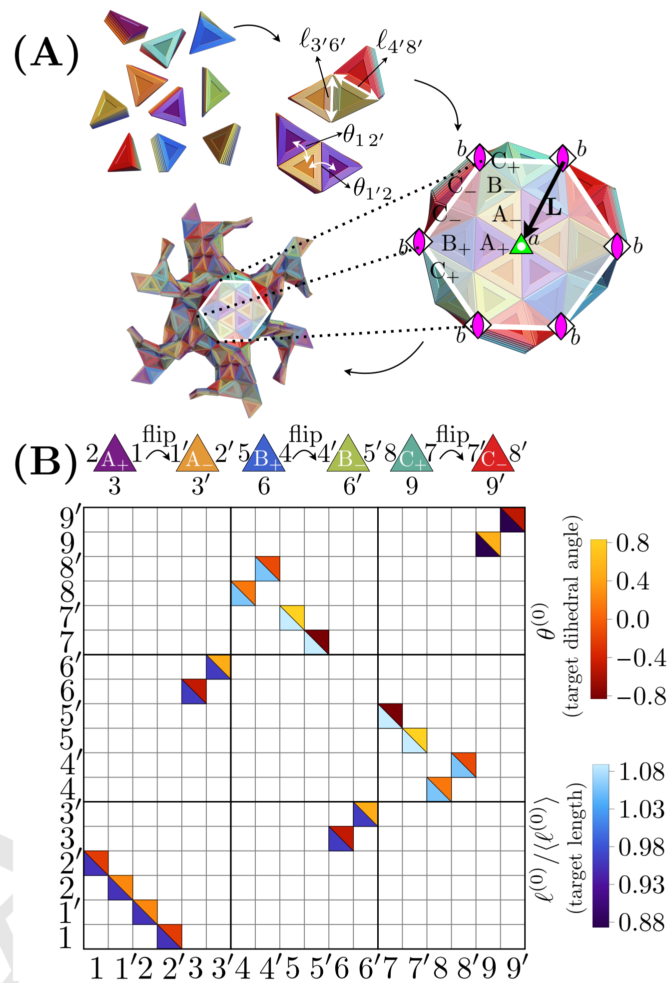


Fig. 3. Program for $T = 3$ Gyroid assembly. **(A)** Illustration of the self-assembly process: triangular blocks of different types following their interaction matrix template form larger structures, such as the shown $\mathrm{G_3}$ cubic unit-cell. The three-dimensional rendering highlights the structure as DNA origami particle, with the important feature that opposite faces of the particle are distinct, as colored and referenced a $\pm$ in **(B)**. The center and vertices of the highlighted hexagon represent the Wyckoff sites present in the triangulation. The glyph next to each vertex denotes the symmetry of the Wyckoff site. Notice that the translation vector $\mathbf{L}$, which defines the T number of the structure, joins Wyckoff sites of different symmetry. **(B)** Interaction matrix for $\mathrm{G_3}$ with each colored block representing a valid edge-pairing. The monovalent design rules used in the construction ensure that a given edge-type is only allowed to bind to at most one other edge-type. The lower (upper) color of each colored-block of the matrix represents the target length (dihedral angle) of a valid edge-pairing. Both planar sides of each triangular block are colored differently to account for the "flipping" symmetry exhibited by TPMS. This is accounted by assigning two edge-types to each triangular block planar side. Unprimed and primed edge-types are respectively the edge sides of the $+$ and $-$ planar sides of each triangular block.

used as a generic measure of self-closing size D, corresponds to the radius of the largest sphere that can be enclosed by the surface.

‡Note that, strictly, the triangulation in the hyperplane satisfies rotoinversion symmetries on the vertices ($\bar{4}$) and centers ($\bar{3}$) of the fundamental hexagonal cells as annotated Supporting Fig. S1A.

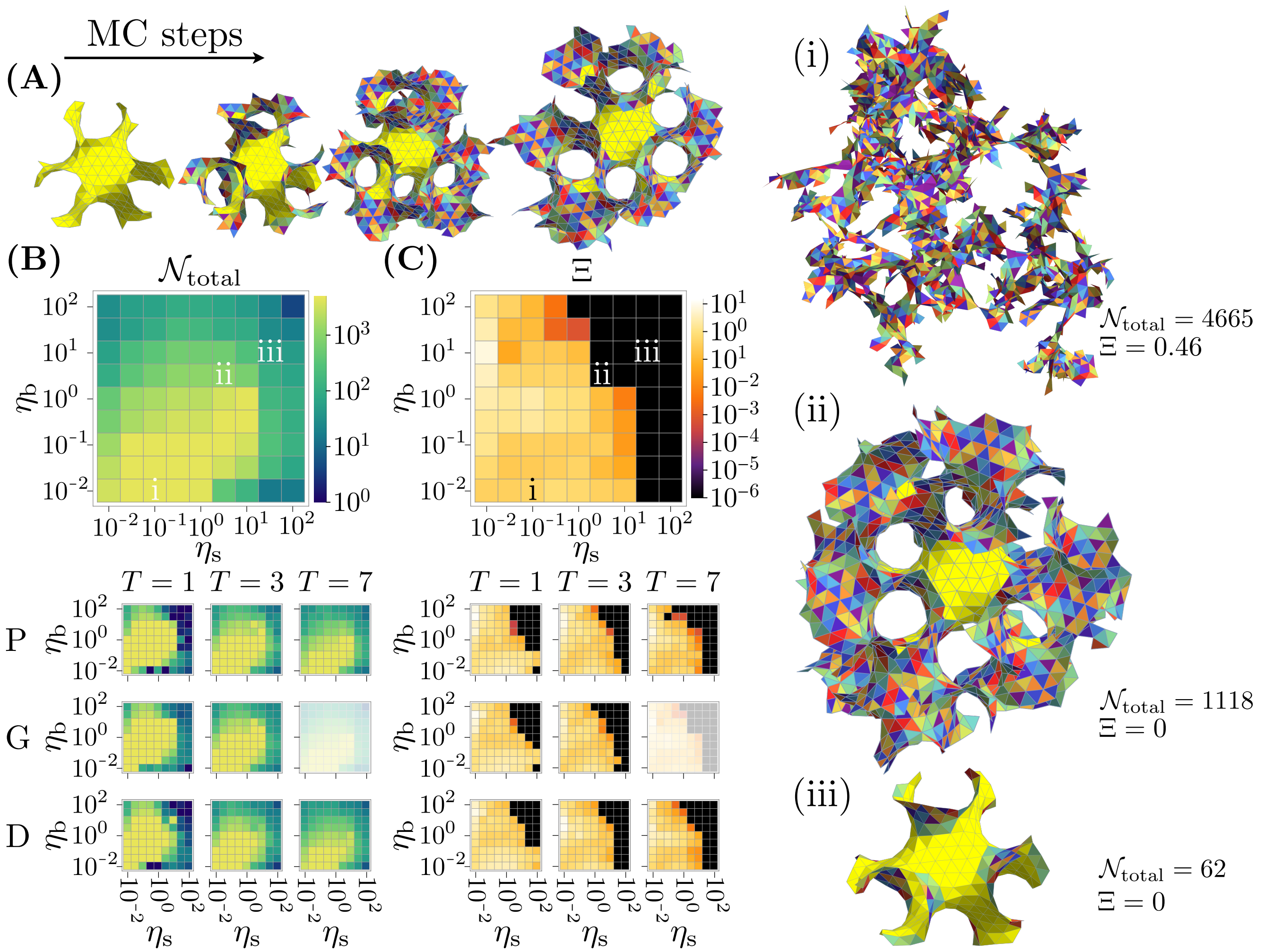


Fig. 4. Efficiency and fidelity of assembly versus geometric specificity of binding. (A) An example assembly simulated trajectory of the G_7 assembly, growing from a fixed "seed" (yellow), with newly bond subunits shown in corresponding colors. **(B)** and **(C)**, respectively, show the number of assembled subunits ($\mathcal{N}_{\text{total}}$) and residual shape strain (Ξ) from simulation trajectories for a range of dimensionless bending (η_{b}) and stretching (η_{s}) stiffness, the upper panel highlights the results for G_7, while results for all three symmetries and $T = 1, 3$ and 7 are shown below. Notice that the large panels in **(B)** and **(C)** correspond to the smaller whitened panels shown below for the G_7 structure. Points (i), (ii) and (iii) highlight three different conditions for G_7 assembly corresponding the final structures shown on the right: (i) rapid, off-target assembly; (ii) productive, on-target assembly; and (iii) non-productive, rigid assembly (also shown in Supporting Movies S4-6).

Unlike the CK structures, triply-periodic P_T, D_T and G_T polyhedra have infinite genus in their target (bulk) state. Nevertheless, each structure can be characterized by a well-defined, finite size, roughly corresponding to the characteristic pore size of the structure. To quantify and compare this programmable size scale, we computed the maximal medial thickness D of each triply-periodic polyhedra (56). Fig. 2C plots the finite size design economy, $e(D)$, as function of T for P_T, D_T and G_T polyhedra in comparison to generalized CK designs. Like the CK assemblies, scaling the target size of the triply-periodic polyhedra to larger and larger dimensions (in units of the basic triangular elements) can be achieved at a similar level of economy, with $D \sim T^{1/2}$ and $N_T \sim T$, so that the power law scaling of e with size is identical in the large size limit.

Efficiency versus fidelity of size-economic assembly

The economy of design of our construction for triply-periodic polyhedra depends on the combination of both the interaction specificity between different edge types and the geometric specificity of the edge-edge contacts. To understand the physical limits for the fidelity and assembly yield of these structures, we implemented a coarse-grained model of triangular particle assembly employed in Tyukodi *et al.* (63), in which assembled structures are triangular meshes, with degrees of freedom at their vertices. Similar elastic mesh simulations have been applied to model assemblies of shells and tubules (63–68). Here, the energy of an assembly derives from interactions of bound edges indexed by i and j,

$$E = \sum_{\langle ij \rangle} \Big\{ -\epsilon_{ij} + \frac{k}{2} |\ell_{ij} - \ell_{ij}^{(0)}|^2 + \kappa \big[1 - \cos(\theta_{ij} - \theta_{ij}^{(0)}) \big] \Big\}. \quad [2]$$

The first term describes the binding energy of the edges. We assume that edge i and edge j have a common binding affinity,

$\epsilon_{ij} = \epsilon_{\text{bind}}$ if they are programmed to interact, but do not bind otherwise. Notably, this interaction specificity, which is crucial for forming a target triply-periodic polyhedra, is an important contrast to previous studies of capsid assembly in which all subunit interactions were identical (64–66, 69, 70), but is analogous to models of capsid assembly in which subunit interactions followed CK rules (68, 71–84). The second and third terms in eq. (2) describe the energy cost of edge stretching and bending, respectively, where $\ell_{ij}^{(0)}$ and $\theta_{ij}^{(0)}$ are the target values for edge ij, as determined from the geometrical embedding of P_T, D_T and G_T (e.g. as shown for G_3 in Fig. 3). The respective moduli for edge stretching and dihedral bending between bound faces are given by k and κ. To consider the influence of these distinct elastic modes on assembly behavior, we introduce the dimensionless ratios relative to binding,

$$\eta_{\text{s}} \equiv k\langle|\ell_{ij}^{(0)}|^2\rangle/\epsilon_{\text{bind}}; \quad \eta_{\text{b}} \equiv \kappa/\epsilon_{\text{bind}} \qquad [3]$$

where $\langle|\ell_{ij}^{(0)}|^2\rangle$ is the mean-square target edge length.

To model near-equilibrium assembly behavior, we perform grand canonical Monte Carlo simulations (see Methods), which consider a single cluster of bound, triangular units held at fixed chemical potential with respect to a population of free subunits (i.e. monomers) composed of a mixture of all the triangular species needed to assemble a given triply-periodic polyhedron. To test the efficiency and quality of targeted assembly, an initial seed of the preassembled structure is prepared, and the MC algorithm considers three types of moves: 1) addition/removal of free subunits to an appropriately unbound triangle edge of the cluster; 2) vertex displacement of assembled particles; and 3) fission/fusion of edges between bound/unbound edges of the assembled cluster. Moves are accepted with the Boltzmann-weighted probability according to the change in energy, Eq. (2), and chemical potential μ for free monomer addition at fixed temperature. Example simulation trajectories are shown for $T = 3$ structures in Supporting Movies S1-3.

To consider the role of geometric specificity, we choose $\epsilon_{\text{bind}} = -6.5k_{\text{B}}T$ and chemical potential $\mu = -4.5k_{\text{B}}T$ and vary the elastic/binding ratios. We introduce $\mathcal{N}_{\text{total}}$, the average number of assembled units in the particle cluster, as a measure of assembly efficiency. To capture the fidelity of the assembly we quantify the mean-quadratic strain

$$\Xi = \langle|\ell_{ij} - \ell_{ij}^{(0)}|^2\rangle/\langle|\ell_{ij}^{(0)}|^2\rangle + \langle|\theta_{ij} - \theta_{ij}^{(0)}|^2\rangle, \qquad [4]$$

which is computed from the elastic ground-state of the ultimate structure to remove the influence of thermal fluctuations. We terminated simulations at 50×10^6 MC sweeps, or when $\mathcal{N}_{\text{total}} = 5000$. We simulated assembly trajectories for P, G, and D structures for a range of T-numbers and varying the elastic/binding ratios over four orders of magnitude: $\eta_{\text{s}} \in [10^{-2}, 10^2]$ and $\eta_{\text{b}} \in [10^{-2}, 10^2]$.

Fig. 4B and C show simulation results for the dependence of $\mathcal{N}_{\text{total}}$ and Ξ on η_{b} and η_{s} for G_7 assemblies. They show that $\mathcal{N}_{\text{total}}$ decreases with both stretching and bending stiffness, with the efficiency near zero outside of the region $\eta_{\text{s}} \lesssim 10$ and $\eta_{\text{b}} \lesssim 1$. In contrast, the high-flexibility regime is the regime of *low fidelity* as indicated by the large, non-zero residual strain values Ξ. These $\Xi > 0$ values are evidence of assemblies that form with the correct edge matching specificity, but nevertheless have a topology that is incompatible with the target edge lengths and dihedral angles. In other words, while assembly is rapid when subunits are flexible, $\eta_{\text{s}} \ll 1$ or $\eta_{\text{b}} \ll 1$ result in highly defective, "off-target" structures, as shown for a structure like Fig. 4i. In the opposite limit of $\eta_{\text{s}} \gg 1$ or $\eta_{\text{b}} \gg 1$, while bonds form the correct geometry so that $\Xi \simeq 0$, the assembly efficiency is low due to the small rate of new subunits joining to a free edge. This results in a structure like Fig. 4iii, with few if any additional subunits bound to the seed. However, at intermediate flexibility – approximately $5 \lesssim \eta_{\text{s}} \lesssim 10$ and $5 \lesssim \eta_{\text{b}} \lesssim 10$ – assembly achieves both significant yield (i.e. $\mathcal{N}_{\text{total}} \gtrsim 10^2$) and high fidelity (i.e. $\Xi \simeq 0$), indicating productive and defect-free assembly of the target crystalline structure (Fig. 4ii).

The smaller panels of Fig. 4B and C compare the assembly efficiency, $\mathcal{N}_{\text{total}}$, and fidelity, Ξ, for all three P_T, G_T, and D_T structures and for sequences of increasing target sizes, corresponding to $T = 1, 3$, and 7. All cases show the same qualitative dependence on angular and length flexibility of bound subunits: rapid yet off-target assembly at high flexibility, on-target yet sluggish assembly for stiff structures, and a regime of productive, on-target assembly in the intermediate flexibility regime. These overall trends reveal that interaction specificity alone is not sufficient for reaching target assemblies. This is indeed consistent with the fact that P_T, G_T and D_T have an identical interaction matrix for each given T, but differ in terms of the target edge lengths and dihedral angles.

It is instructive to compare these observations to results of models for positive curvature capsids. While we find that interaction specificity (edge-type binding specificity) is essential for assembly of target triply-periodic polyhedral structures, small capsid structures can assemble without interaction specificity, i.e. from systems of identical subunits, within certain ranges of bending and stretching moduli (64–66, 69). However, to form larger ($T > 7$) capsids with icosahedral symmetry, interaction specificity (42) or templating (42, 65) is essential, and interaction specificity significantly increases target yields and robustness to parameter variations for smaller capsids (68). Our observation that, in addition to interaction specificity, a minimal level of geometric specificity is essential to form a target triply-periodic polyhedral structure is also consistent with capsids. Even when interaction specificity allows for only a single ground state capsid structure, malformed structures assemble under conditions of low geometric specificity and/or strong interactions because mis-bound subunits do not have time to anneal before becoming trapped in the assembly (68, 71–86). On the other hand, too much geometric specificity leads to low kinetic cross-sections and thus slow assembly (72, 78). These relative assembly kinetics are also borne out by variable temperature simulations (at fixed monomer concentration) of G_7 in Fig. S7, which show that changing temperature for the triply-periodic polyhedra has analogous effects as in other self-assembly systems. For intermediate and stiff binding (ii and iii in Fig. 4), which respectively exhibit productive reversible assembly or unproductive slow assembly at the default temperature, increasing temperature decreases growth rates due to the increased entropic cost of binding. Growth eventually goes to zero above a melting temperature that decreases with increasing stiffness. In contrast, the highly flexible binding case (i in Fig. 4), which exhibits overly rapid, nearly irreversible, defective assembly at the default temperature, has a nonmonotonic growth rate over the range of temperatures studied. Increasing temperature initially facilitates growth by enabling more reversible, less defective assembly which favors further growth. The increased reversibility is reflected in a decreasing residual strain. Growth rates ultimately decrease with increasing temperature due to the entropic cost for binding.

Briefly, we note that while we have focused on growth of pre-seeded structures as a primary metric of efficiency of assembly, the dependence of nucleation rates on the geometric specificity of binding is consistent with our observations on growth rates (8). Fig. S8 compares unseeded simulation trajectories of G_7 structures for high- and intermediate binding flexibilties, i.e. conditions (i) and (ii) in Fig. 4, while the stiffest case (iii) was not observed to nucleate on the time scale of the simulation (10^9 MC sweeps). Mean nucleation times increase with stiffness, implying a qualitatively similar trend to the decreasing growth rates with increased stiffness observed in Fig. 4.

Defect-mediated mis-assembly

In our model of triply-periodic polyhedra, bound subunits have perfect type-specificity, and hence, off-target assemblies must have the same local network of subunits but the wrong global geometry. Here, we show that the primary mechanism of mis-assembly derives from topological defects of the quasi-2D crystalline assembly. These defects take the form of point disclinations, defined relative to the target polyhedral assembly (see Fig. 5A), with an angular wedge of the triangular mesh removed or added relative to its ideal geometry as one encircles a vertex. As in the standard convention for wedge disclinations in 2D crystals (87), we associate the topological charge s of a disclination to the excess degree of bond rotation around a vertex relative to the target structure. This charge can be defined and measured at any given vertex (see Methods and Supporting Fig. S9). As edge binding only takes place between complementary edge-types, such disclinations are only possible at vertices of rotational symmetry in the target assembly, i.e. at vertices located at Wyckoff positions in the target assembly. This means, for example for the $Ia\bar{3}d$ space group of the G_7 structure, the $\bar{3}$ symmetry of the 16a position supports $s = \pm 2\pi/3$ disclinations, while the $\bar{4}$ symmetry of the 24d position supports $s = \pm\pi$ disclinations (5A) [§].

Fig. 5B shows the average number of defects per unit cell $n(s)$ for different disclination types in simulations of $T = 7$ assemblies, for the same range of flexibilities considered in Fig. 4. Notably, these defects appear when bonds are flexible, coincident with the regime of large residual strains Ξ in off-target assembly, indicating that disclination formation is the primary mechanism of mis-assembly. Surprisingly, this defect population is biased in the *sign* of topological charge. That is, defects with either $s = +2\pi/3$ or $+\pi$ charge form at finite density, corresponding to patches with wedges *removed* relative to the stress-free target geometry, but the negatively-charged variants of these defects do not form at significant densities even in highly bendable or stretchable assemblies for these parameters [¶]. This same bias is also observe for P, G and D assemblies for variable T numbers, as shown in Supporting Fig. S10. Similarities between the local elastic energies of positive and negative disclination types in Supporting Fig. S11A suggest that this imbalance is not driven by differences in strains generated by these defects types. Instead, we find by simulation of variable excluded volume sizes (Supporting Fig. S11B) that the bias towards $s = +2\pi/3$ or $+\pi$ relative to their negative counterparts is a feature of steric interactions between triangular units: volume exclusion tends to penalize crowding *excess* triangular units around a shared vertex required by $s < 0$ defects.

Finally, we note that the residual net topological charge of defects has consequences for the gross morphology of off-target assemblies. In a 2D crystal that grows isotropically, a finite disclination charge density would tend to generate elastic energies that grow superextensively, i.e. faster than the number of subunits (87). Instead, we find a morphological transition in defective assemblies that mitigates growth of elastic energy with assembly. Fig. 5C compares the graph distance d_B of vertices to a *free boundary* in the assembly for the conditions corresponding to off-target vs. on-target assembly in Fig.4, points (i) and (ii) respectively. Notably, off-target mis-assembly results in narrow, stringy structures ($d_B \lesssim 2$), while on-target assembly of sufficiently rigid subunits results in bulk 2D assembly with the free boundary extending far away from interior subunits. The stringy morphology of off-target mis-assembly has the important effect of reducing the far-field elastic cost of disclinations, as the long-range effects of defects are screened by the presence of free boundaries (70, 90). Hence, analogous to the anisotropic domain growth in curvature-frustrated 2D crystals (91–93) or filaments (94), we argue that this narrow, strip-like morphology eludes the superextensive costs that would be otherwise be generated by finite disclination charge densities in isotropically growing 2D crystals.

In Fig. S12 we show results for much longer simulation times for G_7 for the cases of low and intermediate flexibility shown in Fig. 4, (i and ii, respectively), to test the ability to assemble large, multi-unit-cell crystals. These results show that running upwards of $\sim 10^8$ MC sweeps leads to assemblies of $\mathcal{N}_{\text{total}} \sim 10^3 - 10^4$ subunits, an order of magnitude larger than the 336 subunits per cubic repeat of G_7. Notably, large crystals formed at intermediate flexibility exhibit only nominal residual strains, consistent with a low density of defects, and bulk-like morphologies, in contrast to the quasi-1D, highly-defective assemblies formed by flexible subunits.

Discussion

In summary, we have extended the economical design principles of the CK construction of closed shells to triply-periodic, negative curvature programmable assembly. In both cases, economy derives from the commensurabilty of the symmetries of a subtriangulation with the symmetry elements of the target structure. For triply-periodic polyhedra, this requires constraining the vertices and centers of the base tile to the Wyckoff sites of appropriate symmetry. Preserving these symmetries in the sub-triangulation guarantees that the sub-triangulation is composed of "redundant" copies of relatively few symmetry-inequivalent particles.

However, while the high-symmetry is necessary for design economy, it is also the source of off-target misassembly. No-

[§] Along with this purely topological notion of defects, one can also define a notion of "geometric defects" that is standard to notions of order embedded in curved manifolds (88), and can be related to discrete measures of Gaussian curvature in triangular meshes, namely the excess/deficit of sum of interior angles meeting a vertices (89). Hence, the target (stress free) state of triply-periodic polyhedra exhibits spatially distributed Gaussian curvature both due to the presence of 8-coordinated vertices as well as variable edge lengths in the target mesh.

[¶] Our definition of disclination charge accounts for the effective negative Gaussian curvature of the target crystal, but defines excess bond rotation relative to a target bond coordination that may be larger than 6 (e.g. the 8-coordinated vertex at position 24d of G structures).

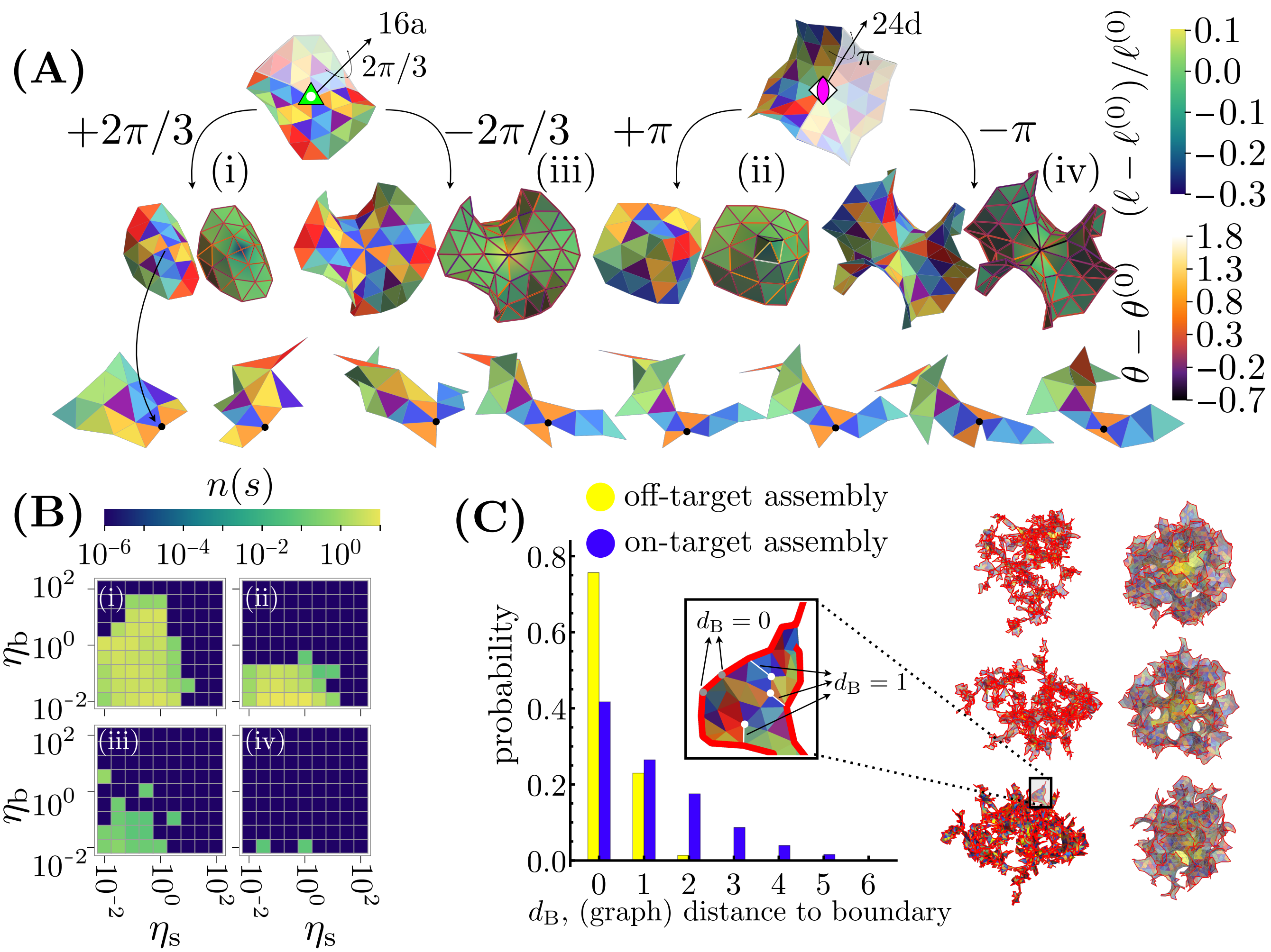


Fig. 5. Disclination pathways to off-target assembly. (A) Positive and negative topological defects compatible with the self-assembly matching rules of a G_7 structure, corresponding respectively to fractional edge removed or added from the target assembly, extending from a high-symmetry vertex. Here, possible defects are constructed by a Volterra-like construction on defectless patches centered at the Wyckoff sites 16a and 24d, where the topological charge s quantifies the excess/defective of rotation angle around the disclination. Each defective patch is colored with respect to the unique triangle species as shown on the left patches as well as their length strain (faces) and angle strain (edges) as shown on the right patches. The snapshot sequence on the bottom shows a possible assembly pathway of a topological defect with charge $s = +2\pi/3$. **(B)** Number of disclinations formed per primitive cell, $n(s)$, as a function of dimensionless ratios η_{s} and η_{b} for each of the defect types labeled from (i) to (iv) for simulated G_7 assembly. **(C)** Blue (Yellow): probability of a vertex to have a graph distance, d_{B}, to the boundary of a structure exhibiting on-target (off-target) assembly. The structure snapshots on the right (left) column show different views of the on-target (off-target) structures. The enlarged box shows different examples of vertices having a graph distance $d_{\mathrm{B}} = 0$ and $d_{\mathrm{B}} = 1$. Off- and on-target assembly correspond to points (i) and (ii) in Fig. 4, respectively.

tably, the very same rotational symmetries that anchor the sub-tilings of P_T, D_T and G_T are sites where disclinations are possible, and these disclinations proliferate if the geometric specificity of the binding between subunits is too low. Indeed, as the example in Supporting Fig. S13 illustrates, certain T values ($T = 4, 12, 16, \ldots$) lead to an additional set of (2-fold) Wyckoff positions, enabling the formation of a third set of $n\pi$ disclinations and thus more assembly errors. This trade-off between design economy and the propensity for misassembly is unavoidable. It leads to narrowly defined regimes where assembly is flexible enough to occur at reasonable rates, but specific enough to suppress disclinations. The design criteria for efficient and high-fidelity assembly (i.e. $5 \lesssim \eta_{\mathrm{s}} \lesssim 10$ and $5 \lesssim \eta_{\mathrm{b}} \lesssim 10$) are thus critical for the experimental design and realization of size-controlled crystals. In the context of programmable DNA triangles, experimental yields of off-target tubule assembly suggest a range of bending stiffness to binding ratio in the range $\eta_{\mathrm{b}} \approx 0.1 - 1$, and dimensional arguments suggest η_{s} to be in a similar range, which is notably in the range of productive and high-fidelity assembly of our physical model (44).

These estimates suggest this system may be ideal for harnessing the economy of P_T, D_T and G_T for programmed assembly of crystal structures with unit cells that are tunable to dimensions larger than those of the subunits, which is currently a challenge with assembly of colloidal or supramolecular building blocks. This limitation applies to current approaches to program the assembly of complex crystals of DNA functionalized particles (34, 35), as well as DNA origami "voxels" (36, 37). Notably, the nanometric size of programmable building blocks typically puts the photonic bandgap behavior far outside of the range of the optical regime for DNA-programmable crystals, as unit cells sizes have been typically limited to within a few times the subunit size. Hence it would be advantageous to use P_T, D_T and G_T assemblies as platforms for bottom-up design of photonic materials, with wavelength tunable via T. For example, taking computed bandgaps for gyroidal crystals (12) and using the ~ 50 nm size of DNA origami triangles (42), suggests that photonic behavior occurs in the visible range for $T = 4-9$.

We note that the photonic structures appearing in biological structures have slightly different symmetries than the TPMS structures we have studied here. That is, while TPMS symmetries correspond to so-called "double network" architectures (e.g. $Ia\bar{3}d$ spacegroup for G) and such structures are observed in membranous biological assemblies (95, 96), photonic bandgap nanostructures formed in butterflies and beetles (12, 14) correspond to "single network" architectures which break the symmetry between the two-interpenetrating channels (e.g. $I4_132$ for single-gyroids). It is, of course, straightforward to generalize the design scheme studied here for TPMS derived structures to assemblies that target these lower-symmetry analogs, for example by projection of sub-triangulated vertex positions onto *constant-mean curvature* variants of TPMS (97, 98). In the context of design economy, this symmetry reduction (i.e. reducing $\bar{3}$ symmetry at the center of the hexagonal base to 3-fold rotation) would come at the notable cost of doubling the number of required subunits relative to the higher symmetry double-network structures.

Finally, we note that recent approaches to synthetic protein design have assembled icosohedral shells of highly-modular size and structure, realizing CK structures in the range of $T = 4-100$ (99). Thus, we anticipate that our "inverted CK" design principles could be a template for engineering new classes of triply-periodic, protein-based mesoporous frameworks of controllable periodicity and symmetry.

Supporting Information Appendix (SI)

Supporting Information Appendix (SI). Supporting text appendices are provided to detail the construction algorithm for P_T, G_T and D_T assemblies, simulation methods, and analysis of defects and off-target assembly.

SI Movies. Supporting Movie S1 (P3.avi) Supporting Movie S2 (G3.avi) Supporting Movie S3 (D3.avi) Supporting Movie S4 (G7_i.avi) Supporting Movie S5 (G7_ii.avi) Supporting Movie S6 (G7_iii.avi)

Materials and Methods

Triply-periodic triangulations and interaction rules. Our construction of triangulations of P, G and D are based on projecting portions of planar triangular graphs on the level-set models of these minimal surfaces, in a way the preserves the symmetries of the *246 tiling of the $\mathbb{H}^2$ in the respective cubic space groups of $\mathbb{E}^3$ (59). In brief, this construction begins with a triangular base, 1/6 of the hexagonal patch. The vertices this triangular base, denoted $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$, are constrained to lie on Wyckoff site positions with appropriate rotoinversion symmetries to embedded the *246 tiling. As shown in Table 1, $\mathbf{a}$ is placed at the $\bar{3}$ center of hexagonal patch, while $\mathbf{b}$ and $\mathbf{c}$ lie on $\bar{4}$ points, corresponding to vertex where four hexagonal patches meet. Last, we note that inversion "flips" the normal to triangular particles, so that this hexagonal base is therefore constructed by a single symmetry equivalent unit. This base triangle itself, when embedded into the respective $Im\bar{3}m$, $Ia\bar{3}d$ and $Pn\bar{3}m$ space groups thus constitutes the $T = 1$ triangulation of P, G and D.

TPMS	$\mathbf{a}$	$\mathbf{b}, \mathbf{c}$
P ($Im\bar{3}m$)	$8c$	$12d$
G ($Ia\bar{3}d$)	$16a$	$24d$
D ($Pn\bar{3}m$)	$4c$	$6d$

Table 1. Wyckoff site symmetries and locations for the vertices the base triangulation tile (i.e. 1/6 of the fundamental hexagon). Notice that vertices b and c share the same Wyckoff site symmetry.

Higher T number triangulations follow from a procedure (described in detail in SI Text Sec. 1) where by the planar base triangle $\mathbf{a}, \mathbf{b}, \mathbf{c}$, is subtriangulated according the identical construction as CK, followed by a projection of the vertex positions from the planar bases (arranged according to the space group symmetries) onto a level set model of P, G or D via simple gradient flow.

This procedure, in general, results in distortions of the dihedral angles between edge-sharing triangular faces, as well as lengths of edges, geometric information which we then record and define to set the *target values* on triangular subunits and their selective interactions (see SI Text Sec. 2).

Grand Canonical MC simulation. For the assembly simulations, we use a model previously developed for icosahedral shell self assembly (45, 64, 100, 101) and then adapted by us for arbitrary triangle design (43, 63). Subunits in the model are flexible triangles which can bind to each other along an edge. The local preferred curvature is modelled by preferred dihedral angles between neighboring faces sharing a bond (edge) and any deviation from the preferred dihedral angle has a corresponding bending energy cost. The energy associated to triangle-triangle binding, edge stretching and dihedral change (bending) are shown in Eq. 2.

Each triangular subunit consists of 3 edges, each of which may be of distinct interaction type. The interaction matrix defines which type is allowed to bind to which type. If there are n_s subunit species in the simulation, there are at most $3n_s$ edge types with $(3n_s)^2$ different interactions, i.e. binding energies, bending moduli and dihedral angles. In addition, each of the $3n_s$ edge types may have their own stretching moduli and rest lengths. In the simulations presented in the main text, we fixed the binding energies for all allowed edge pair types to the same value. Similarly, bending moduli for all pairs and stretching moduli for all edge types are also set to the same value. Moreover, each allowed edge type pair has its own dihedral angle.

The simulation follows the growth of a single structure in the grand canonical ensemble, i.e. the structure is immersed in a bath of freely diffusing subunits held at fixed concentration. Concentrations (or, equivalently, the chemical potentials) for all species are set to be equal. The dynamics is governed by a series of Monte Carlo moves with fixed relative rates. The moves allow for subunit exchange between the structure and the bath, internal binding and unbinding of edges and thermal fluctuation of vertices with no change in topology. There are a total number of 11 moves and each move is carefully designed to satisfy detailed balance with its reverse move (63).

Calculation of defect charges. Given a structure assembled with the matching rules of triply-periodic polyhedra, we can determine disclination charges by considering closed paths around encircling the disclinations like the one shown on the Supporting Fig. S9. Each path can be in general seen as a series of steps in which each individual step consists of a composition of two rotations: an initial rotation of angle ϕ_{ij} corresponding to the angle between two consecutive edges E_i and E_j and a second rotation of angle θ_j corresponding to the dihedral angle associated with the edge E_j (s). This approach closely follows the formalism introduced by belcastro and Hull in which origami folding patterns are viewed as collections of affine transformations around the internal vertices of the patterns (102). Furthermore, moving around the vertex or rotating the surface around the vertex are analogous operations so we can perform all the ϕ and θ rotations around axes passing through the enclosed vertex and parallel to $\hat{\mathbf{z}}$ and $\hat{\mathbf{x}}$ respectively. One full rotation amounts then to the composition of rotation operations (from right to left) $\hat{R}_v(\Phi, \hat{\mathbf{n}})$:

$$\begin{aligned}\hat{R}_v(\Phi, \hat{\mathbf{n}}) \quad &= \quad \hat{R}\,(\theta_0,\, \hat{\mathbf{x}})\,\hat{R}\,(\phi_{n0},\, \hat{\mathbf{z}})\ldots \\ &\quad\quad \hat{R}\,(\theta_2,\, \hat{\mathbf{x}})\,\hat{R}\,(\phi_{12},\, \hat{\mathbf{z}})\,\hat{R}\,(\theta_1,\, \hat{\mathbf{x}})\,\hat{R}\,(\phi_{01},\, \hat{\mathbf{z}})\,, \quad [5]\end{aligned}$$

where Φ is the desired angle around disclination vertex v whose absolute value can be found can be determined as $|\Phi| = \arccos[(\mathrm{Tr}(\hat{R}_v) - 1)/2]$. With the rotation angle Φ the disclination charge of a defect is defined as the angle deficit $s = 2\pi - |\Phi|$.

ACKNOWLEDGMENTS. The authors thank S. Fraden and D. Hayakawa for help with rendering of cryoEM reconstructions DNA origami assemblies and also B. Rogers for many useful comments on this manuscript. This work was primarily support by the NSF through the Brandeis Center for Bioinspired Soft Materials, an NSF MRSEC (DMR-2011846). The authors acknowledge additional support from the German Federal Ministry of Education and Research under grant number 031L0160 and European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 829010 (CMD); European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101026118 (BT); NSF grant DMR-2309635 (MFH); and NSF grant DMR-2217543 (CDS). Simulations where performed using the UMass Cluster at the Massachusetts Green High Performance Computing Center and the High Performance Computing Cluster of MPI-CBG and CSBD. The authors gratefully acknowledge the seminal work of Alan Schoen on geometry of triply-periodic minimal surfaces that was crucial inspiration for this work.

1. JMT Thompson, AR Parker, A vision for natural photonics. *Philosophical Transactions Royal Society London. Series A: Mathematical, Physical Engineering Sciences* **362**, 2709–2720 (2004).
2. RO Prum, ER Dufresne, T Quinn, K Waters, Development of colour-producing β-keratin nanostructures in avian feather barbs. *Journal The Royal Society Interface* **6** (2009).
3. AC Neville, *Biology of Fibrous Composites*. (Cambridge University Press, Cambridge), (1993).
4. P Fratzl, Cellulose and collagen: from fibres to tissues. *Current Opinion Colloid & Interface Science* **8**, 32–39 (2003).
5. JD Perlmutter, MF Hagan, Mechanisms of virus assembly. *Annual Review Physical Chemistry* **66**, 217–239 (2015) PMID: 25532951.
6. CA Kerfeld, C Aussignargues, J Zarzycki, F Cai, M Sutter, Bacterial microcompartments. *Nature Reviews Microbiology* **16**, 277–290 (2018).
7. M Rother, MG Nussbaumer, K Renggli, N Bruns, Protein Cages and Synthetic Polymers: A Fruitful Symbiosis for Drug Delivery Applications, Bionanotechnology and Materials science. *Chem. Soc. Rev.* **45**, 6213–6249 (2016).
8. MF Hagan, GM Grason, Equilibrium mechanisms of self-limiting assembly. *Reviews Modern Physics* **93**, 025008 (2021) Publisher: American Physical Society.
9. F Oosawa, S Asakura, *Thermodynamics of the Polymerization of Protein*. (Academic Press, London), (1975).
10. Y Bouligand, Liquid crystals and biological morphogenesis: Ancient and new questions. *Comptes Rendus Chimie* **11**, 281–296 (2008).
11. K Michielsen, D Stavenga, Gyroid cuticular structures in butterfly wing scales: biological photonic crystals. *Journal The Royal Society Interface* **5**, 85–94 (2008).
12. V Saranathan, et al., Structure, function, and self-assembly of single network gyroid (I4132) photonic crystals in butterfly wing scales. *Proceedings National Academy Sciences* **107**, 11676–81 (2010).
13. BD Wilts, K Michielsen, J Kuipers, H De Raedt, DG Stavenga, Brilliant camouflage: photonic crystals in the diamond weevil, entimus imperialis. *Proceedings Royal Society B: Biological Sciences* **279**, 2524–2530 (2012).
14. V Saranathan, S Narayanan, A Sandy, ER Dufresne, RO Prum, Evolution of single gyroid photonic crystals in bird feathers. *Proceedings National Academy Sciences* **118**, e2101357118 (2021).
15. Y Ke, LL Ong, WM Shih, P Yin, Three-Dimensional Structures Self-Assembled from DNA Bricks. *Science* **338**, 1177–1183 (2012).
16. MR Jones, NC Seeman, CA Mirkin, Programmable materials and the nature of the DNA bond. *Science* **347**, 1260901–1260901 (2015).
17. Z Zeravcic, VN Manoharan, MP Brenner, Colloquium : Toward living matter with colloidal particles. *Reviews Modern Physics* **89**, 031001 (2017).
18. SM Douglas, et al., Self-assembly of dna into nanoscale three-dimensional shapes. *Nature* **459**, 414–418 (2009).
19. PS Huang, BS E., D Baker, The coming of age of $de\ novo$ protein design. *Nature* **537**, 320–327 (2016).
20. SC Glotzer, MJ Solomon, Anisotropy of building blocks and their assembly into complex structures. *Nature Materials* (2007).
21. S Sacanna, DJ Pine, Shape-anisotropic colloids: Building blocks for complex assemblies. *Current Opinion Colloid & Interface Science* **16**, 96–105 (2011).
22. T Heuckel, GM Hocky, S Sacanna, Total synthesis of colloidal matter. *Nature Reviews Materials volume* **6**, 1053–1069 (2021).
23. Z Su, et al., The role of architectural engineering in macromolecular self-assemblies via non-covalent interactions: A molecular LEGO approach. *Progress Polymer Science* **103**, 101230 (2020).
24. A Murugan, J Zou, MP Brenner, Undesired usage and the robust self-assembly of heterogeneous structures. *Nature Communications* **6**, 6203 (2015).
25. A Murugan, Z Zeravcic, MP Brenner, S Leibler, Multifarious assembly mixtures: Systems allowing retrieval of diverse stored structures. *Proceedings National Academy Sciences* **112**, 54–59 (2015).
26. WM Jacobs, A Reinhardt, D Frenkel, Rational design of self-assembly pathways for complex multicomponent structures. *Proceedings National Academy Sciences* **112**, 6313–6318 (2015).
27. WM Jacobs, D Frenkel, Self-assembly of structures with addressable complexity. *Journal American Chemical Society* **138**, 2457–2467 (2016) PMID: 26862684.
28. Z Zeravcic, VN Manoharan, MP Brenner, Size limits of self-assembled colloidal structures made using specific interactions. *Proceedings National Academy Sciences* **111**, 15918–15923 (2014).
29. Y Ke, et al., DNA brick crystals with prescribed depths. *Nature Chemistry* **6**, 994–1002 (2014).
30. LL Ong, et al., Programmable self-assembly of three-dimensional nanostructures from 10,000 unique components. *Nature* **552**, 72–77 (2017).
31. D Nykypanchuk, MM Maye, D Van Der Lelie, O Gang, DNA-guided crystallization of colloidal nanoparticles. *Nature* **451**, 549–552 (2008).
32. E Auyeung, et al., Synthetically programmable nanoparticle superlattices using a hollow three-dimensional spacer approach. *Nature Nanotechnology* **7**, 24–28 (2012).
33. Y Tian, et al., Ordered three-dimensional nanomaterials using dna-prescribed and valence-controlled material voxels. *Nature Materials* **19**, 789–796 (2020).
34. WB Rogers, JC Crocker, Direct measurements of dna-mediated colloidal interactions and their quantitative modeling. *Proceedings National Academy Sciences* **108**, 15687–15692 (2011).
35. A Hensley, WM Jacobs, WB Rogers, Self-assembly of photonic crystals by controlling the nucleation and growth of dna-coated colloids. *Proceedings National Academy Sciences* **119**, e2114050118 (2022).
36. PW Majewski, et al., Resilient three-dimensional ordered architectures assembled from nanoparticles by dna. *Science Advances* **7**, eabf0617 (2021).
37. A Michelson, et al., Three-dimensional visualization of nanoparticle lattices and multimaterial frameworks. *Science* **376**, 203–207 (2022).
38. SH Park, H Park, K Hur, S Lee, Design of dna origami diamond photonic crystals. *ACS Applied Bio Materials* **3**, 747–756 (2020) PMID: 35019418.
39. T Gerling, KF Wagenbauer, AM Neuner, H Dietz, Dynamic dna devices and assemblies formed by shape-complementary, nonâbase pairing 3d components. *Science* **347**, 1446–1452 (2015).
40. PWK Rothemund, et al., Design and characterization of programmable dna nanotubes. *Journal American Chemical Society* **126**, 16344–16352 (2004) PMID: 15600335.
41. E Benson, et al., Dna rendering of polyhedral meshes at the nanoscale. *Nature* **523**, 441–444 (2015).
42. C Sigl, et al., Programmable icosahedral shell system for virus trapping. *Nature Materials* **20**, 1281–1289 (2021).
43. TE Videbaek, et al., Tiling a tubule: how increasing complexity improves the yield of self-limited assembly. *Journal Physics: Condensed Matter* **34**, 134003 (2022).
44. D Hayakawa, et al., Geometrically programmed self-limited assembly of tubules using dna origami colloids. *Proceedings National Academy Sciences* **119**, e2207902119 (2022).
45. R Zandi, D Reguera, RF Bruinsma, WM Gelbart, J Rudnick, Origin of icosahedral symmetry in viruses. *Proceedings National Academy Sciences* **101**, 15556–15560 (2004).
46. R Twarock, A Luque, Structural puzzles in virology solved with an overarching icosahedral design principle. *Nature Communications* **10**, 4414 (2019).
47. A Siber, Icosadeltahedral geometry of geodesic domes, fullerenes and viruses: A tutorial on the t-number. *Symmetry* **12** (2020).
48. JE Johnson, JA Speir, Quasi-equivalent viruses: A paradigm for protein assemblies. *J. Mol. Biol.* **269**, 665–675 (1997).
49. DLD Caspar, A Klug, Physical Principles in the Construction of Regular Viruses. *Cold Spring Harbor Symposia on Quantitative Biology* **27**, 1–24 (1962).
50. J Bohlin, AJ Turberfield, AA Louis, P Å ulc, Designing the self-assembly of arbitrary shapes using minimal complexity building blocks. *ACS Nano* **17**, 5387–5398 (2023) PMID: 36763807.
51. DEP Pinto, P Sulc, F Sciortino, J Russo, Design strategies for the self-assembly of polyhedral shells. *Proceedings National Academy Sciences* **120**, e2219458120 (2023).
52. MC Pedersen, ST Hyde, Polyhedra and packings from hyperbolic honeycombs. *Proceedings National Academy Sciences* **115**, 6905–6910 (2018).
53. H Tanaka, T Dotera, ST Hyde, Programmable self-assembly of nanoplates into bicontinuous nanostructures. *ACS Nano* **0**, null (0) PMID: 37527198.
54. EA Lord, AL Mackay, Periodic minimal surfaces of cubic symmetry. *Current Science* **85**, 346–362 (2003).
55. AH Schoen, Reflections concerning triply-periodic minimal surfaces. *Interface Focus* **2**, 658–668 (2012).
56. GE Schroeder, SJ Ramsden, AG Christy, ST Hyde, Medial surfaces of hyperbolic structures. *The European Physical Journal B - Condensed Matter* **35**, 551–564 (2003).
57. T Dotera, H Tanaka, Y Takahashi, Hexagulation numbers: the magic numbers of equal spheres on triply periodic minimal surfaces. *Structural Chemistry* **28**, 105–112 (2017).
58. JF Sadoc, J Charvolin, Infinite periodic minimal surfaces and their crystallography in the hyperbolic plane. *Acta Crystallographica Section A* **45**, 10–20 (1989).
59. SJ Ramsden, V Robins, ST Hyde, Three-dimensional Euclidean nets from two-dimensional hyperbolic tilings: kaleidoscopic examples. *Acta Crystallographica Section A* **65**, 81–108 (2009).
60. MC Pedersen, ST Hyde, Hyperbolic crystallography of two-periodic surfaces and associated

structures. *Acta Crystallographica Section A* **73**, 124–134 (2017).
61. MC Pedersen, ST Hyde, S Ramsden, JJK Kirkensgaard, Mapping hyperbolic order in curved materials. *Soft Matter* **19**, 1586–1595 (2023).
62. CT Chantler, F Boscherini, B Bunker, eds., *International Tables for Crystallography: X-ray absorption spectroscopy and related techniques.* (International Union of Crystallography, Chester, England) Vol. I, 1 edition, (2020).
63. B Tyukodi, F Mohajerani, DM Hall, GM Grason, MF Hagan, Thermodynamic Size Control in Curvature-Frustrated Tubules: Self-Limitation with Open Boundaries. *ACS Nano* (2022).
64. GM Rotskoff, PL Geissler, Robust nonequilibrium pathways to microcompartment assembly. *Proceedings National Academy Sciences United States America* **115**, 6341–6346 (2018).
65. S Li, P Roy, A Travesset, R Zandi, Why large icosahedral viruses need scaffolding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 10971–10976 (2018).
66. S Panahandeh, et al., How a virus circumvents energy barriers to form symmetric shells. *ACS Nano* **14**, 3170–3180 (2020).
67. H Fang, B Tyukodi, WB Rogers, MF Hagan, Polymorphic self-assembly of helical tubules is kinetically controlled. *Soft Matter* **18**, 6716–6728 (2022).
68. F Mohajerani, et al., Multiscale modeling of hepatitis b virus capsid assembly and its dimorphism. *ACS Nano* **16**, 13845–13859 (2022) doi: 10.1021/acsnano.2c02119.
69. J Wagner, R Zandi, The robust assembly of small symmetric nanoshells. *Biophys. J.* **109**, 956 (2015).
70. S Li, R Zandi, A Travesset, GM Grason, Ground States of Crystalline Caps: Generalized Jellium on Curved Space. *Phys. Rev. Lett.* (2019).
71. R Schwartz, PW Shor, PE Prevelige, B Berger, Local Rules Simulation of the Kinetics of Virus Capsid Self-Assembly. *Biophys. J.* **75**, 2626–2636 (1998).
72. MF Hagan, D Chandler, Dynamic Pathways for Viral Capsid Assembly. *Biophys. J.* **91**, 42–54 (2006).
73. RL Jack, MF Hagan, D Chandler, Fluctuation-dissipation ratios in the dynamics of self-assembly. *Phys. Rev. E* **76**, 021119 (2007).
74. MF Hagan, OM Elrad, RL Jack, Mechanisms of Kinetic Trapping in Self-Assembly and Phase Transformation. *J. Chem. Phys.* **135**, 104115 (2011).
75. D Rapaport, The Role of Reversibility in Viral Capsid Growth: A Paradigm for Self-Assembly. *Phys. Rev. Lett.* **101**, 186101 (2008).
76. DC Rapaport, Studies of reversible capsid shell growth. *J. Phys.: Condens. Matter* **22**, 104115 (2010).
77. DC Rapaport, Molecular dynamics simulation of reversibly self-assembling shells in solution using trapezoidal particles. *Phys. Rev. E* **86**, 051917 (2012).
78. S Whitelam, et al., The Impact of Conformational Fluctuations on Self-Assembly: Cooperative Aggregation of Archaeal Chaperonin Proteins. *Nano Lett.* **9**, 292–297 (2009).
79. HD Nguyen, VS Reddy, CL Brooks, Deciphering the Kinetic Mechanism of Spontaneous Self-Assembly of Icosahedral Capsids. *Nano Lett.* **7**, 338–344 (2007).
80. H Nguyen, C Brooks, Generalized structural polymorphism in self-assembled viral particles. *Nano Lett.* **8**, 4574 (2008).
81. HD Nguyen, VS Reddy, CL Brooks, Invariant Polymorphism in Virus Capsid Assembly. *J. Am. Chem. Soc.* **131**, 2606–14 (2009).
82. AW Wilber, et al., Reversible Self-Assembly of Patchy Particles into Monodisperse Icosahedral Clusters. *J. Chem. Phys.* **127**, 085106 (2007).
83. AW Wilber, JPK Doye, AA Louis, ACF Lewis, Monodisperse self-assembly in a model with protein-like interactions. *J. Chem. Phys.* **131**, 175102 (2009).
84. S Cheng, A Aggarwal, MJ Stevens, Self-assembly of artificial microtubules. *Soft Matter* **8**, 5666 (2012).
85. MF Hagan, Modeling Viral Capsid Assembly. *Adv. Chem. Phys.* **155**, 1–68 (2014).
86. S Whitelam, RL Jack, The Statistical Mechanics of Dynamic Pathways to Self-assembly. *Ann Rev Phys Chem* **66**, 143–63 (2015).
87. HS Seung, DR Nelson, Defects in flexible membranes with crystalline order. *Phys. Rev. A* **38**, 1005–1018 (1988).
88. M Kléman, Curved crystals, defects and disorder. *Advances Physics* **38**, 605–667 (1989).
89. D Meek, D Walton, On surface normal and gaussian curvature approximations given data sampled from a smooth surface. *Computer Aided Geometric Design* **17**, 521–543 (2000).
90. GM Grason, Defects in crystalline packings of twisted filament bundles. i. continuum theory of disclinations. *Phys. Rev. E* **85**, 031603 (2012).
91. S Schneider, G Gompper, Shapes of crystalline domains on spherical fluid vesicles. *Europhysics Letters* **70**, 136 (2005).
92. G Meng, J Paulose, DR Nelson, VN Manoharan, Elastic Instability of a Crystal Growing on a Curved Surface. *Science* **343** (2014).
93. GM Grason, Perspective: Geometrically frustrated assemblies. *J. Chem. Phys.* **145**, 110901 (2016).
94. Y Yang, R Meyer, MF Hagan, Self-Limited Self-Assembly of Chiral Filaments. *Phys. Rev. Lett.* **104**, 258102 (2010).
95. ZA Almsherqi, T Landh, SD Kohlwein, Y Deng, Chapter 6 cubic membranes: The missing dimension of cell membrane organization in *International Review of Cell and Molecular Biology*, International Review of Cell and Molecular Biology. (Academic Press) Vol. 274, pp. 275–342 (2009).
96. E Selstam, J Schelin, WP Williams, AP Brain, Structural organisation of prolamellar bodies (plb) isolated from zea mays. parallel tem, saxs and absorption spectra measurements on samples subjected to freeze-thaw, reduced ph and high-salt perturbation. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1768**, 2235–2245 (2007).
97. DM Anderson, HT Davis, JCC Nitsche, LE Scriven, Periodic surfaces of prescribed mean curvature in *Physics of Amphiphilic Layers*, eds. J Meunier, D Langevin, N Boccara. (Springer Berlin Heidelberg, Berlin, Heidelberg), pp. 130–130 (1987).
98. K Grosse-Brauckmann, Gyroids of constant mean curvature. *Experimental Mathematics* **6**, 33–50 (1997).
99. QM Dowling, et al., Hierarchical design of pseudosymmetric protein nanoparticles, (Bioengineering), preprint (2023).
100. R Zandi, P van der Schoot, D Reguera, W Kegel, H Reiss, Classical Nucleation Theory of Virus Capsids. *Biophys. J.* **90**, 1939–1948 (2006).
101. R Zandi, B Dragnea, A Travesset, R Podgornik, On virus growth and form. *Phys. Rep.* (2020).
102. sarah-marie belcastro, TC Hull, Modelling the folding of paper into three dimensions using affine transformations. *Linear Algebra its Applications* **348**, 273–282 (2002).

Supporting Information for

Limits of economy and fidelity for programmable assembly of size-controlled triply-periodic polyhedra

Carlos M. Duque, Douglas M. Hall, Botond Tyukodi, Michael F. Hagan, Christian D. Santangelo, and Gregory M. Grason

Gregory M. Grason
E-mail: grason@umass.edu

This PDF file includes:

Supporting text
Figs. S1 to S22
Legends for Movies S1 to S6
SI References

Other supporting materials for this manuscript include the following:

Movies S1 to S6

Supporting Information Text

1. Extension of Caspar Klug construction to P,G,D triangulations

The construction of Caspar and Klug is a procedure to create structures with increasing number of subunits proportional to the triangulation number T while preserving icosahedral symmetry (1). The extension of this construction to the structures studied here was formulated previously (2), considering the dual problem of packing spheres on the surface with packing deriving from hexagonal packing of the plane. Here, we give an explicit procedure for deriving the sub-triangulations ($T > 1$) derived from base triangulations ($T = 1$) of the P,G, and D surfaces based on the principles of Caspar and Klug. In each case, the vertices of the base triangulations are identified with vertices and centers of the hexagonal base tiles associated with each structure.

The steps of the construction are outlined in Fig S14, yielding explicit coordinates for vertices of each triangle in the cubic unit cell. The $T = 1$ structure is taken as a starting point, which is then sub-divided to the desired triangulation number. Finally, each vertex is translated by gradient flow along an objective function to arrive closer to an approximation of the target surface.

From the $T = 1$ structure, both the coordinates of each triangle's vertices in the unit cell is known and also an indexing of each triangle's neighbors is derived by identifying which edges are shared by which triangles. The extended construction is to then identify a part of the planar triangular lattice with each triangles of $T = 1$ by an orthogonal projection: each edge of the $T = 1$ structure corresponds to a vector $\mathbf{L} = h\mathbf{a}_1 + k\mathbf{a}_2$ of the planar structure with standard triangular lattice vectors $\mathbf{a_1}, \mathbf{a_2}$. The triangulation indices h, k are integers that enumerate different triangulation numbers via the relation $T = h^2 + k^2 + hk$ that derives from the area of the equilateral triangle with edge $\mathbf{L}$ expressed in units of the original triangular lattice vectors' length.

Adjacent edges on the $T = 1$ structure are identified with vectors in the planar diagram $\mathbf{L}, \mathbf{U}_1 = -k\mathbf{a}_1 + (h + k)\mathbf{a}_2$ and $\mathbf{U}_2 = (h+k)\mathbf{a}_1 + -h\mathbf{a}_2$. Each vertex contained in the triangle with edges $\mathbf{L}, \mathbf{U}_1$ can be expressed as $i\mathbf{a}_1 + j\mathbf{a}_2$, and equivalently as $\frac{1}{T}(i(h + k) + jk)\mathbf{L} + \frac{1}{T}(jh - ik)\mathbf{U}_1$. This mapping of vertices is an orthographic projection, so that while the triangulation of the icosahedron in this manner will yield equilateral triangles, here isosceles triangles are found due to the isosceles triangles of $T = 1$ deriving from the asymmetric units of P,G and D. The edges of the ($T > 1$) triangulation follow from the mapping identifying $\mathbf{L}, \mathbf{U}_1$ and the additional identification of $\mathbf{U}_2$ wherever an edge of the planar diagram crosses over the vector $\mathbf{L}$.

The resulting sub-triangulations are planar except for triangles near the boundary. In the final step to construct the explicit coordinates of the $T > 1$ triangulation, in the spirit of quasi-equivalence, vertices of the sub-triangulation are each translated by gradient flow along the function f^i ($i = P, G, D$) to the nodal approximation of the target surface $f = 0$ while minimizing the distance that each vertex is translated from its initial value. This is performed by using Mathematica software to minimize $\|\mathbf{x} - \mathbf{x}_0\|^2$ subject to the constraint $f(\mathbf{x}) = 0$. While the vertices are transformed, the edge topology is kept the same by preserving the edge data according to indexing of each vertex before and after the gradient flow. This final step preserves the symmetries of the structure because each $f(x, y, z)$ has the targeted symmetries.

2. Derivation of matching rules and geometric parameters for simulation

With the cubic unit cell of structures derived using the construction of the previous section, we next consider the matching rules for building blocks that exactly address the target structure in the rigid limit, of zero compliance in either bending or stretching. We consider the minimal number of subunit types required where each subunit has three edges, each edge has a single length, dihedral angle with respect to its neighbor, and only binds with exactly one subunit type and a single corresponding edge index (see figure 2). When comparing to triangulations of icosahedral shells and other deltahedral shells, we further allow a triangle to adopt three-fold symmetry, so that each of its three edges have the same lengths, dihedrals, and bind to a single edge identity.

The types and matching rules may be identified in the unit cell directly from the identifications found by applying the discrete symmetries of the corresponding target surface. For the purpose of generating matching rules for the simulations, we only use the translational and rotational symmetries (no inversion or mirror symmetry), resulting in twice the number of subunit species with distinct species that are related by the additional symmetries not considered. Taking advantange of the rotoinversion symmetries of the ultimate triangulations in $\mathbb{E}^3$ then allows us further to identify "flipped" triangle pairs, as a single sub unit type.

In principle, the explicit construction detailed above in the previous section will preserve these symmetries from the original structure, since we are applying orthogonal projection of a three-fold symmetric structure onto the the $T = 1$ structure and gradient flow according to a function that has the same symmetry. We have developed a method to check whether the numerical procedure for construction preserves symmetry by re-deriving type and matching rule assignment from a given triangulation with vertex coordinates and indexing of triangle neighbors. The procedure is an exhaustive search of all possible assignments, made simpler by the method of constraint propagation: successive sub-symmetries are found by attempting to identify pairs of triangles and their orientations, and propagating the additional identifications that are necessary for such a structure.

The constraint propagation of an identification follows from recursive evaluations of local matches on the collection of triangles $\{t_i\}$ of the explicit triangulation of the unit cell. A list of proposed type assignments is maintained, with each t_i having a type and a rotation relative to other members of the type. Each type has an additional Boolean value assigning three-fold symmetry. At each recursion, a proposed match is t_a and t_b and a rotation r that identifies side j of t_a with side $j + r \bmod 3$ of t_b. The recursion consists of the following:

1. For all three sides $j = 1..3$, check that corresponding side lengths and dihedrals match.

2. If this local match fails, the entire propagation terminates and the type assignment is rejected.

3. If the local match succeeds but has been previously identified, no additional matching is required (this is the base case to the recursion).

4. If the local match succeeds but t_a and t_b were previously identified to be the same type but with a different rotation, and furthermore no threefold symmetry was previously identified, then threefold symmetry is proposed for this type. Two recursive calls are made to identify $j = 1$ $j = 2$ and $j = 1$ $j = 3$ neighbors of t_a, with rotations such that the corresponding edges shared with t_a are identified.

5. If the match succeeds and t_a and t_b are different types, then the proposed type assignments is updated to reflect their proposed identification. All triangles of the same type as t_b are assigned the same type as t_a and their relative rotations with respect to t_a follow from the their relation to t_b and the new proposed rotation r from t_a to t_b. Three recursive calls are made to identify neighbor j of t_a with neighbor $j + r \bmod 3$ of t_b, with the appropriate rotations that derive from the identifications of the sides of t_a and t_b, starting again at step **1** each time.

The result of applying this matching algorithm for all possible pairs and possible rotations is that a minimal number of types is identified. The appropriate side lengths and dihedrals for each type can be found from any triangle that has that type, and the matching rules are found from the types of each neighbor. This algorithm was implemented in Mathematica, which is included as supplementary information.

3. Ground state energies of topological defects and effect of excluder size

To show that the overabundance of positive defects cannot be readily attributed to some elastic favorability of positive defects, we calculated the ratio $E(+|s|)/E(-|s|)$, where $E(\pm|s|)$ are the ground state elastic energies of defect patches of charge $\pm|s|$. We performed energy minimization as a function of the ratio $\eta_{\rm s}/\eta_{\rm b} \sim k/\kappa$, with k and κ being the moduli for stretching and dihedral bending respectively. We considered defect patches for the charges $s = \pm 2\pi/3$ and $s = \pm\pi$ for G_7 using the patch topologies shown on Fig. 5A. Each energy minimization was performed in Mathematica with a conjugate gradient algorithm and using the geometry of the patches of Fig. 5A as initial conditions. In Fig. S11A we show the ratios for charge magnitudes $|s| = 2\pi/3$ and $|s| = \pi$. In general, we do not observe a regime in which positive defects have a much smaller elastic cost than negative defects. On one hand, positive defects appear to have a higher elastic energy cost than their negative counterparts for $\eta_{\rm s}/\eta_{\rm b} \lesssim 10$. On the other hand, the positive to negative elastic cost becomes comparable for $\eta_{\rm s}/\eta_{\rm b} \gtrsim 10$. These results suggest then the necessity of a different mechanism in order to rationalize the excess of positive charges.

In Fig. S11B we investigate how steric interactions affect the overall excess of positive charges for the case of G_7 structures. We probe the defect imbalance by measuring the quantities n_+ and n_- of a given structure which are defined as the total number of charges of positive and negative sign respectively. Defining $n_{\rm total} = n_+ + n_-$ and $\Delta n = n_+ - n_-$, we measure the mean relative counts difference between positive and negative charges, $\langle \Delta n/n_{\rm total}\rangle$, and explore how it varies as the size of the normalized excluder radius between adjacent subunits, $R_{\rm exc}/\langle \ell_0\rangle$, is increased. We fixed the chemical potential, $\mu = -4.5k_{\rm B}T$, and considered three different values for binding affinity between subunits, $E_{\rm bind}$. To calculate the mean values 10 independent realizations per point were considered. Simulations were run for 50×10^6 Monte Carlo steps or until the number of assembled subunits was at least $\mathcal{N}_{\rm total} = 5000$. For each $E_{\rm bind}$ we found that as $R_{\rm exc}/\langle \ell_0\rangle$ increases, $\langle \Delta n/n_{\rm total}\rangle$ tends to saturate to values ~ 1, which we interpret as a clear signature of the fact that negative defects tend to be suppressed as the excluder radius increases. In other words, the allowed configurations of multiple subunits around a given vertex is considerably affected by the size of the excluder radius of neighboring subunits.

4. Preclosure bias and origin of stringy structures

In Fig. S11C we further explore the bias for defect preclosure and calculate $\langle n(s)_{\rm uc}\rangle$, the mean total number of defects with charge s per unit-cell and study how $\langle n(s)_{\rm uc}\rangle$ varies with increasing binding affinity $\epsilon_{\rm bind}$. We work again with G_7 and fix the chemical potential and excluder radius to $\mu = -4.5k_{\rm B}T$ and $R_{\rm exc}/\langle \ell_0\rangle = 0.26$ respectively. Furthermore, we set $\eta_{\rm s} = 1$ and $\eta_{\rm b} = 0.01$ and solved for the elastic moduli k and κ by normalizing with respect to the binding affinity value, $\epsilon_{\rm bind} = -6.5k_{\rm B}T$, used throughout the main text. We used the same stopping criteria of the previous section and considered 50 independent Monte Carlo realizations. As $\epsilon_{\rm bind}/k_{\rm B}T$ grows large and edge binding moves become more favorable the amounts of positively charged defects, namely $s = +2\pi/3, +\pi$, increase while the amounts of negatively charged defect types hardly experience any increment. Only for $s = -2\pi$ we see an increase past $\epsilon_{\rm bind}/k_{\rm B}T \gtrsim 6.5$. This type of charge is compatible by any vertex whether or not they exhibit any n-fold rotational symmetry. This can be realized by adding twice as much subunits around the vertex before closure. The results hint that rapid "off-target" assembly induced by large edge binding is reflected on an overabundance of topological defects. The steric interactions due to excluders introduce an additional bias for the selection of positive defects.

In Fig. S11D we show the dependence of mean graph distance to the boundary, $\langle d_{\rm B}\rangle$, with respect to the binding affinity. We perform this calculation by measuring the graph geodesic of all the vertices composing a structure to the nearest vertex located at the boundary of the structure. In order to avoid any bias coming from the starting seed, we do not include the graph distance of vertices belonging to the seed. We additionally considered the zero, $d_{\rm B} = 0$, distance contributions of all the vertices located at the boundary of the structure. As the binding affinity increases we observe that $\langle d_{\rm B}\rangle$ increases with $\epsilon_{\rm bind}$. In general, we observe that in order to accommodate the increasing number of defects, the structures adopt "stringy" morphologies like in Fig. 5C in which defects are located near the boundaries of the structure. We noticed that $d_{\rm B} \lesssim 1$ for small $\epsilon_{\rm bind}$ while $d_{\rm B} \lesssim 3$ for larger values of $\epsilon_{\rm bind}$. Even for large $\epsilon_{\rm bind}$, however, we observe that $\langle d_{\rm B}\rangle \lesssim 1$ which suggests, at least for the binding affinity range we considered, that vertices located at the boundary of the structure tend to skew $\langle d_{\rm B}\rangle$ towards smaller values regardless of the strength of the binding affinity.

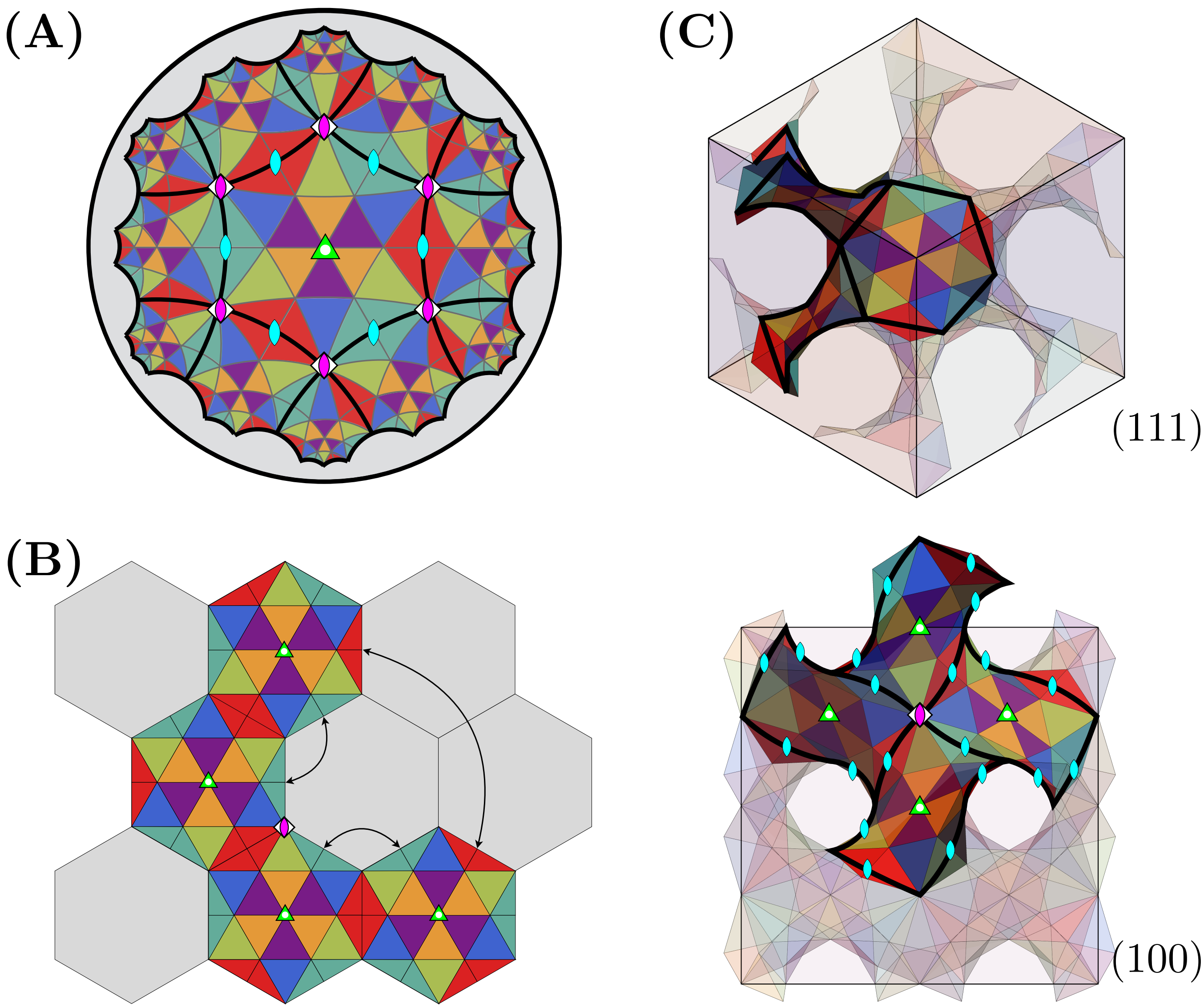


Fig. S1. (A) T=3 triangulation projected onto $\mathbb{H}^2$. The triangles are colored following the ID labels assigned to triangular subunits of T=3 triply-periodic polyhedra. The thick black lines correspond to the lines of the $\{6,4\}$ hyperbolic tessellation which can be further subdivided using a single non-Euclidean triangular mono-tile with internal angles $\{\pi/3, \pi/4, \pi/4\}$ corresponding to the T=1 case. The vertices marked with glyphs are points that map to Wyckoff sites associated to the space groups of the different TPMS. **(B)** Planar template illustrating how the sides of 4 hexagonal T=3 tiles should be identified in order to perform a folding and distorting procedure leading to the $\mathbb{E}^3$ arrangement shown on **(C)** for the G_3 case in which they all meet at a single vertex of $\bar{4}$ symmetry.

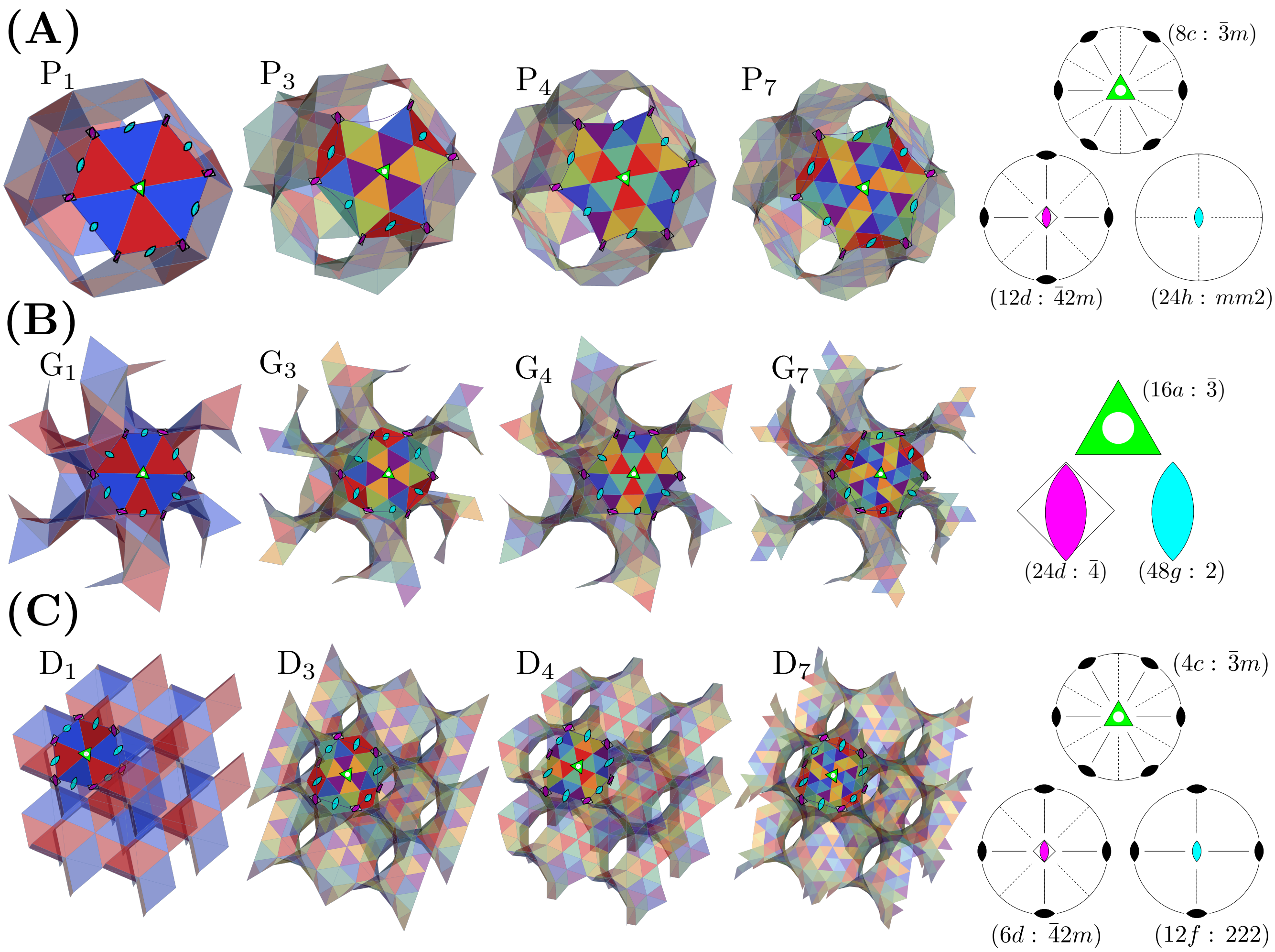


Fig. S2. Wyckoff sites of triply-periodic polyhedral crystals. **(A-C)** Wyckoff sites of different symmetry are present on the fundamental hexagonal patches of P$_T$, G$_T$ and D$_T$ respectively. The symmetry of each Wyckoff site is encoded by one of the glyphs shown on the right panels. Notice that sites with symmetries 8c, 16a, and 4c as well as 12d, 24d, and 6d are present for all shown T numbers of P$_T$, G$_T$ and D$_T$ respectively. Among these, P$_4$, G$_4$, and D$_4$ respectively, lead to an addition family of 2-fold symmetric vertices at 24h, 48g, and 12f.

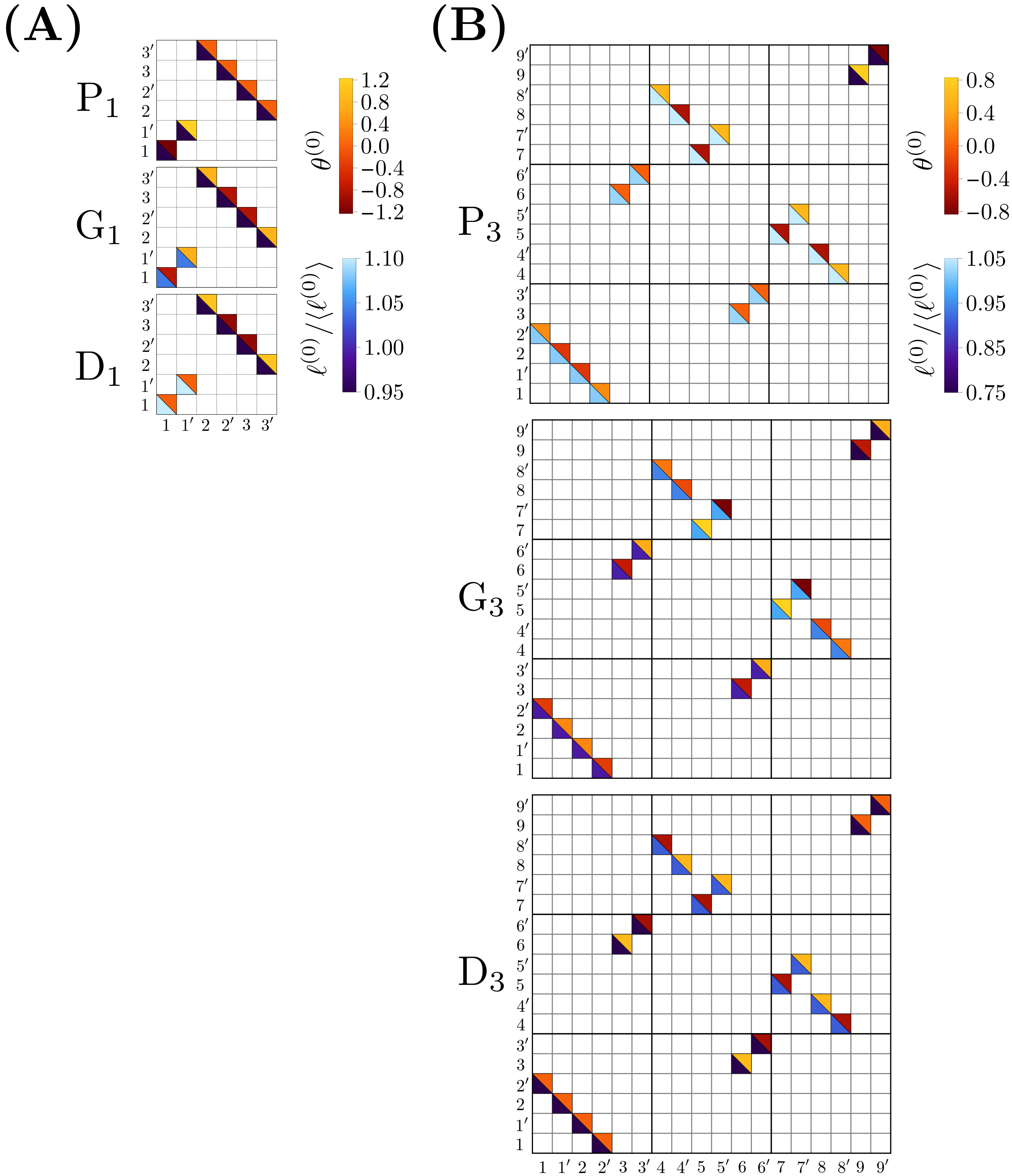


Fig. S3. (A-B) Interaction matrix for T=1 and T=3 structures respectively. Each colored block represents a valid edge-pairing in which the lower (upper) half color of each colored-block in the matrix represents the target length (dihedral angle) of a valid edge-pairing. Notice that the matrix "topologies" for a given triangulation number T are the same and we only need to specify the geometric data of each structure.

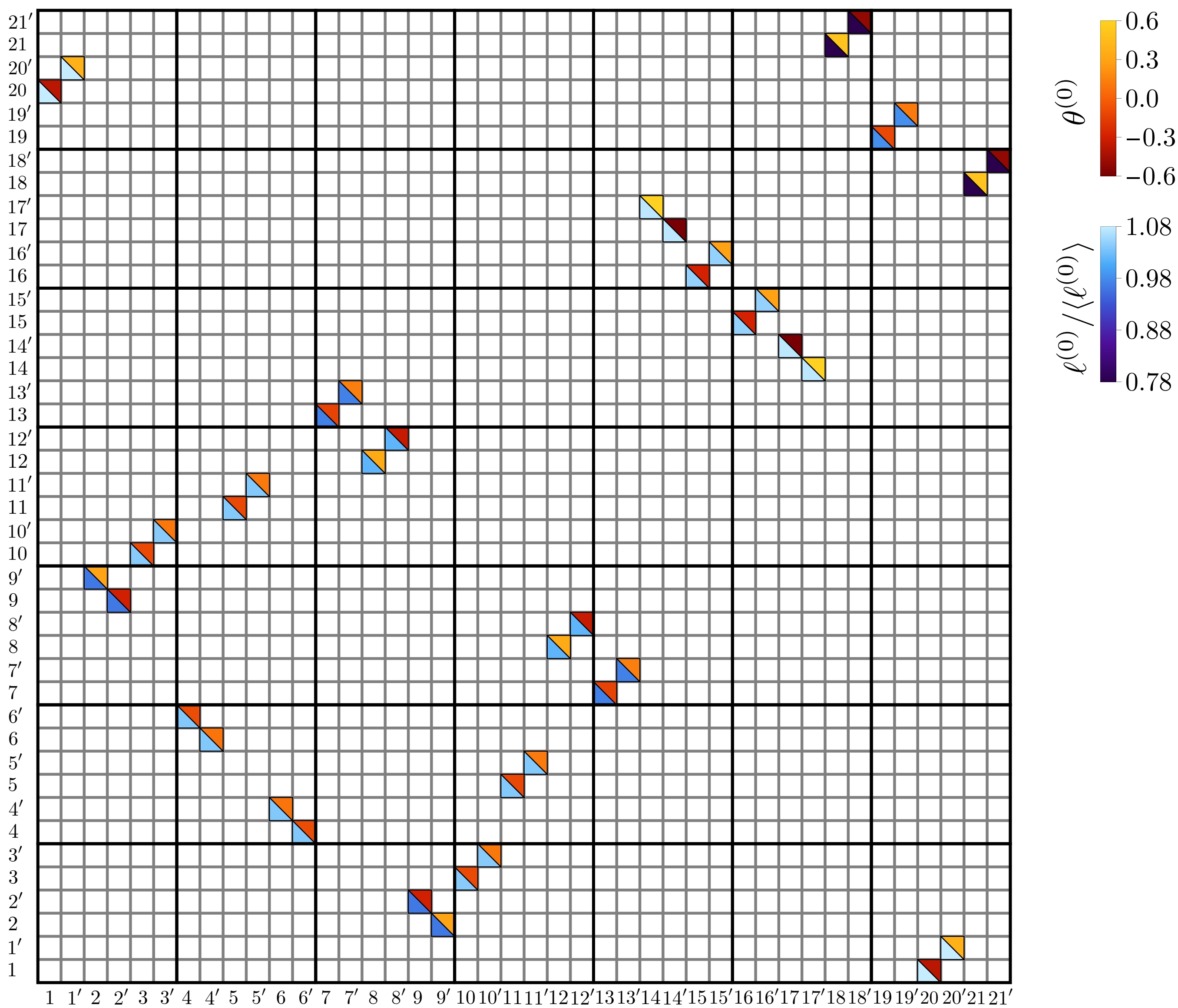


Fig. S4. Interaction matrix for P_7. Each colored block represents a valid edge-pairing in which the lower (upper) half color of each colored-block in the matrix represents the target length (dihedral angle) of a valid edge-pairing.

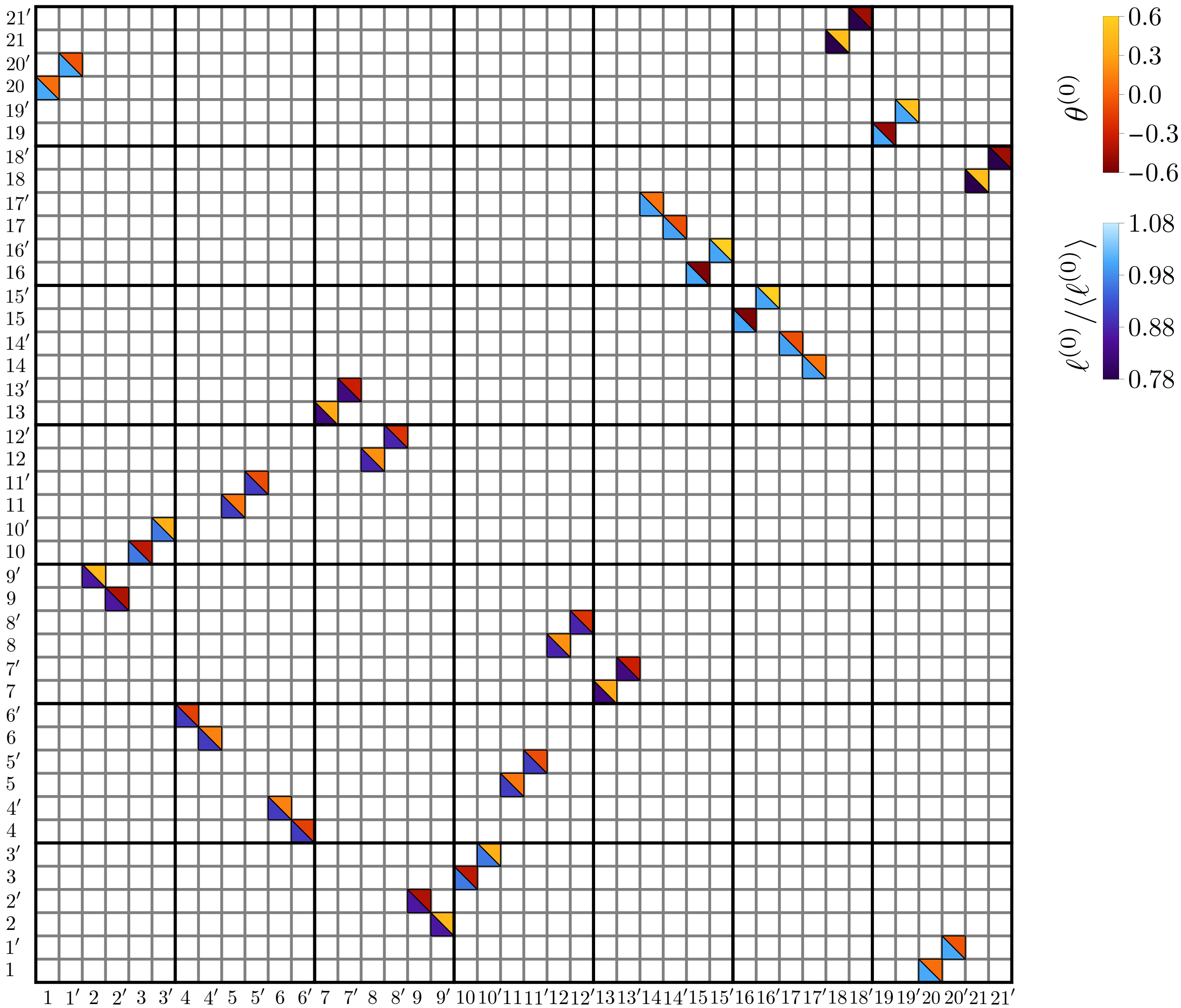


Fig. S5. Interaction matrix for G_7. Each colored block represents a valid edge-pairing in which the lower (upper) half color of each colored-block in the matrix represents the target length (dihedral angle) of a valid edge-pairing.

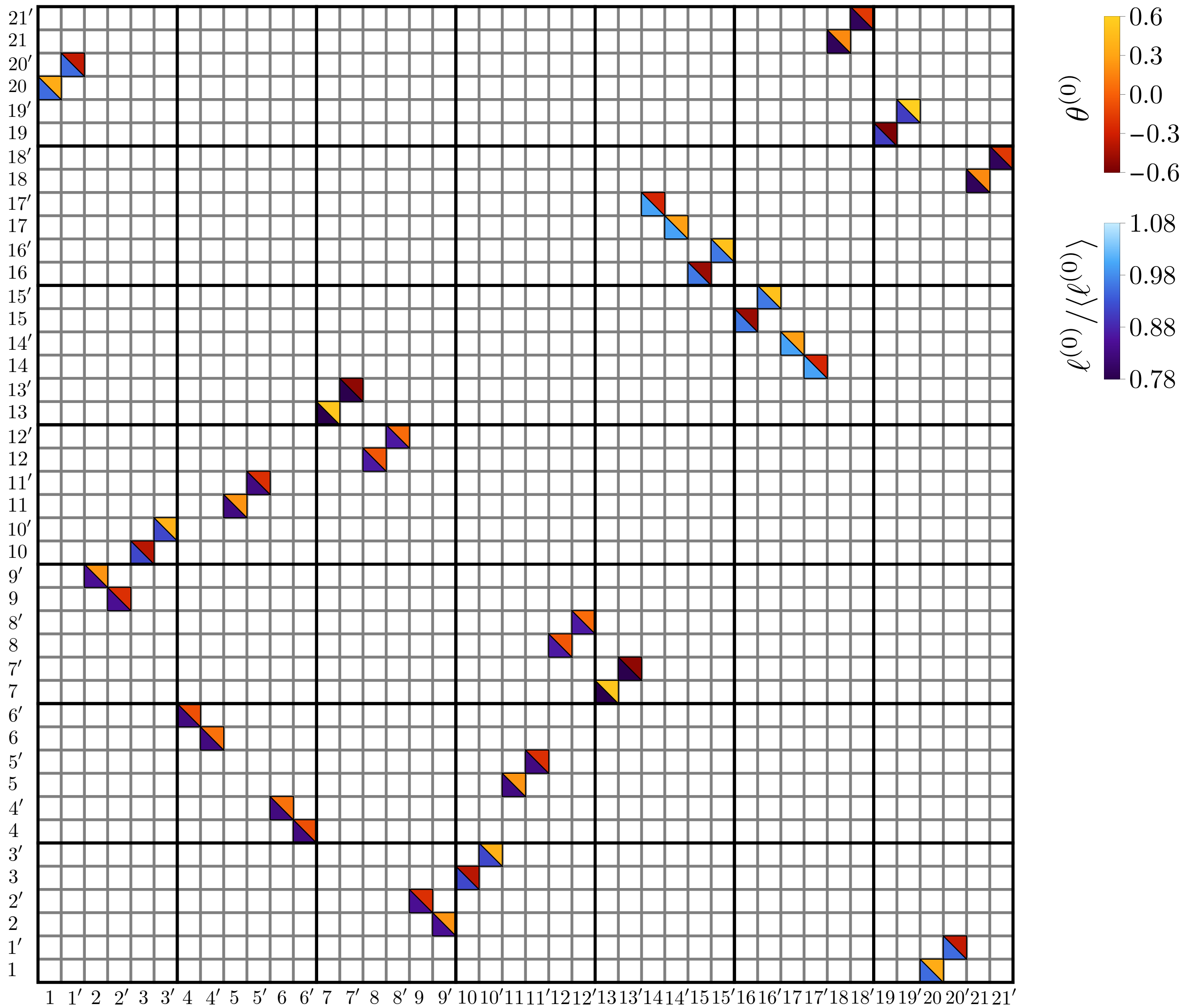


Fig. S6. Interaction matrix for D_7. Each colored block represents a valid edge-pairing in which the lower (upper) half color of each colored-block in the matrix represents the target length (dihedral angle) of a valid edge-pairing.

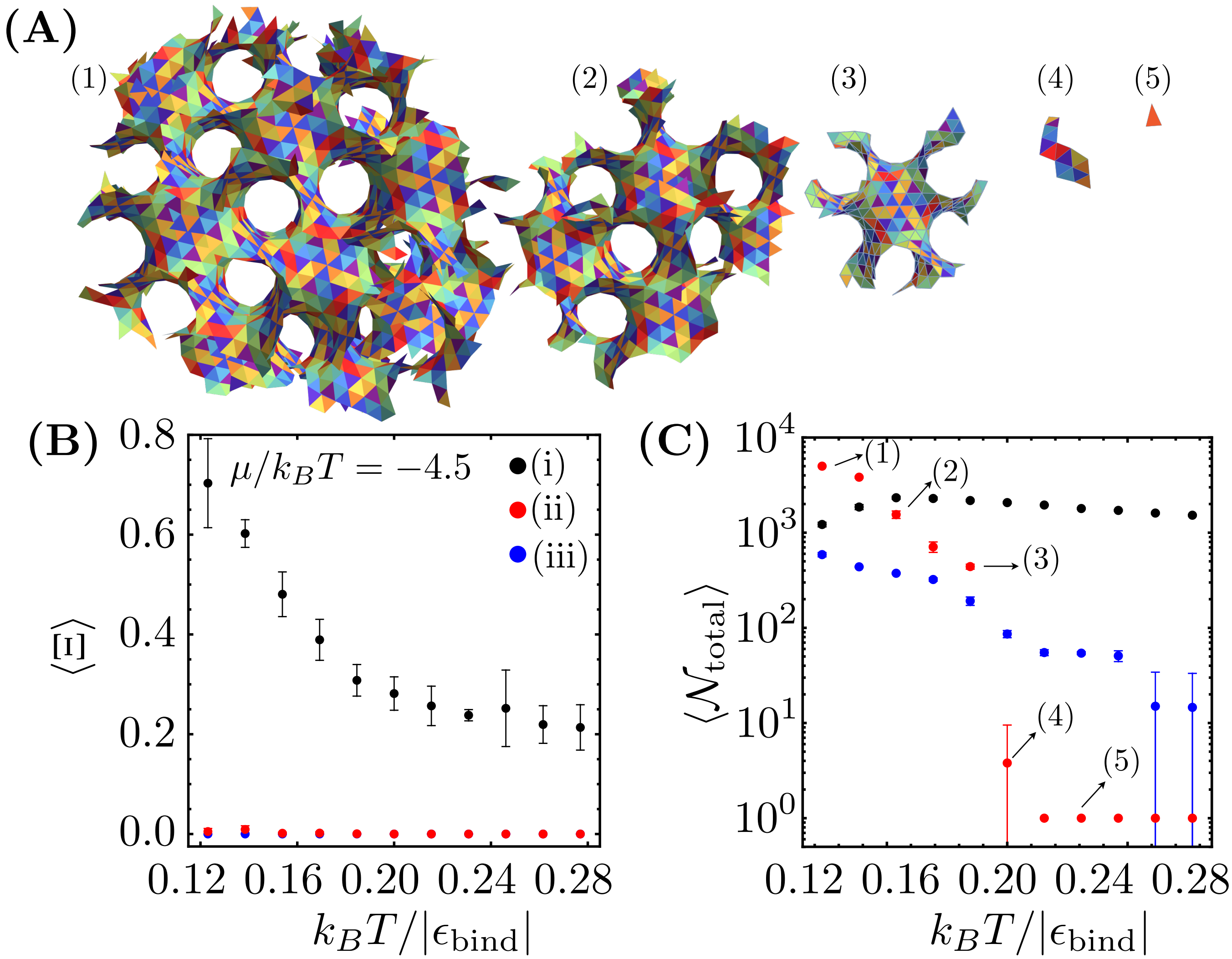


Fig. S7. Results from simulations performed at varied temperatures. **(A)** G_7 Snapshots corresponding to simulations using the conditions of the intermediate flexibility regime (ii) of Fig. 4, corresponding to parameter values indicated in **(C)**. **(B)** Mean quadratic strain, $\langle\Xi\rangle$, as a function of increasing $k_{\rm B}T/|\epsilon_{\rm bind}|$. Black, red, and blue points respectively correspond to the simulation parameters of cases i, ii, and iii of Fig. 4. **(C)** Mean number of assembled triangles, $\langle\mathcal{N}_{\rm total}\rangle$, for the same temperature range and simulation parameters used in **(B)**. Each labeled point for case ii matches the simulation conditions of the snapshots shown in **(A)**. Each data point in **(B)** and **(C)** was calculated with respect to 5 independent realizations. All simulations were performed at fixed bath concentration $\mu/k_{\rm B}T = 4.5$. Simulations were ended either when $\mathcal{N}_{\rm total} = 5000$ or when 50×10^6 MC sweeps were reached. In order to allow for melting, the triangles composing the seed structure are also allowed to disassemble.

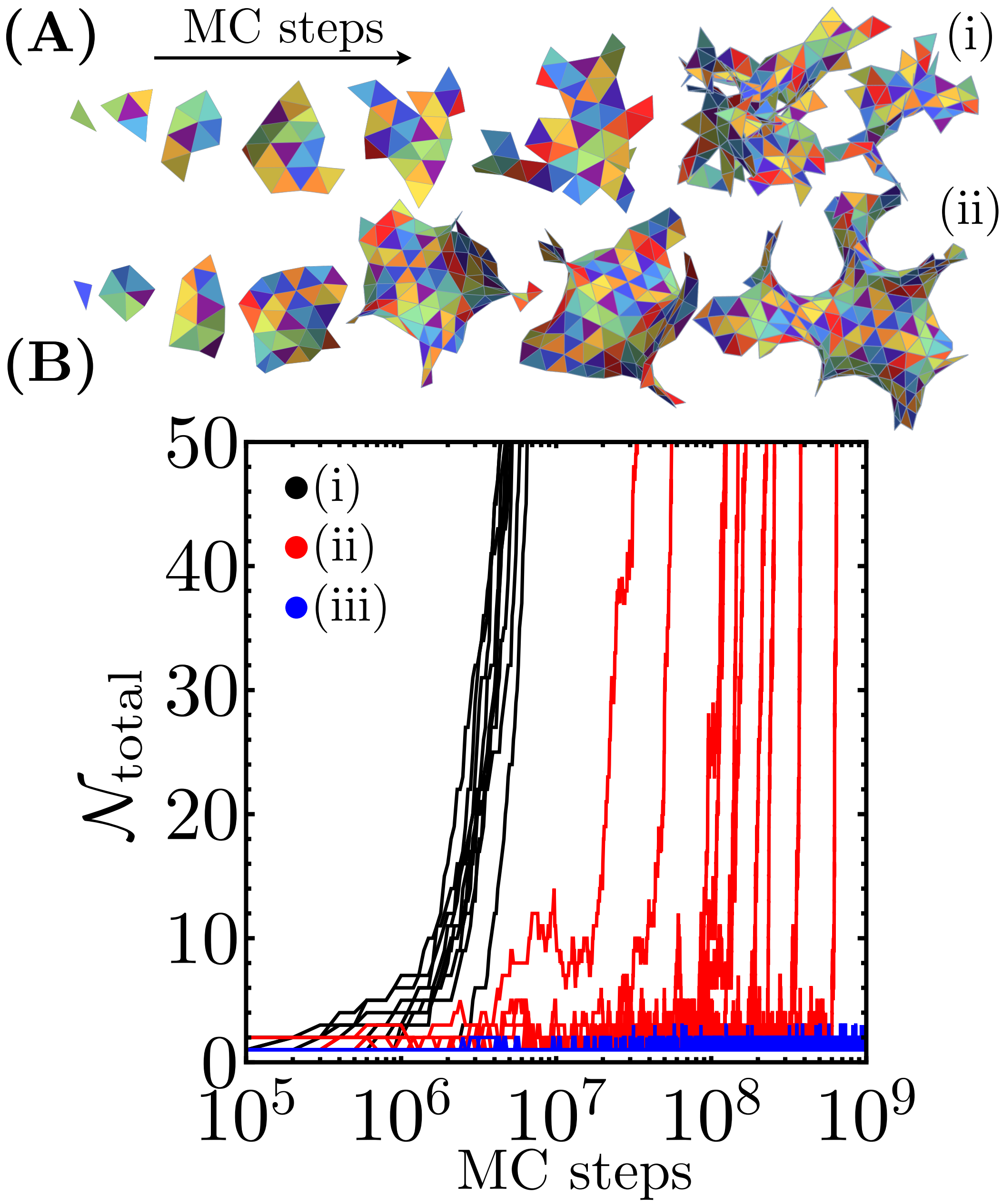


Fig. S8. Results from unseeded simulations to investigate nucleation. (A) Top (bottom): Sequence of snapshots showing a nucleation pathway for a small G_7 structure using the simulation parameters of i (ii) of Fig. 4. In both cases, the size of the rightmost structures is 336 triangles (i.e. the seed size of G_7). **(B)** Number of assembled subunits, $\mathcal{N}_{\text{total}}$, as a function of the number of MC steps for unseeded structures using the simulation parameters of cases i (black), ii (red), and iii (blue) of Fig. 4. We show time traces of 10 independent realizations for each of the simulated regimes. Simulations were ended either when $\mathcal{N}_{\text{total}} = 336$ or when 10^9 MC sweeps were reached.

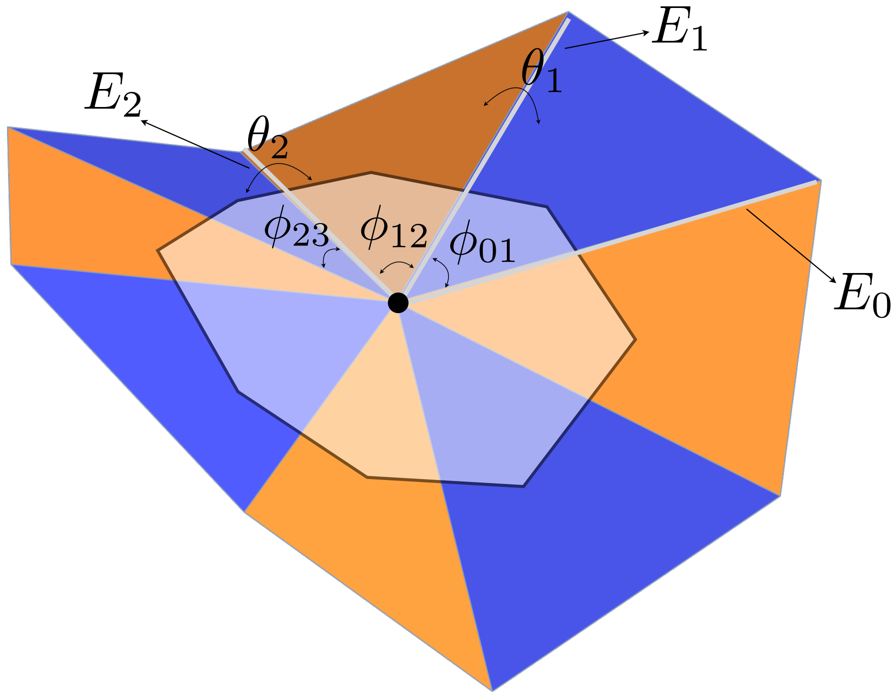


Fig. S9. Illustration of the topological charge calculation procedure: given a closed path of subunits around a vertex we can move across the anticlockwise-ordered edges, E_i, meeting at the vertex in order to determine the charge. Moving between two edges E_1 and E_2 an angular contribution ϕ_{01} is picked up. As we move to the neighboring subunit and angle contribution corresponding to the dihedral angle is picked up as well. The process is continued until we are back at the starting subunit. Using the collection of ordered angles $\phi_{01},\ \theta_1,\ \phi_{12},\ \theta_2, \ldots$ allows us to build the rotation composition of Eq. 2. Notice that in order to find the "true" effective rotational angle Φ we need to use the ϕ and θ angles of the "target" geometry instead of the "current" geometry.

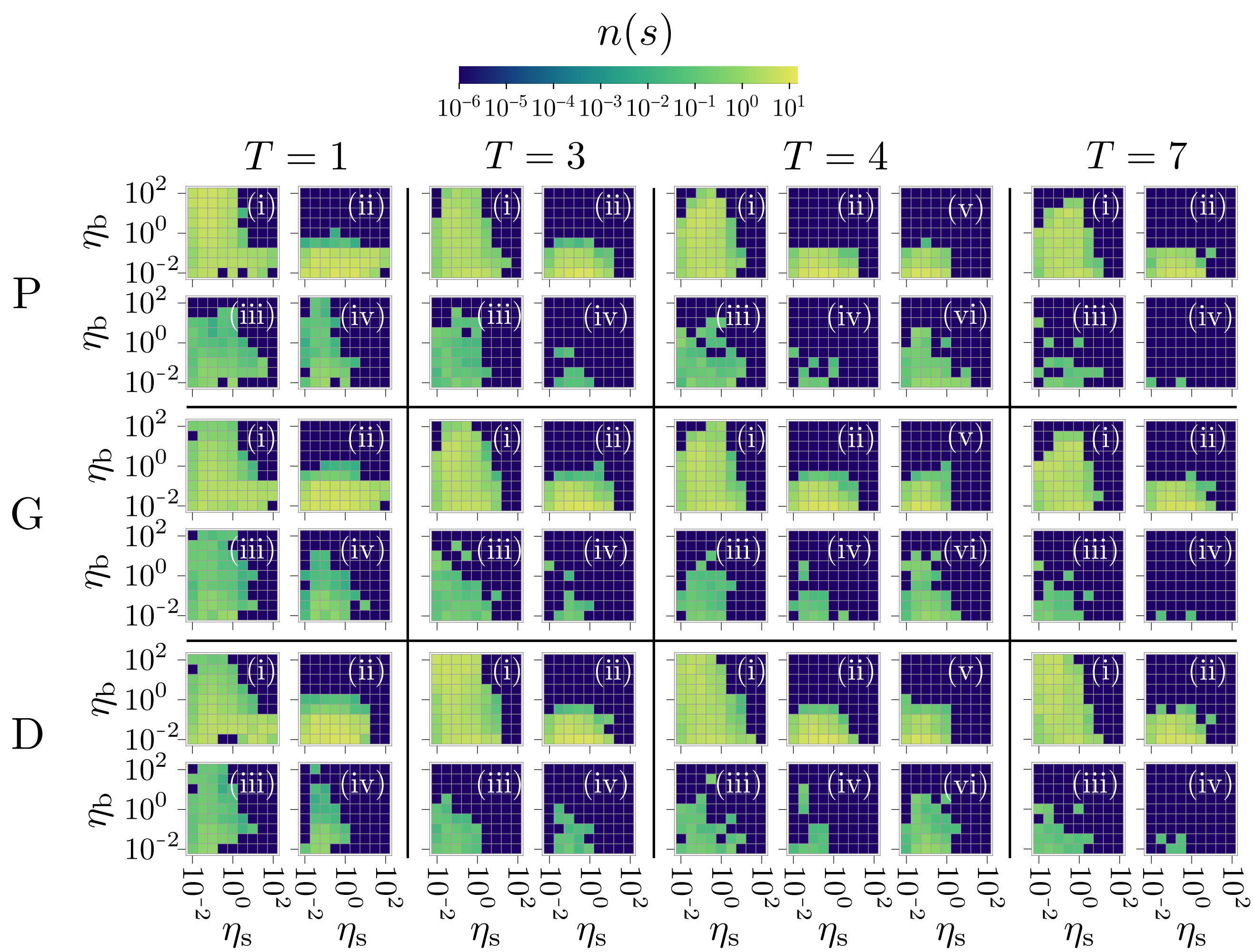


Fig. S10. Defect population per unit cell, $n(s)$, as a function of dimensionless ratios η_s and η_b for each of the defect types. Panels (i) to (vi) respectively correspond to charges $+2\pi/3$, $+\pi$, $-2\pi/3$, $-\pi$, $+\pi'$, and $-\pi'$.

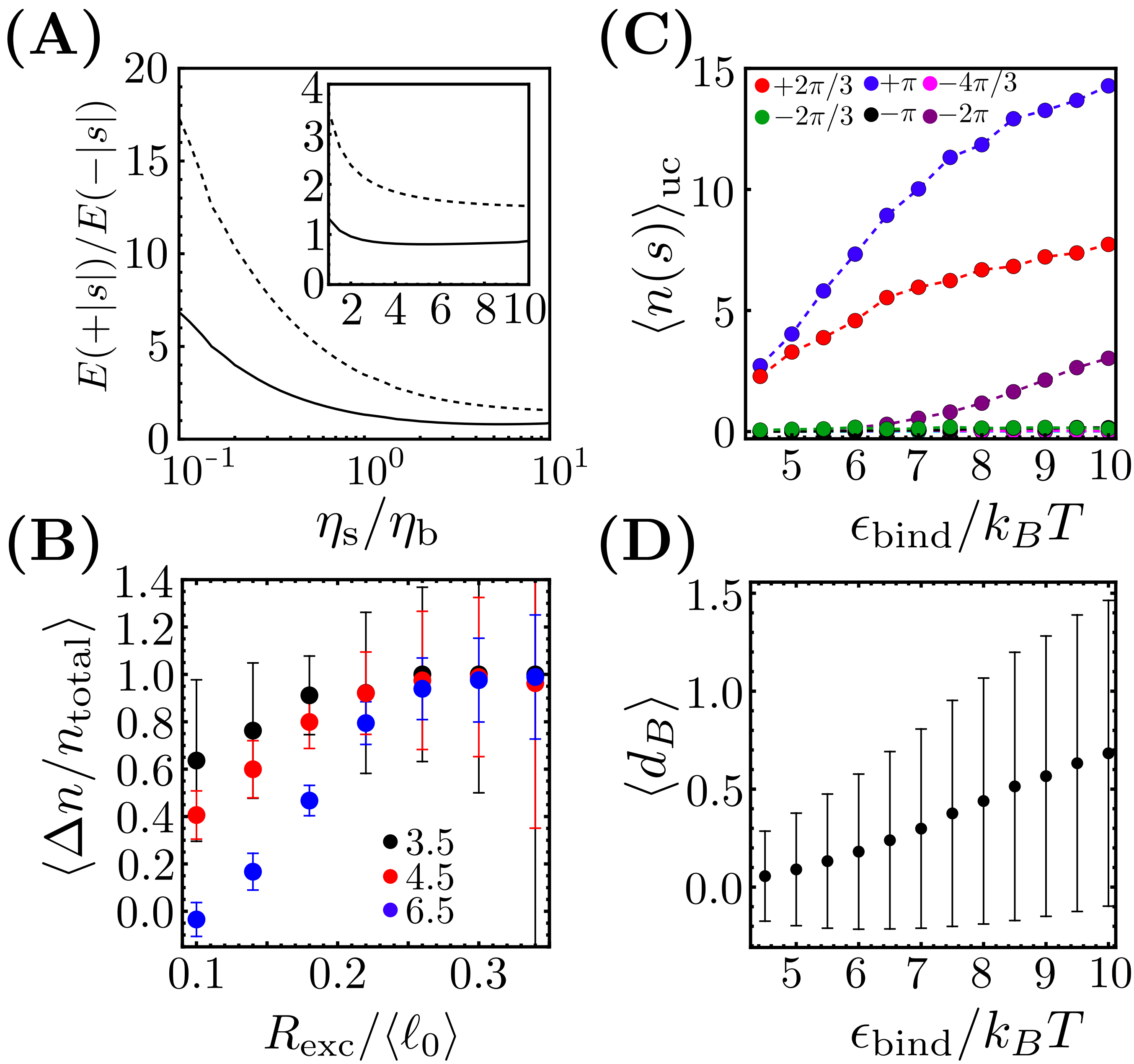


Fig. S11. **(A)** Elastic energy ratio $E(+|s|)/E(-|s|)$ for $|s| = 2\pi/3$ (solid line) and $|s| = \pi$ (dashed line) as a function of η_s/η_b. **(B)** Mean relative counts difference between positive and negative charges, $\langle \Delta n/n_{total} \rangle$, for increasing excluder radius size. The black, red, and blue points correspond to fixed binding affinity values $E_{bind} = 3.5$, 4.5, and 6.5 respectively. **(C)** Mean defect type population per unit cell, $\langle n(s)_{uc} \rangle$, for a G_7 structure for increasing binding affinity, ϵ_{bind}. The elastic moduli are fixed for all points and calculated using $\eta_s = 1$ and $\eta_s = 0.01$ with a reference value $E_{bind} = 6.5 k_B T$. The chemical potential, $\mu = -4.5 k_B T$, is also fixed for all points. Each point is computed as the mean of 50 independent realizations. **(D)** Mean graph distance, $\langle d_B \rangle$, for increasing binding affinity and same simulations parameters of **(C)**.

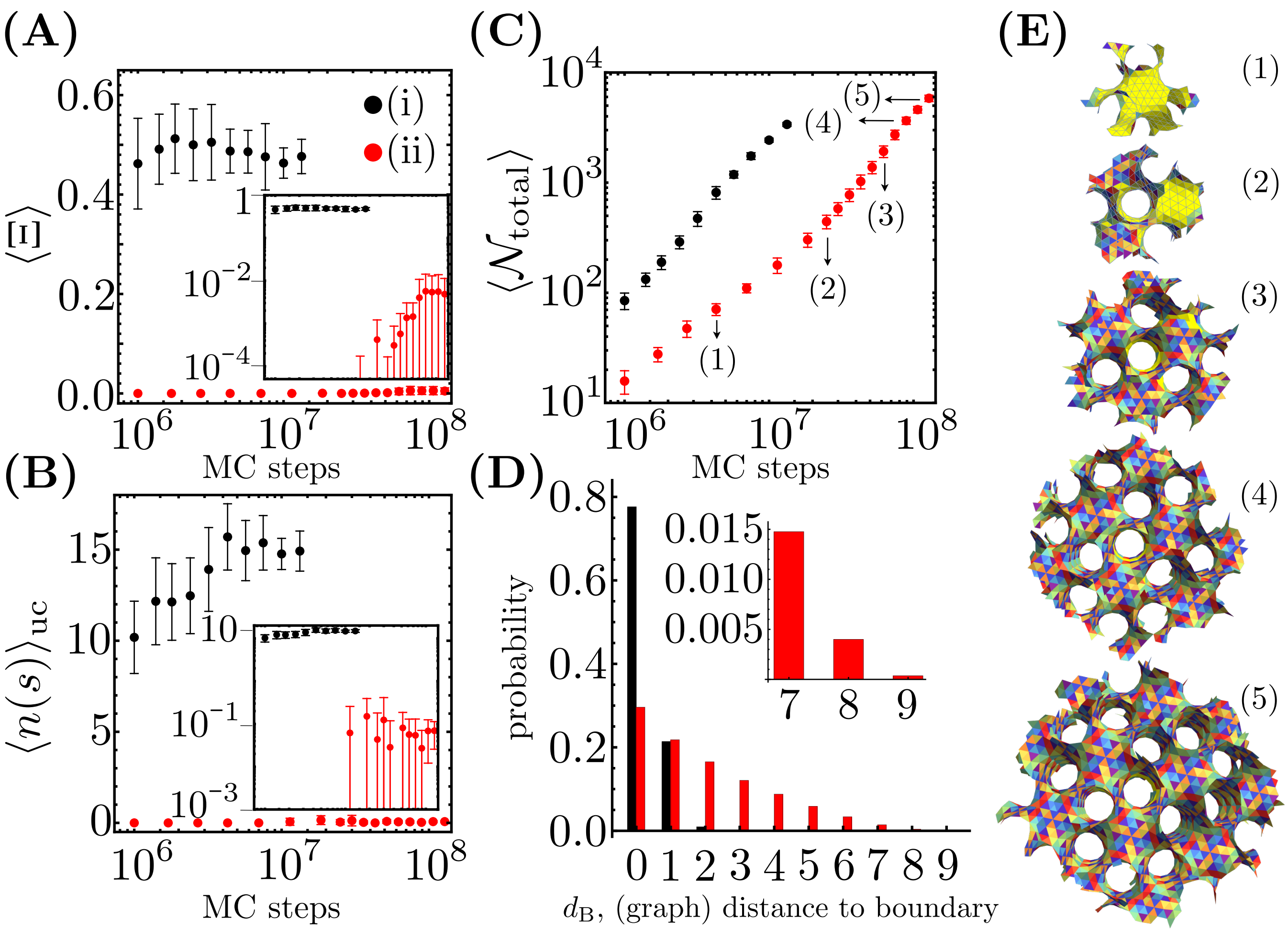


Fig. S12. Long simulations demonstrating the growth of structures that are much larger than the subunit scale. (A-B) Mean quadratic strain, $\langle \Xi \rangle$, and mean total number of defects (per unit cell) , $\langle n(s) \rangle_{\mathrm{uc}}$, as a function of the number MC steps. Simulations are shown for the simulation conditions of i (black) and ii (red) of Fig. 4. The insets of both **(A)** and **B)** show log-log plots of the same quantities. **(C)** Mean number of assembled subunits, $\langle \mathcal{N}_{\mathrm{total}} \rangle$, for the cases i and ii of Fig. 4 as a function of MC steps. The largest $\langle \mathcal{N}_{\mathrm{total}} \rangle$ values for case ii case show that the intermediate flexibility structures grown to sizes up to ~ 6000 subunits (roughly 18 times bigger than the starting seed size). **(D)** Probability of a vertex to have a graph distance, d_{B}, to the boundary with respect to the cases i and ii. The probabilities were calculated using the terminal (largest) structures for both i and ii cases. **(E)** Sequence of snapshots showing a sample pathway for a single case of intermediate flexibility ii. The different assembly stages correspond to the labeled points of **(C)**. We used 10 independent realizations for each of the simulated regimes. The simulations for regime i (ii) were performed to $\gtrsim 10^7$ ($\gtrsim 10^8$) MC sweeps.

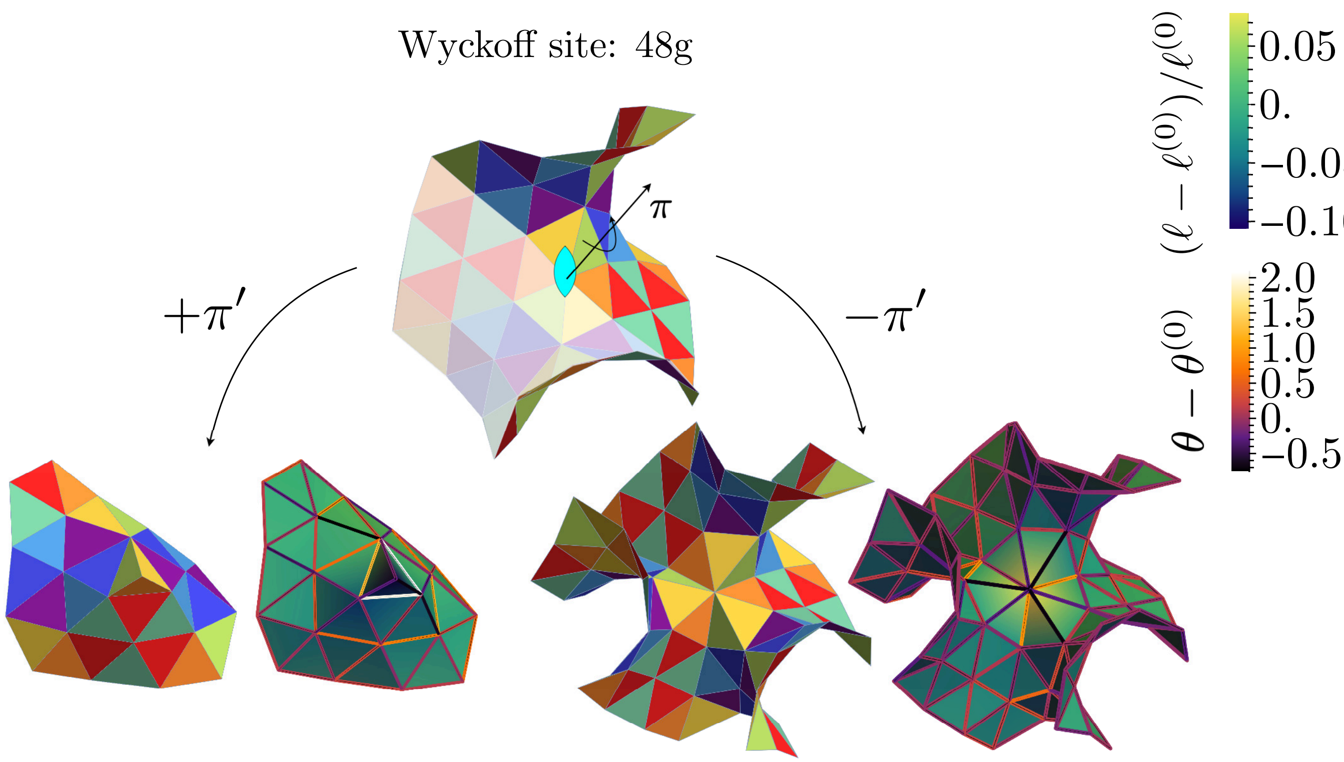


Fig. S13. For a subset of T values additional Wyckoff sites with a 2-fold symmetry can be found. The triangulation on top shows a defectless patch exhibiting a Wyckoff site 48g present in G_4 structures. Defects of charge $s = \pm\pi$ are generated through a Volterra construction on the defectless patch. The charges are primed to distinguish them from the defects obtained through the Wyckoff sites 24d. Each defective patch is colored with respect to the unique triangle ID label as shown on the left patches as well as their length strain (faces) and angle strain (edges) as shown on the right patches.

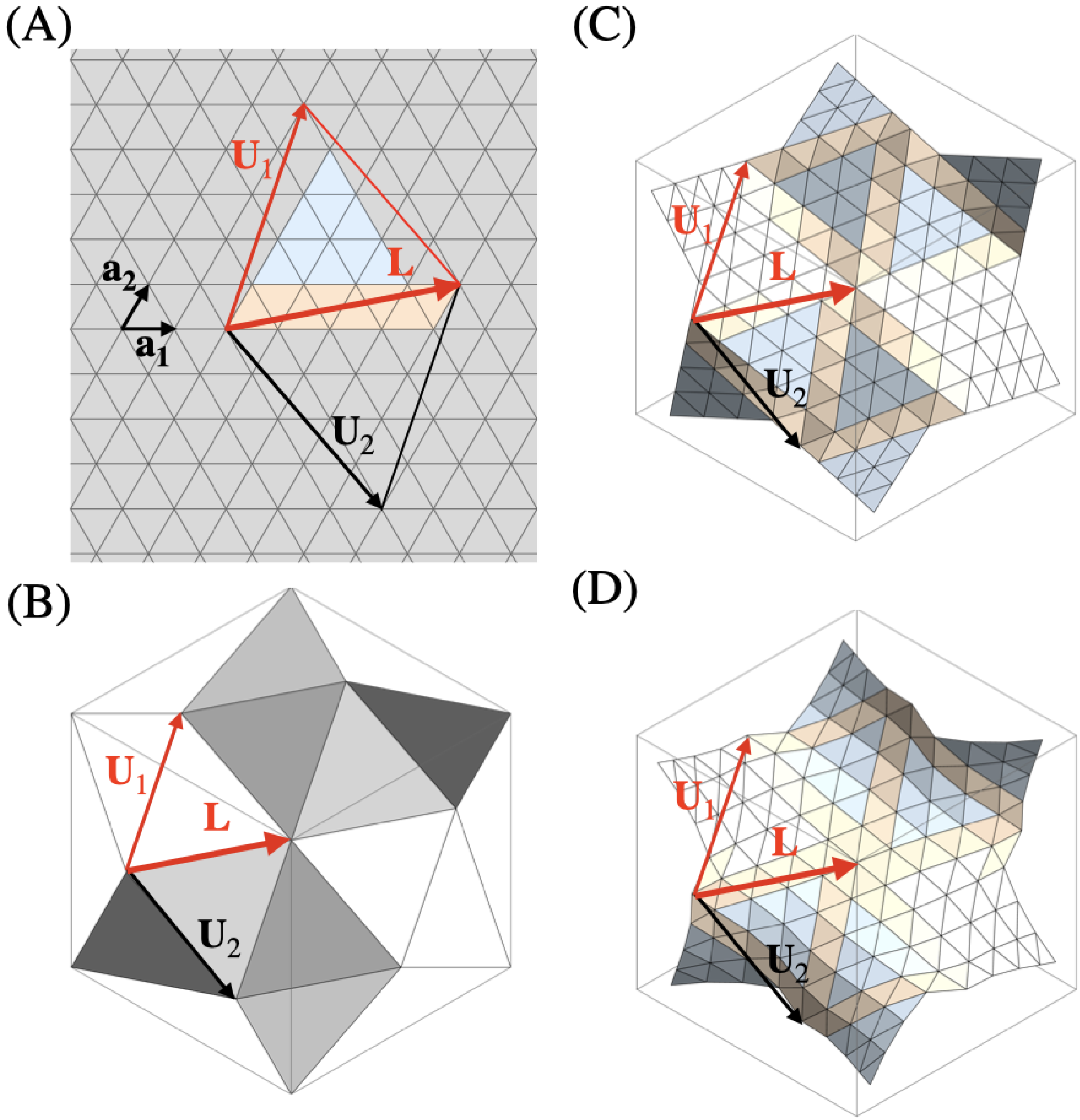


Fig. S14. Construction of triangulation for given T from $T = 1$ triangulation. **(A)** A planar construction defines the vectors $\mathbf{L}, \mathbf{U}_1, \mathbf{U}_2$ in terms of the triangular lattice vectors $\mathbf{a}_1, \mathbf{a}_2$. The lattice triangles contained in the larger triangle with edges $\mathbf{L}, \mathbf{U}_1$ (blue) will be mapped to each face and the lattice triangles that contain the line segment $\mathbf{L}$ (orange) will be mapped to each edge of the base ($T = 1$) triangulation. Here, the construction is illustrated for $(h, k) = (4, 1)$. **(B)** The base triangulation is a set of triangles so that each triangle has three neighbors and a consistent orientation can be assigned to the triangulation. Here, the unit cell of the $T = 1$ triangulation associated with the gyroid surface is shown, deriving from the surface's asymmetric units. **(C)** The triangles of the planar construction are mapped to the base triangulation by identifying each edge with the triangulation vector $\mathbf{L}$ and associated neighboring edges $\mathbf{U}_1, \mathbf{U}_2$. **(D)** Finally, the vertices are translated closer to the surface while maintaining the targeted symmetries of the structures.

5. The model

In this section we provide additional details about the model and Monte Carlo simulation that we used to generate the results in the main text. The model and algorithm have been presented in ref. (3–5). In particular, we consider flexible triangular subunits which can bind to each other along edges with a set of preferred dihedral angles that set the preferred curvatures of the assembling sheet. Monte Carlo simulations are performed in the grand canonical ensemble at fixed μVT, with μ the chemical potential of subunits in the bath. Each Monte Carlo simulation involves a single cluster undergoing assembly and disassembly, with subunits taken from or returned to the bath respectively, as well as structural relaxation moves. We describe specific differences with respect to the RG model below.

A. Energies. Each three edges of the triangular subunits may be of different types, $t(p) = 1, 2, 3$, for edge index p and whether a given type can bind to another type is given by the matching rules.

The total energy of the system is given by

$$E = \sum_{p}^{3n_{\mathrm{s}}} E^{p}_{\mathrm{stretch}} + \frac{1}{2}\sum_{\langle pq\rangle}\left(E^{pq}_{\mathrm{bend}} + E^{pq}_{\mathrm{bind}}\right) \tag{1}$$

where the first sum goes over all edges, with n_{s} the number of subunits in the cluster. The second sum only goes over bound edges (i.e. non-boundary, adjacent edges, so there are $2n_{\mathrm{b}}$ terms in the sum, with n_{b} as the number of bonds). The $1/2$ factor corrects for double counting.

The stretching energy is defined as:

$$E^{p}_{\mathrm{stretch}} = \epsilon_{\mathrm{s}}\frac{(l^{p} - l_0)^2}{2} \tag{2}$$

where ϵ_{s} is the stretching modulus, l^{p} is the instantaneous length, and l_0 is the stress-free (rest) length of an edge. For the tubule model we set the stretching modulus and rest length equal for all edges.

The bending energy is quadratic in deviations from the preferred dihedral angle:

$$E^{pq}_{\mathrm{bend}} = \kappa_{\mathrm{b}}\frac{\left(\theta^{pq} - \theta_0^{t(p)t(q)}\right)^2}{2} \tag{3}$$

with p and q adjacent edges and $t(p), t(q)$ the edge types. κ_{b} is the bending modulus and is set equal for all edge types. $\theta_0^{t(p)t(q)}$ is the preferred dihedral angle between edges with types $t(p)$ and $t(q)$.

Binding energies corresponding to all matching edge pairs are set equal, to $E^{pq}_{\mathrm{bind}} \equiv \epsilon_{\mathrm{b}}$.

In addition to the above terms, each subunit has at its center of mass a spherical excluder of radius $0.2l_0$ to prevent subunit overlaps. Finally, to prevent extreme distortions of subunits, maximum edge length fluctuations are limited to $l_0/2 < l < 3l_0/2$.

B. Coarse-graining. Our model is motivated by the triangular DNA origami subunits developed in Sigl et al.(6), in which subunits bind through lock-and-key 'patches' along subunit edges in which attractive interactions are generated through blunt-end stacking of unsatisfied nucleotides. Therefore, in our model we define attractive bonds along subunit edges (rather than at vertices as in the RG model). In particular, attractive bonds occur at each shared pair of subunit edges with the same type. Because the interactions in the experimental system are driven by nucleotide stacking, they are extremely short-ranged in comparison to the subunit size (the subunit edge lengths are approximately 60 nm). Therefore, in our simulations we avoid resolving the short length scale fluctuations in separation distance between bound edges and their associated vertices by coarse-graining as follows.

A microstate i is defined as the position of all the $3n_{\mathrm{s}}$ vertices of n_{s} subunits: $i \to (\vec{x}_1, \vec{x}_2, ...\vec{x}_{3n_{\mathrm{s}}})$ The grand canonical probability *density* of finding the system around state i is

$$f(i) = \frac{P(\vec{x}_1, \vec{x}_1 + d\vec{x}_1; ...; \vec{x}_1, \vec{x}_{n_s} + d\vec{x}_{n_s})}{d\vec{x}_1 d\vec{x}_2 ... d\vec{x}_{3n_{\mathrm{s}}}} = \frac{1}{Z_\Omega}\frac{e^{\beta n_{\mathrm{s}}\mu}}{\lambda^{9n_{\mathrm{s}}}}e^{-\beta E_i} \tag{4}$$

where μ is the chemical potential and λ^3 is the standard state volume. This probability density has the dimensions of $1/\mathrm{volume}^{3n_{\mathrm{s}}}$ corresponding to all the $3n_{\mathrm{s}}$ vertices of the subunits. Due to bonds, however, some pairs of vertices are confined within a *binding volume* v_{a}. We consider a square-well potential so that the binding energy is constant within this volume. Analogous to Ref. (7), we can then coarse-grain to avoid resolving intra-bond fluctuations. We assume that fluctuations of bound edges are sufficiently small that each pair of vertices at either end of a bound edge pair are constrained within a *binding volume* v_{a}. Note that we constrain vertices rather than edges so that the coarse-grained microstate can be represented in terms of positions of vertices rather than edges, which is easier to implement computationally. In the coarse-grained system, a coarse microstate is specified by the coordinates corresponding to the independent vertex degrees of freedom (with 1 degree of freedom for each bound vertex group and unbound vertex): $\Gamma \to (\vec{x}_1, \vec{x}_2, ...\vec{x}_{n_{\mathrm{v}}})$, where n_{v} is the number of independent bound vertex groups and free vertices. The probability of such a coarse-grained state is given by the net weight of all the corresponding fine-grained microstates:

$$\rho(\Gamma) = \int_{\{v_a\}} f(i) \mathrm{d}^{n_{\mathrm{VB}}} \vec{x} \quad [5]$$

where n_{VB} is the number of vertex-bonds and is given by $n_{\mathrm{VB}} = 3n_{\mathrm{s}} - n_{\mathrm{v}}$. For simplicity, we take the limit in which $\sqrt[3]{v_{\mathrm{a}}}$ is small in comparison to the length scale over which the elastic energy varies, so that the energy is constant within the bound volume v_{a}. Then $f(i)$ is a constant, and the probability density is given by

$$\rho(\Gamma) = \frac{1}{Z_\Omega} v_{\mathrm{a}}^{n_{\mathrm{VB}}} \frac{e^{\beta n_{\mathrm{s}} \mu}}{\lambda^{9 n_{\mathrm{s}}}} e^{-\beta E_\Gamma} \quad [6]$$

where E_Γ is the total energy of state Γ (including stretching, bending and binding energies). The coarse graining process is illustrated in Fig. S15.

C. Implementation and data structure. The simulation is implemented on top of the OpenMesh library (8). Subunits are implemented as triangular mesh elements. OpenMesh uses the halfedge data structure which is suitable to implement triangles with directed normals (Fig S16). The directed halfedges allow for a clockwise iteration through the boundary of a triangle, which makes the two faces of the triangles distinguishable. Only halfedges with opposite orientations can bind together, making it impossible to form a Mobius strip, for example. The data structure and the resulting iterators in OpenMesh allow for an easy and efficient iteration over the neighborhood of mesh elements (vertices, edges and faces). The implementation of mesh element rearrangements is less straightforward, but we implemented it via the insertion and removal of virtual triangles. In addition, OpenMesh allows for the storage of various properties on mesh elements, allowing storage of edge types and face types stored on the elements. To improve readability in the upcoming sections, we will not represent halfedges separately.

6. The Monte Carlo moves

In this section we detail the Monte Carlo moves of the simulation. Our algorithm has 11 moves: vertex displacement, simple subunit insertion/deletion, wedge insertion/deletion, wedge fusion/fission, crack fusion/fission, and edge fusion/fission.

Detailed balance. For the transition between state Γ and Γ' detailed balance corresponds to (7, 9):

$$P(\Gamma) \times \alpha(\Gamma \to \Gamma') \times p_{\mathrm{acc}}(\Gamma \to \Gamma') = P(\Gamma') \times \alpha(\Gamma' \to \Gamma) \times p_{\mathrm{acc}}(\Gamma' \to \Gamma) \quad [7]$$

where $\alpha(\Gamma \to \Gamma')$ is the probability of generating a $\Gamma \to \Gamma'$ move attempt (trial), $p_{\mathrm{acc}}(\Gamma \to \Gamma')$ is the probability of accepting the move, and $P(\Gamma) = \rho(\Gamma) \mathrm{d}^{n_{\mathrm{v}}(\Gamma)} \vec{x}$ is the equilibrium probability of finding a system in a voxel of volume $\mathrm{d}^{n_{\mathrm{v}}(\Gamma)} \vec{x}$.

Next, we use Eq. Eq. (7) to define the acceptance criteria for each MC move. The acceptance criteria are derived in detail for the wedge fusion/fission move; the steps to follow are the same for all other moves.

A. Vertex displacement. In this move, a vertex is randomly selected, a random uniform displacement is drawn, and the vertex is displaced to its new position according to:

$$x \to x + \mathcal{U}(-d_{\max}, d_{\max}) \quad [8]$$
$$y \to y + \mathcal{U}(-d_{\max}, d_{\max}) \quad [9]$$
$$z \to z + \mathcal{U}(-d_{\max}, d_{\max}) \quad [10]$$

with $d_{\max}$ the maximum displacement. The move is accepted with a probability $p_{\mathrm{acc}} = \exp(-\Delta E / k_{\mathrm{B}} T)$ where ΔE is the (bending plus stretching) energy change due to the displacement. The parameter $d_{\max}$ can be adjusted during a burn-in period to optimize convergence to equilibrium. Generally optimal values are on the order of the typical length scale of thermal fluctuations dictated by the elastic energy, leading to acceptance probabilities on the order of 50%. In our simulations typical values are between $d_{\max} = [0.01 l_0, 0.1 l_0]$. The vertex displacement move is illustrated in Fig S17: the number of subunits n_{s}, number of vertices n_{v}, number of vertex bonds n_{VB} and number of bonds n_{b} remains unchanged during this move.

B. Simple insertion / removal.

B.1. Simple insertion. In this move, an edge is randomly selected from the set of all boundary edges, where a new subunit will be attached. The number of such boundary edges is n_{e}. Subunits can be inserted in n_{r} different rotations, where n_{r} is the number of distinct rotational states for a subunit which has one edge aligned with the edge of a neighboring subunit. For our triangular subunits with three distinct edge types, $n_{\mathrm{r}} = 3$. In our algorithm, during insertion of a subunit its rotational state is chosen randomly from the set of three possibilities. If the aligned edge is not complementary to the type of the boundary edge, then the move is rejected. In this work, the two edges must be of the same type to be complementary.

The positions of two of the new subunit's vertices (those at either end of the edge being bound) are set equal to the positions of the corresponding vertices of the boundary edge to which it is binding. The third vertex position is randomly chosen from within a volume v_{add} centered at the equilibrium position of the new vertex.

Thus, the attempt probability for a simple insertion is given by:

$$\alpha(i \to j) = n_{\mathrm{e}} k_{\mathrm{i}} \tau n_{\mathrm{r}} \times \frac{1}{n_{\mathrm{e}} n_{\mathrm{r}} (v_{\mathrm{add}} / \mathrm{d}\vec{x})}. \quad [11]$$

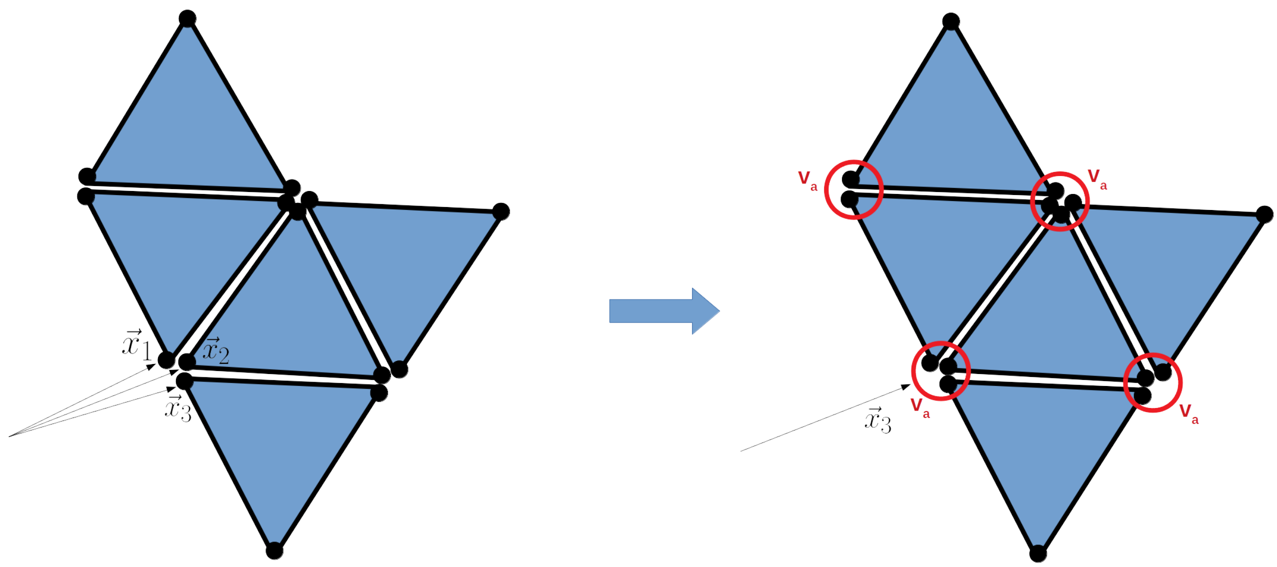


Fig. S15. Coarse-graining of an example cluster configuration. In this configuration, the number of subunits is $n_s = 5$, the number of initial (before coarse-graining) vertices is $3n_s = 15$, and the number of vertices after coarse-graining is $n_v = 7$. The red circles indicate bound vertex groups, and the number of vertex degrees of freedom that have been eliminated by coarse-graining in this configuration is $n_{VB} = 1 + 3 + 2 + 2 = 8 = 3n_s - n_v$. Motivated by DNA origami subunits in Sigl et al. (6), the attractive interactions (i.e. 'bonds') in this model occur along edge-pairs of the same type shared by two subunits. In this configuration there are $n_b = 4$ bonds.

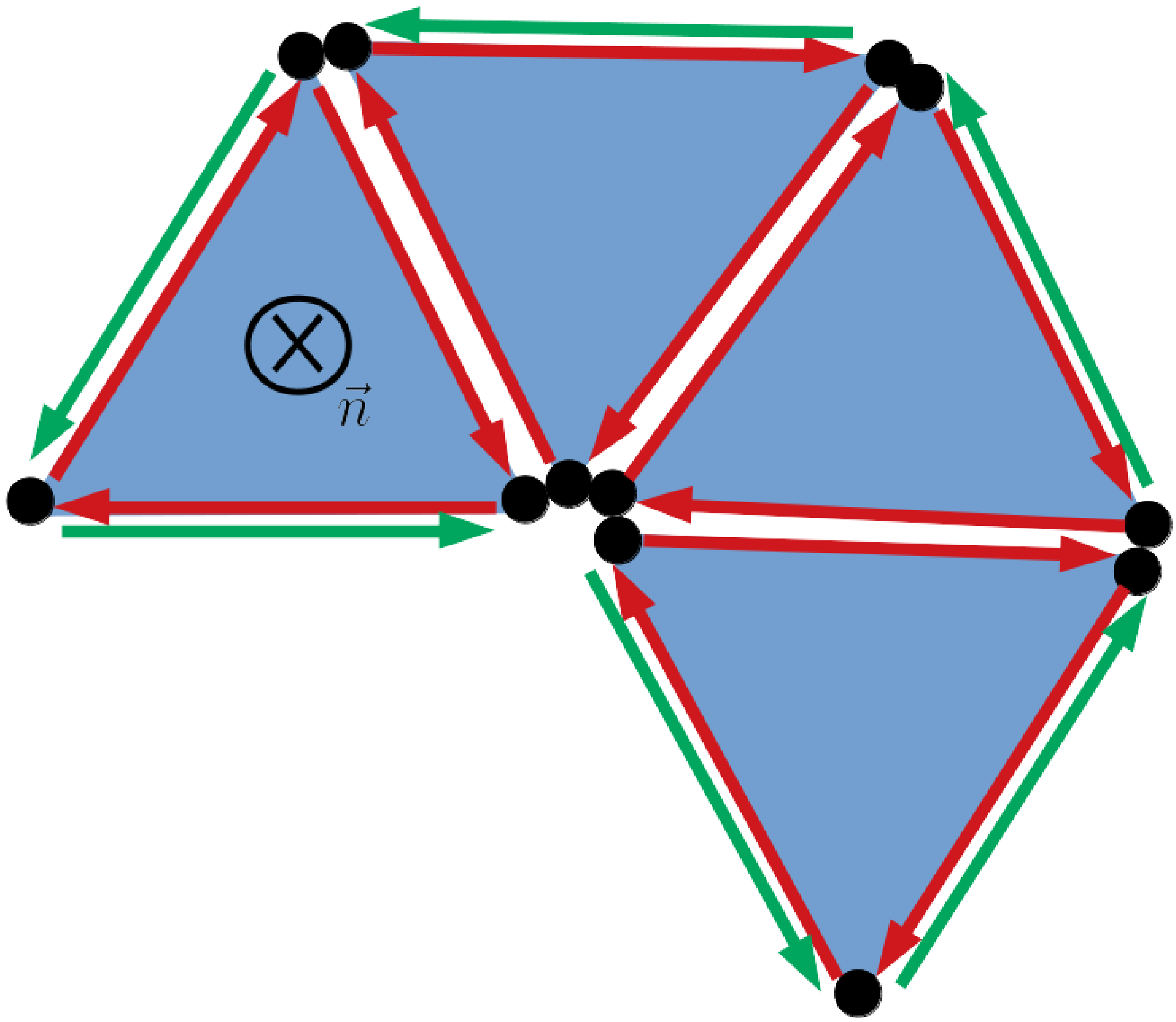
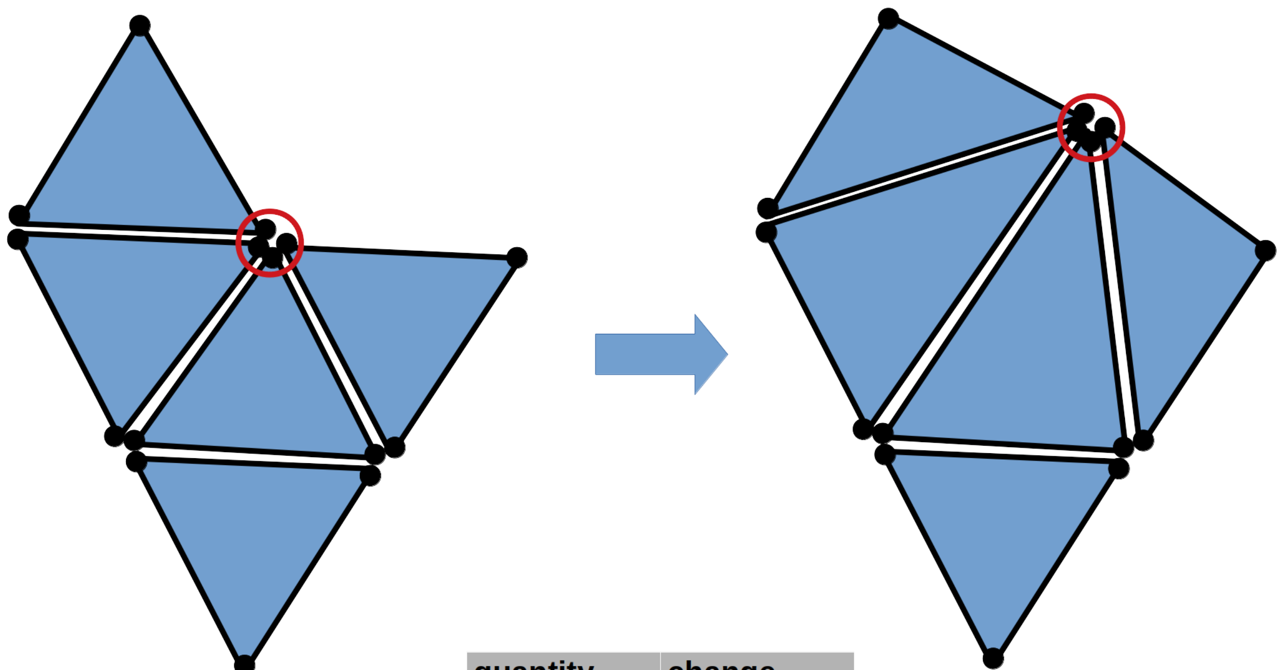


Fig. S16. The halfedge data structure used by OpenMesh. Each edge is represented by two directed edges. Boundary edges are no exception and thus are represented by a non-boundary halfedge and a boundary halfedge (in green). This latter is irrelevant for our model. Directed edges allow for the unambiguous definition of face normals, for efficient iterations of the element's neighborhood as well as boundary iterations.



quantity	change
n_s	0
n_{VB}	0
n_b	0

Fig. S17. Vertex move. A vertex is randomly displaced and the move is accepted according to the usual Metropolis probability.

Then, applying Eq. Eq. (7) and the attempt probability for the reverse move (simple deletion, presented next, Eq. Eq. (13)), the acceptance probabilities for a simple insertion is

$$p_{\mathrm{acc}}(i \to j) = \min\left[1, \frac{v_{\mathrm{a}}^2 v_{\mathrm{add}}}{\lambda^9} \exp[-(\Delta E_{i\to j} - \mu)/k_{\mathrm{B}}T]\right]. \quad [12]$$

$\Delta E_{i\to j}$ is the energy change due to the move and includes the stretching energy of the newly inserted subunit, its bending energy along the shared edge, and the binding energy due to the creation of an extra bond. During this move, one new (edge) bond and two new vertex bonds are created; i.e. $n_{\mathrm{b}} \to n_{\mathrm{b}} + 1$ and $n_{\mathrm{VB}} \to n_{\mathrm{VB}} + 2$. Moreover, the number of vertices in the structure increases by one, $n_{\mathrm{v}} \to n_{\mathrm{v}} + 1$.

B.2. Simple removal. The reverse move to simple insertion is simple removal. Subunits that can be deleted with this move are those with two boundary edges. The number of simply removable subunits is n_{sr}. One of these is selected randomly, so the attempt probability is

$$\alpha(j \to i) = n_{\mathrm{sr}} k_{\mathrm{i}} \tau \times \frac{1}{n_{\mathrm{sr}}} \quad [13]$$

and, using Eq. Eq. (7) and Eq. Eq. (11), the acceptance probability is

$$p_{\mathrm{acc}}(j \to i) = \min\left[1, \frac{\lambda^9}{v_{\mathrm{a}}^2 v_{\mathrm{add}}} \exp[-(\Delta E_{j\to i} + \mu)/k_{\mathrm{B}}T]\right] \quad [14]$$

During this move, the structure loses one (edge) bond and two vertex bonds; $n_{\mathrm{b}} \to n_{\mathrm{b}} - 1$ and $n_{\mathrm{VB}} \to n_{\mathrm{VB}} - 2$. The number of vertices in the structure decreases by one, $n_{\mathrm{v}} \to n_{\mathrm{v}} - 1$.

If there are multiple species with chemical potentials μ_k, detailed balance must be satisfied for each species, individually. Moreover, each species can have different insertion rates k_{i}^k.

To keep $\alpha < 1$, we ensure that the insertion rate k_{i} constrained by

$$n_{\mathrm{e}} k_{\mathrm{i}} \tau n_{\mathrm{r}} < 1 \quad [15]$$

$$n_{\mathrm{sr}} k_{\mathrm{i}} \tau < 1 \quad [16]$$

In equilibrium, one can use adaptive rates, i.e. reduce k_i on the run if the above condition is not satisfied. In that case, sampling is not taken for the ensuing several time steps. Alternatively, the rates may be set to a low enough value from the beginning and only tested on the run to ensure that the $\alpha < 1$ condition is satisfied. This latter technique is appropriate for dynamical runs as it keeps the rates constant throughout the simulation.

Moreover, we must ensure that v_{add} is large enough so that the vertex does not leave the v_{add} volume during structural relaxation moves; otherwise the insertion/deletion moves would not be reversible and the detailed balance would be violated. For a better convergence, one could choose a gaussian distribution $\mathcal{N}(\vec{r})$ for the position of the new vertex instead of a uniform distribution $1/v_{\mathrm{add}}$. In this case, this distribution has to be accounted for in the acceptance probabilities $p_{\mathrm{acc}}(i \to j)$ and $p_{\mathrm{acc}}(j \to i)$.

C. Wedge insertion/removal.

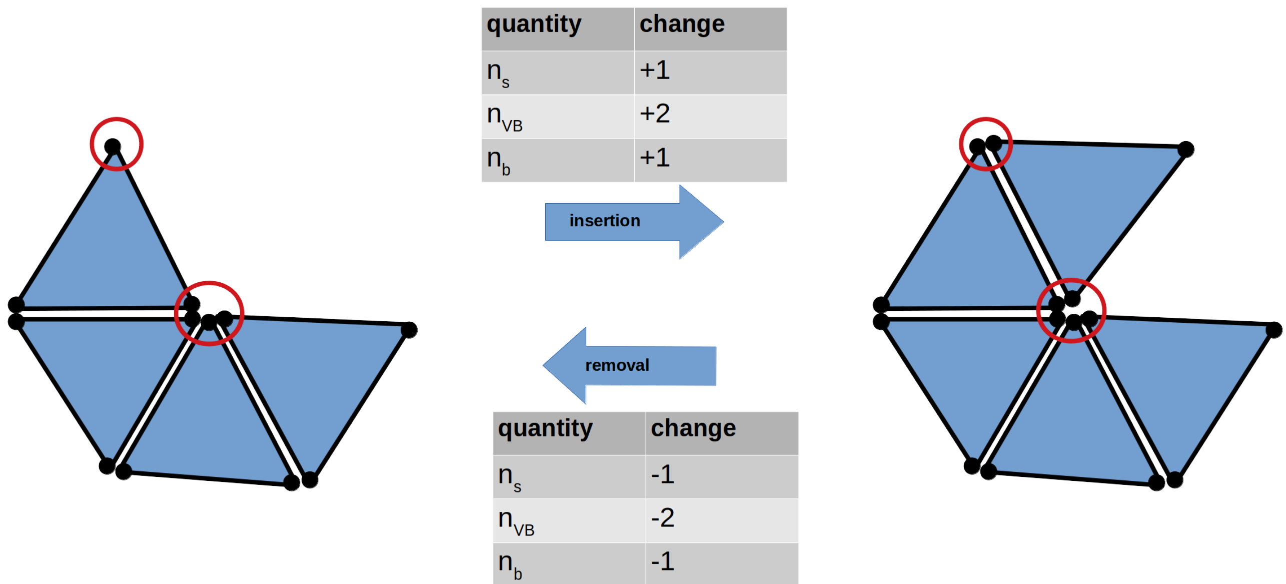
C.1. Wedge insertion. Wedges are positions in the structure where a triangle can be inserted via attaching to two edges (Fig. S19). In a wedge move, we pick randomly from the set of available wedge positions in the structure, and pick a random orientation for the new subunit. Denoting the number of wedge positions in a given structure as n_{w}, the attempt probability for a wedge move is

$$\alpha(i \to j) = n_{\mathrm{w}} k_{\mathrm{i}} \tau n_{\mathrm{r}} \times \frac{1}{n_{\mathrm{r}} n_{\mathrm{w}}} \quad [17]$$

In contrast to the simple insertion move, there is no need for random vertex displacement in a wedge move because all three vertices of the new subunit are fixed by the three vertices of the wedge position. Combining Eq. Eq. (17) and the attempt probability for wedge removal (Eq. Eq. (19)), The acceptance probability for a wedge insertion is

$$p_{\mathrm{acc}}(i \to j) = \min\left[1, \frac{v_{\mathrm{a}}^3}{\lambda^9} \exp[-(\Delta E_{i\to j} - \mu)/k_{\mathrm{B}}T]\right]. \quad [18]$$

During a wedge insertion, two edge bonds and three vertex bonds are created; i.e., $n_{\mathrm{b}} \to n_{\mathrm{b}} + 2$ and $n_{\mathrm{VB}} \to n_{\mathrm{VB}} + 3$, but the number of vertices is unchanged, $n_{\mathrm{v}} \to n_{\mathrm{v}}$. $\Delta E_{i\to j}$ includes the binding energy of the two newly formed bonds, the two bending energies along the two newly bound edges and the stretching energy of the newly inserted subunit.



quantity	change
n_s	+1
n_{VB}	+2
n_b	+1

quantity	change
n_s	-1
n_{VB}	-2
n_b	-1

Fig. S18. Simple insertion and removal.

C.2. Wedge removal. The reverse move of wedge insertion is wedge removal. In a wedge removal, we randomly choose one of the removable wedges from a given structure. With the number of removable wedges as n_{wr}, the attempt probability is

$$\alpha(j \to i) = n_{\mathrm{wr}} k_{\mathrm{i}} \tau \times \frac{1}{n_{\mathrm{wr}}}. \qquad [19]$$

Using Eq. Eq. (17), the acceptance probability for a wedge removal is then

$$p_{\mathrm{acc}}(j \to i) = \min\left[1, \frac{\lambda^9}{v_{\mathrm{a}}^3} \exp[-(\Delta E_{j\to i} + \mu)/k_{\mathrm{B}}T]\right]. \qquad [20]$$

We have the following constraints on rates k_{i} for wedge insertion/removal:

$$n_{\mathrm{w}} k_{\mathrm{i}} \tau n_{\mathrm{r}} < 1 \qquad [21]$$

$$n_{\mathrm{wr}} k_{\mathrm{i}} \tau < 1 \qquad [22]$$

As for simple insertion and removal, in the case of multiple species, detailed balance is satisfied for each species separately for wedge insertion/removal.

D. Wedge fusion / fission.

D.1. Wedge fusion. In this move, a *fusable wedge* is closed, without inserting a new subunit (Fig S20). That is, the two vertices on either side of the wedge opening are merged into a single vertex. Fusable wedges are vertex pairs that i) form a wedge (as in the case of wedge insertion) and ii) are within a separation distance of l_{fuse}.

Denoting the number of fusible wedge positions as n_{w}, in each MC step, a wedge fusion is attempted with probability $n_{\mathrm{w}} k_{\mathrm{wf}} \tau$, where k_{wf} is an adjustable parameter controlling the relative probability of attempting wedge fusion. Then, a wedge position is selected randomly from the set of all n_{w} fusible wedges. The attempt probability is thus

$$\alpha(i \to j) = n_{\mathrm{w}} k_{\mathrm{wf}} \tau \times \frac{1}{n_{\mathrm{w}}}. \qquad [23]$$

Using Eqs. Eq. (23) and Eq. (25), the acceptance probability for fusion moves is

$$p_{\mathrm{acc}}(i \to j) = \min\left[1, \frac{v_{\mathrm{a}}}{v_{\mathrm{fuse}}} \exp(-\Delta E_{i\to j}/k_{\mathrm{B}}T)\right] \qquad [24]$$

where $v_{\mathrm{fuse}} = (4\pi/3)(l_{\mathrm{fuse}}/2)^3$ is the volume of a sphere with diameter l_{fuse}, and $\Delta E_{i\to j}$ is the energy change due to the fusion, including changes in bending, stretching, and binding energies. A fusion move increases the number of edge bonds and vertex bonds by one, $n_{\mathrm{b}} \to n_{\mathrm{b}} + 1$ and $n_{\mathrm{VB}} \to n_{\mathrm{VB}} + 1$; the factor of v_{a} appears in Eq. Eq. (24) to account for the latter.

D.2. Wedge fission. Wedge fission, in which a wedge is opened, is the reverse of the wedge fusion move. Fissionable edges are those edges that can be split along their boundary vertex to obtain a wedge. Denoting the number of such edges as n_{f}, the probability of attempting a wedge fission move during an MC step is $n_{\mathrm{f}} k_{\mathrm{wf}} \tau$. If a fission move is attempted, then an edge is selected randomly from the n_{f} fissionable edges. The position of one of the new vertices is selected randomly within the sphere of volume v_{fuse} centered at the original position of the merged vertices, and the other new vertex is placed in the opposite direction from the original position, at the same distance. Thus, the attempt generation probability is

$$\alpha(j \to i) = n_{\mathrm{f}} k_{\mathrm{wf}} \tau \times \frac{1}{n_{\mathrm{f}}(v_{\mathrm{fuse}}/\mathrm{d}\vec{x})} \qquad [25]$$

and the acceptance probability is

$$p_{\mathrm{acc}}(j \to i) = \min\left[1, \frac{v_{\mathrm{fuse}}}{v_{\mathrm{a}}} \exp(-\Delta E_{j\to i}/k_{\mathrm{B}}T)\right] \qquad [26]$$

We verify that detailed balance holds between wedge fusion and fission as follows. There are two cases to consider:

1. $(v_{\mathrm{fuse}}/v_{\mathrm{a}}) \exp(-\Delta E_{j\to i}/k_{\mathrm{B}}T) < 1 \Leftrightarrow (v_{\mathrm{a}}/v_{\mathrm{fuse}}) \exp(-\Delta E_{i\to j}/k_{\mathrm{B}}T) > 1$

   In this case, $p_{\mathrm{acc}}(i \to j) = 1$ and $p_{\mathrm{acc}}(j \to i) = (v_{\mathrm{fuse}}/v_{\mathrm{a}}) \exp(-\Delta E_{j\to i}/k_{\mathrm{B}}T)$. Then

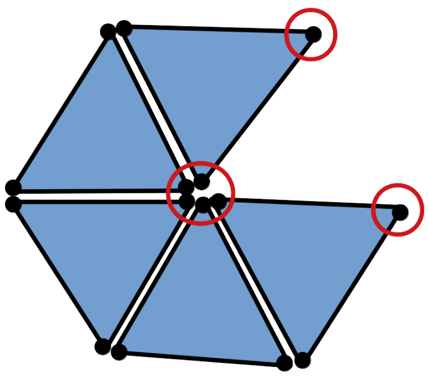
$$P_i \times \alpha(i \to j) \times p_{\mathrm{acc}}(i \to j) = \frac{1}{Z_\Omega} v_{\mathrm{a}}^{n_{\mathrm{VB},i}} \exp[-(E_i - \mu n_{s,i})/k_{\mathrm{B}}T] \frac{1}{\lambda^{9n_{s,i}}} \times \mathrm{d}^{n_{\mathrm{v,i}}}\vec{x} \times k_{\mathrm{wf}}\tau \qquad [27]$$

   and

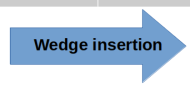
$$P_j \times \alpha(j \to i) \times p_{\mathrm{acc}}(j \to i) = \frac{1}{Z_\Omega} v_{\mathrm{a}}^{n_{\mathrm{VB},j}} \exp[-(E_j - \mu n_{\mathrm{s},j})/k_{\mathrm{B}}T] \frac{1}{\lambda^{9n_{s,j}}} \times \mathrm{d}^{n_{\mathrm{v,j}}}\vec{x} \qquad [28]$$

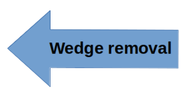
$$\times \quad k_{\mathrm{wf}}\tau \mathrm{d}\vec{x}/v_{\mathrm{fuse}} \times (v_{\mathrm{fuse}}/v_{\mathrm{a}}) \exp(-\Delta E_{j\to i}/k_{\mathrm{B}}T) \qquad [29]$$

   Using: $\Delta E_{j\to i} = E_i - E_j$, $n_{s,i} = n_{s,j}$ (because the move leaves the subunit number unchanged), $n_{\mathrm{V}B,i} = n_{\mathrm{VB},j} - 1$ (one vertex bond is broken upon fission) and $n_{v,i} = n_{v,j} + 1$ (an extra vertex is being born upon fission), we see that the two are equal and detailed balance holds.



quantity	change
n_s	+1
n_{VB}	+3
n_b	+2





quantity	change
n_s	-1
n_{VB}	-3
n_b	-2

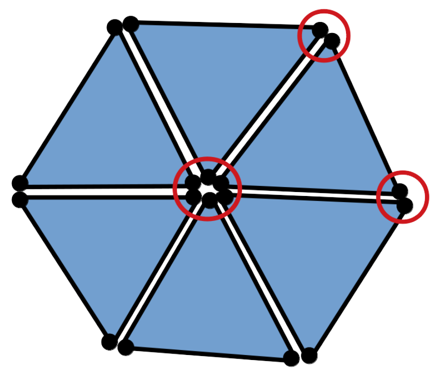


Fig. S19. Wedge insertion and removal.

2. $(v_{\rm fuse}/v_{\rm a})\exp(-\Delta E_{j\to i}/k_{\rm B}T) > 1 \Leftrightarrow (v_{\rm a}/v_{\rm fuse})\exp(-\Delta E_{i\to j}/k_{\rm B}T) < 1$

In this case, $p_{\rm acc}(i\to j) = (v_{\rm a}/v_{\rm fuse})\exp(-\Delta E_{i\to j}/k_{\rm B}T)$ and $p_{\rm acc}(j\to i) = 1$. Then

$$P_i \times \alpha(i\to j)\times p_{\rm acc}(i\to j) = \frac{1}{Z_\Omega} v_a^{n_{{\rm VB},i}} \exp[-(E_i - \mu n_{s,i})/k_{\rm B}T]\frac{1}{\lambda^{9n_{s,i}}}\times {\rm d}^{\rm n_{v,i}}\vec{x}\times k_{\rm wf}\tau \quad [30]$$

$$\times \quad (v_{\rm a}/v_{\rm fuse})\exp(-\Delta E_{i\to j}/k_{\rm B}T) \quad [31]$$

and

$$P_j \times \alpha(j\to i)\times p_{\rm acc}(j\to i) = \frac{1}{Z_\Omega} v_a^{n_{{\rm VB},j}} \exp[-(E_j - \mu n_{s,j})/k_{\rm B}T]\frac{1}{\lambda^{9n_{s,j}}}\times {\rm d}^{\rm n_{v,j}}\vec{x} \quad [32]$$

$$\times \quad k_{\rm wf}\tau {\rm d}\vec{x}/v_{\rm fuse}$$

[33]

Using again $\Delta E_{j\to i} = E_i - E_j$, $n_{s,i} = n_{s,j}$, $n_{VB,i} = n_{{\rm VB},j} - 1$ and $n_{v,i} = n_{v,j} + 1$, detailed balance holds.

Note that detailed balance is satisfied regardless of the values of $k_{\rm wf}\tau$ or $v_{\rm fuse}$, but as with all of the move frequencies these parameters can be optimized during burn-in to accelerate convergence to the equilibrium distribution $P(i)$. In our simulations, we find that the optimal value of $v_{\rm fuse}$ is on the order of the optimal value of $d_{\rm max}$ for analogous reasons: if $v_{\rm fuse}$ is too small there will be very few vertex pairs identified as fusable, so $n_{\rm w}$ will be low. If $v_{\rm fuse}$ is too large, there will be many fusion candidates but most fusion attempts will be rejected due to the large elastic energy change necessary for the merging deformation.

Most importantly, we note the constraint on the parameters $k_{\rm wf}\tau$ to ensure that generation probabilities do not become larger than unity. Because each attempt is generated as a three step process, using three probabilities, one has to ensure that all those probabilities are less than 1. Specifically,

$$n_{\rm w}k_{\rm wf}\tau < 1 \quad [34]$$

$$n_{\rm f}k_{\rm wf}\tau < 1. \quad [35]$$

E. Crack fusion / fission.

E.1. Crack fusion. Crack fusion closes a crack within the structure; i.e., two adjacent pairs of edges are merged (Fig. S21). Cracks are identified as 4-edge-length holes inside the structure. If the vertices of the hole are labeled A, B, C, D then the polygon ABCD forms a closed loop (see Fig. S21). The crack can be closed by either merging vertices A and C (and correspondingly edges CD to DA and AB to BC) or by merging vertices B and D (and correspondingly edges AD to AB and CD to CB). Each 4-edge-length loop thus defines two potential fusable cracks. However, an additional condition for a crack to be fusable is that its merging vertices must be within a distance $l_{\rm fuse}$ (A and C or D and B in this example). In this work, we have set the crack fusion volume to be the same as that for wedge fusion to reduce the number of parameters, but it is not necessary that they be the same.and the acceptance probability is

$$p_{\rm acc}(i\to j) = \min\left[1, \frac{v_{\rm a}}{v_{\rm fuse}}\exp(-\Delta E_{i\to j}/k_{\rm B}T)\right] \quad [36]$$

There are two edge bonds and one vertex bond formed during a crack fusion.

E.2. Crack fission. The reverse move for crack fusion is crack fission. With the number of potential cracks as $n_{\rm cf}$:

$$\alpha(j\to i) = n_{\rm cf}k_{\rm cf}\tau \times \frac{1}{n_{\rm cf}(v_{\rm fuse}/{\rm d}\vec{x})} \quad [37]$$

$$p_{\rm acc}(j\to i) = \min\left[1, \frac{v_{\rm fuse}}{v_{\rm a}}\exp(-\Delta E_{j\to i}/k_{\rm B}T)\right] \quad [38]$$

As for the case of wedge fusion/fission, the crack fusion attempt frequency parameter $k_{\rm cf}$ is constrained by the conditions maintaining probabilities smaller than unity:

$$n_{\rm c}k_{\rm cf}\tau < 1 \quad [39]$$

$$n_{\rm cf}k_{\rm cf}\tau < 1 \quad [40]$$

[41]

F. Edge fusion / fission.

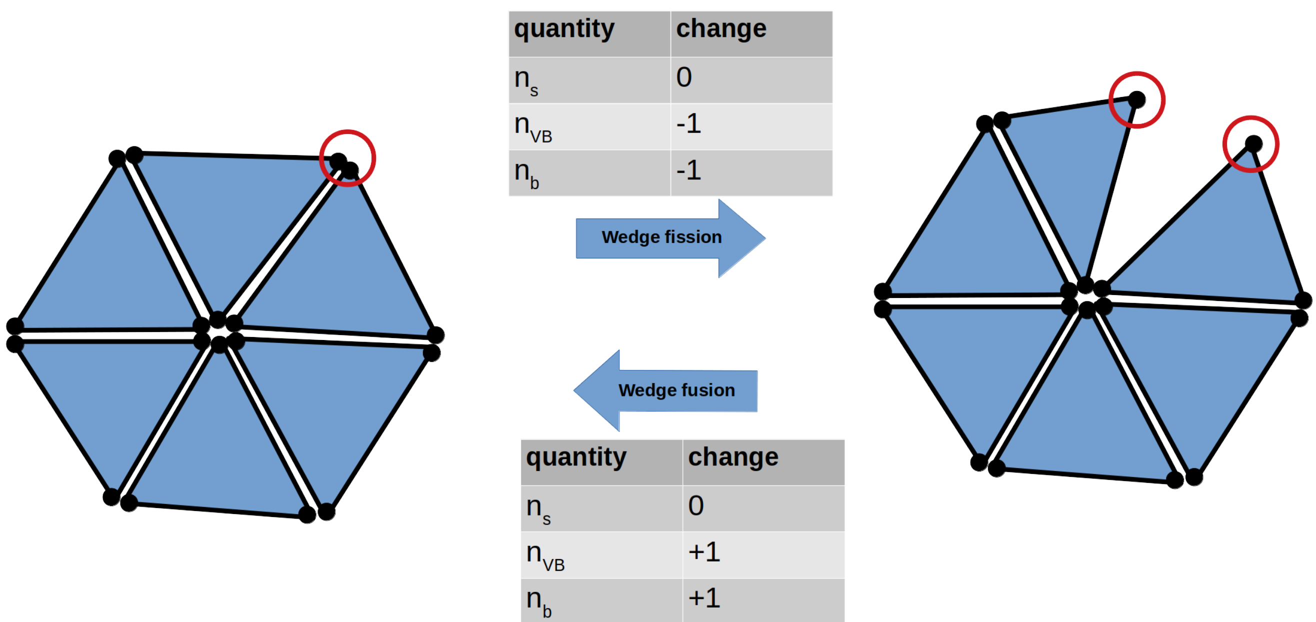
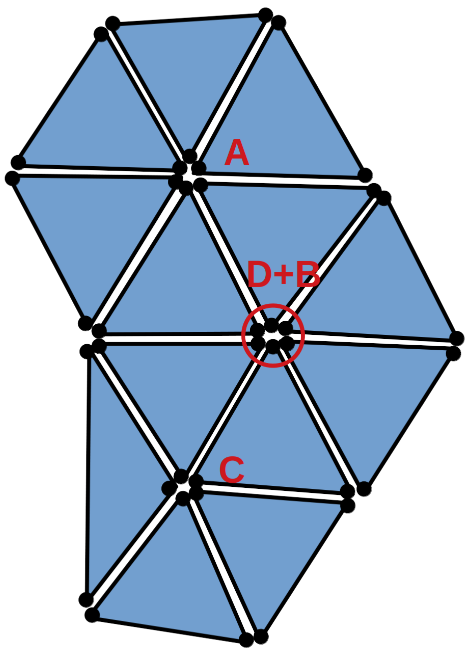


Fig. S20. Wedge fusion and fission.



quantity	change
n_s	0
n_{VB}	-1
n_b	-2

Crack fission

Crack fusion

quantity	change
n_s	0
n_{VB}	+1
n_b	+2

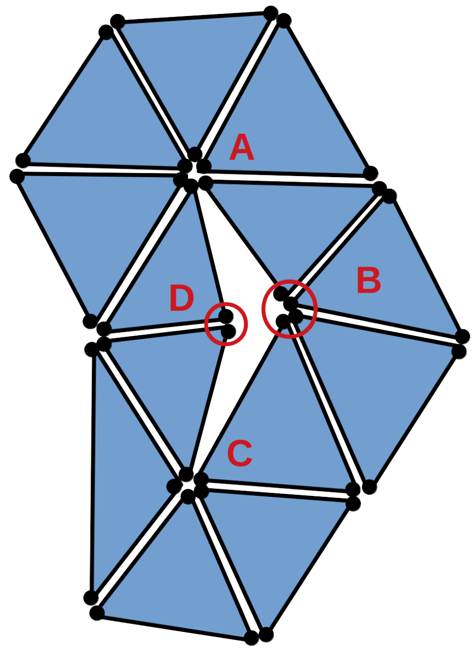


Fig. S21. Crack fusion and fission.

F.1. Edge fusion. During this move two non-neighbor edges are fused (Fig. S22). Fusable edges are non-neighboring edge pairs whose corresponding vertices are within a separation distance l_{fuse}. Since edges are directed, they can only fuse such that, after fusion, they point in the opposite direction. Assuming the edges to be fused are $A \to B$ and $C \to D$ (see Fig. S22), vertex A will merge into vertex D and vertex B will merge into vertex C. Edges are counted as fusable if A is within a distance l_{fuse} to D and B is also within a distance l_{fuse} to C. The attempt probability is analogous to that for wedge and crack fusion/fission,

$$\alpha(i \to j) = n_{\mathrm{e}} k_{\mathrm{ef}} \tau \times \frac{1}{n_{\mathrm{e}}} \tag{42}$$

with n_{e} the number of fusable edges and k_{ef} the edge fusion frequency parameter. The acceptance probability is

$$p_{\mathrm{acc}}(i \to j) = \min\left[1, \left(\frac{v_{\mathrm{a}}}{v_{\mathrm{fuse}}}\right)^2 \exp(-\Delta E_{i\to j}/k_{\mathrm{B}}T)\right] \tag{43}$$

During edge fusion, one edge bond and two vertex bonds are created.

F.2. Edge fission. Edge fission is the reverse move of edge fusion. n_{ef} is the number of breakable edges, that is, those edges that have both vertices on the boundary and which would not result in breaking the structure apart.

$$\alpha(j \to i) = n_{\mathrm{ef}} k_{\mathrm{ef}} \tau \times \frac{1}{n_{\mathrm{ef}}(v_{\mathrm{fuse}}/\mathrm{d}\vec{x})^2} \tag{44}$$

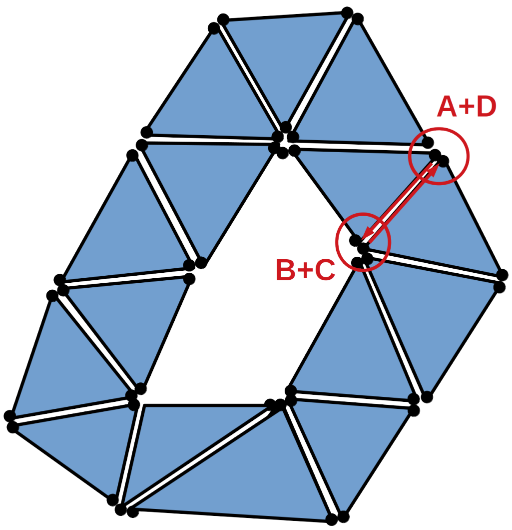
The factor $1/(v_{\mathrm{fuse}})^2$ arises because we must select a random position for each pair of vertices, independently. The acceptance probability is then

$$p_{\mathrm{acc}}(j \to i) = \min\left[1, \left(\frac{v_{\mathrm{fuse}}}{v_{\mathrm{a}}}\right)^2 \exp(-\Delta E_{j\to i}/k_{\mathrm{B}}T)\right]. \tag{45}$$

To maintain probabilities within unity, the edge fusion frequency parameter k_{ef} is constrained by

$$n_{\mathrm{e}} k_{\mathrm{ef}} \tau < 1 \tag{46}$$

$$n_{\mathrm{ef}} k_{\mathrm{ef}} \tau < 1. \tag{47}$$



quantity	change
n_s	0
n_{VB}	-2
n_b	-1

Edge fission

Edge fusion

quantity	change
n_s	0
n_{VB}	+2
n_b	+1

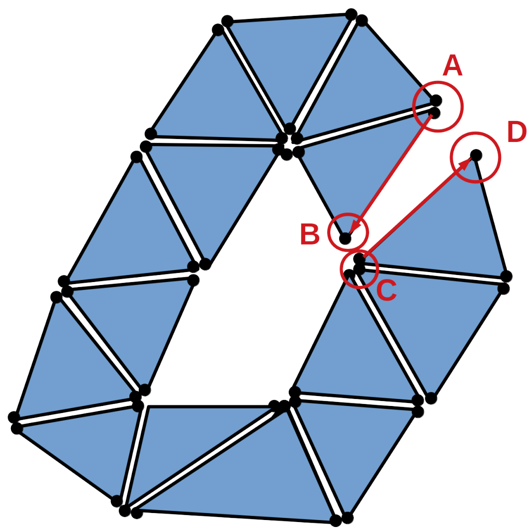


Fig. S22. Edge fusion and fission.

Movie S1. (P3.avi) Self-assembly trajectory for a P_3 structure in the on-target regime with $\eta_s = 10$, $\eta_b = 5$, $\mu = -4.5k_BT$, and $\epsilon_{bind} = -6.5k_BT$ The yellow subunits form a minimal translational simple cubic unit-cell which we use as the seed for self-assembly.

Movie S2. (G3.avi) Self-assembly trajectory for a G_3 structure in the on-target regime with $\eta_s = 10$, $\eta_b = 5$, $\mu = -4.5k_BT$, and $\epsilon_{bind} = -6.5k_BT$ The yellow subunits form a minimal translational bcc unit-cell which we use as the seed for self-assembly.

Movie S3. (D3.avi) Self-assembly trajectory for a D_3 structure in the on-target regime with $\eta_s = 10$, $\eta_b = 5$, $\mu = -4.5k_BT$, and $\epsilon_{bind} = -6.5k_BT$ The yellow subunits form a minimal translational fcc unit-cell which we use as the seed for self-assembly.

Movie S4. (G7_i.avi) Self-assembly trajectory for a G_7 structure in the off-target regime corresponding to the structure shown on Fig. 4i with $\eta_s = 0.1$, $\eta_b = 0.01$, $\mu = -4.5k_BT$, and $\epsilon_{bind} = -6.5k_BT$ The yellow subunits form a minimal translational bcc unit-cell which we use as the seed for self-assembly.

Movie S5. (G7_ii.avi) Self-assembly trajectory for a G_7 structure in the on-target regime corresponding to the structure shown on Fig. 4ii with $\eta_s = 5$, $\eta_b = 5$, $\mu = -4.5k_BT$, and $\epsilon_{bind} = -6.5k_BT$ The yellow subunits form a minimal translational bcc unit-cell which we use as the seed for self-assembly.

Movie S6. (G7_iii.avi) Self-assembly trajectory for a G_7 structure in the unefficient regime corresponding to the structure shown on Fig. 4iii with $\eta_s = 50$, $\eta_b = 10$, $\mu = -4.5k_BT$, and $\epsilon_{bind} = -6.5k_BT$ The yellow subunits form a minimal translational bcc unit-cell which we use as the seed for self-assembly.

References

1. DLD Caspar, A Klug, Physical Principles in the Construction of Regular Viruses. *Cold Spring Harbor Symposia on Quantitative Biology* **27**, 1–24 (1962).
2. T Dotera, H Tanaka, Y Takahashi, Hexagulation numbers: the magic numbers of equal spheres on triply periodic minimal surfaces. *Structural Chemistry* **28**, 105–112 (2017).
3. TE Videbaek, et al., Tiling a tubule: how increasing complexity improves the yield of self-limited assembly. *Journal Physics: Condensed Matter* **34**, 134003 (2022).
4. B Tyukodi, F Mohajerani, DM Hall, GM Grason, MF Hagan, Thermodynamic Size Control in Curvature-Frustrated Tubules: Self-Limitation with Open Boundaries. *ACS Nano* (2022).
5. H Fang, B Tyukodi, WB Rogers, MF Hagan, Polymorphic self-assembly of helical tubules is kinetically controlled. *Soft Matter* **18**, 6716–6728 (2022).
6. C Sigl, et al., Programmable icosahedral shell system for virus trapping. *Nature Materials* **20**, 1281–1289 (2021).
7. GM Rotskoff, PL Geissler, Robust nonequilibrium pathways to microcompartment assembly. *Proceedings National Academy Sciences United States America* **115**, 6341–6346 (2018).
8. M Botsch, S Steinberg, S Bischoff, L Kobbelt, OpenMesh – a generic and efficient polygon mesh data structure in *OpenSG Symposium.* (2002).
9. D Frenkel, B Smit, Understanding molecular simulation: From algorithms to applications (1996).