PAPER • OPEN ACCESS

Tiling a tubule: how increasing complexity improves the yield of self-limited assembly

To cite this article: Thomas E Videbæk *et al* 2022 *J. Phys.: Condens. Matter* **34** 134003

View the article online for updates and enhancements.

Tiling a tubule: how increasing complexity improves the yield of self-limited assembly

Thomas E Videbæk[*] , **Huang Fang** , **Daichi Hayakawa** ,
Botond Tyukodi , **Michael F Hagan** and **W Benjamin Rogers**[*]

Martin A. Fisher School of Physics, Brandeis University, Waltham, MA 02453, United States of America

E-mail: videbaek@brandeis.edu and wrogers@brandeis.edu

CrossMark

Abstract

The ability to design and synthesize ever more complicated colloidal particles opens the possibility of self-assembling a zoo of complex structures, including those with one or more self-limited length scales. An undesirable feature of systems with self-limited length scales is that thermal fluctuations can lead to the assembly of nearby, off-target states. We investigate strategies for limiting off-target assembly by using multiple types of subunits. Using simulations and energetics calculations, we explore this concept by considering the assembly of tubules built from triangular subunits that bind edge to edge. While in principle, a single type of triangle can assemble into tubules with a monodisperse width distribution, in practice, the finite bending rigidity of the binding sites leads to the formation of off-target structures. To increase the assembly specificity, we introduce tiling rules for assembling tubules from multiple species of triangles. We show that the selectivity of the target structure can be dramatically improved by using multiple species of subunits, and provide a prescription for choosing the minimum number of subunit species required for near-perfect yield. Our approach of increasing the system's complexity to reduce the accessibility of neighboring structures should be generalizable to other systems beyond the self-assembly of tubules.

Keywords: self-assembly, multicomponent, Monte Carlo simulations, free-energy landscape, self-limited assembly, valence-limited interactions, complexity

S Supplementary material for this article is available online

(Some figures may appear in colour only in the online journal)

1. Introduction

In recent years, new techniques have been developed that allow for greater control over the types of interactions that can be prescribed between nanometer- and micrometer-scale colloidal particles. In particular, valence-limited interactions provided by patchy colloidal particles [1–7] or DNA origami [8–10] expand the diversity of structures that can be self-assembled, including complex crystal types, like the Kagome [11] or diamond lattices [12], as well as non-crystalline structures, like colloidal molecules, one-dimensional filaments, and two-dimensional sheets [13–31]. The importance of valence-limited interactions in assembling complex structures is that they constrain the number of contacts per particle and can be used to enforce binding at specific angles. This angular specificity has the additional ability to impart an effective curvature to subunits that can cause the assembly to self-close.

A particularly interesting subset of assemblies enabled by specific, directional interactions is self-limited structures, in which one or more dimensions of the final assembly have a finite extent [32]. These types of self-limiting architectures are common-place in living systems, with examples ranging from

microtubules, which are made up of a pair of subunits [33], to viral capsids [34], which can be constructed from a handful of subunits, to ribosomes, which are fully-addressable structures [35]. While groups have made progress in assembling a variety of self-limiting structures from synthetic colloids [16–19, 21, 24, 28, 29], in almost all cases, the self-limiting dimension is comparable to the size of the individual subunits. Creating structures with a self-limited length scale that is larger than the subunits remains a challenge. One recent success in this regard has been seen in the assembly of icosahedral shells using DNA origami, whereby triangular subunits with specific, valence-limited interactions assembled into a range of shells with varying size [10].

A fundamental consequence of introducing a self-limited length scale that is larger than the size of the constituent parts is that the self-limited length can vary due to thermal fluctuations. These fluctuations can cause the system to access unintended final states, leading to a distribution of assemblies rather than just the single target structure [36]. This behavior has been seen in synthetic systems, such as the assembly of rings from wedge-shaped particles made by DNA origami, which form a distribution of ring sizes [9]. There, the off-target states occur due to neighboring minima in the free-energy landscape. In general, the off-target states can either be accessed in equilibrium, as in the case of the rings mentioned above [9] as well as in self-limited, multi-component assemblies [37, 38], or from kinetic traps, where assembly gets caught in a local minimum at early stages of self-assembly and is unable to relax to a lower free energy.

In this report, we examine how using multiple species of particles can limit the formation of off-target states by engineering the free-energy landscape of assembly. Specifically, we study the assembly of triangular subunits into cylindrical tubules. Using both simulation and energetic calculations, we explore how the number of off-target tubules grows with the designed target width and the bending rigidity of the subunits. By finding allowed tilings of the plane with multiple species of triangles, we construct interactions that remove nearby off-target states from the energy landscape. We show that when the periodic length scale associated with a multiple-species tiling becomes comparable to the fluctuations of the self-limited size of the assembly, the target structure assembles with near-perfect yield. This criterion defines a minimum number of subunit species that are needed to guarantee assembly of a prescribed architecture, exemplifying the trade-off between the complexity of the assembly and the distribution of assembly outcomes. Our results provide a route to create tubules with precisely controlled widths and could be extended to other self-limiting architectures.

2. Self-assembly of tubules from triangular subunits

2.1. Subunit design and structure classification

We consider the assembly of cylindrical tubules from triangular subunits, as illustrated in figure 1(a). The subunits are flat equilateral triangles, which bind edge to edge. Since a tube
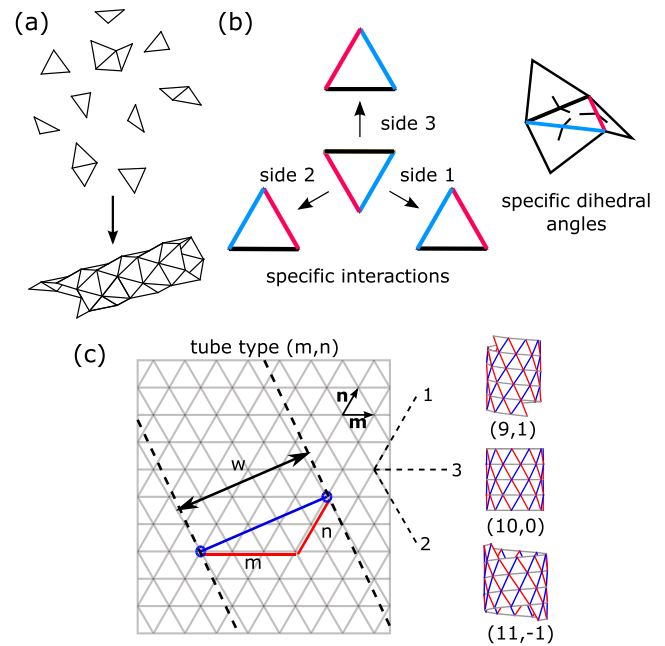
Figure 1. *Design principles of tubule assembly.* (a) Schematic of the self-assembly of triangular subunits into a tubule. (b) Diagram of the specific interactions. Each edge has a specific self-interaction (side 1 binds to side 1, etc) and has a specified dihedral angle between the edges of two bound subunits. These specific angles act as a form of effective curvature for the subunits. (c) Shows how a discrete tubule can be indexed by considering periodic boundary conditions between two vertices of a triangular lattice. The line between the two circled vertices denotes the circumference of the tubule that will form, while the lines labeled by *m* and *n* show the number of steps needed to reach that vertex along lattice directions **m** and **n** and has a width *w*. The labeling of (*m*, *n*) will be used to denote the type of tubule that forms, some examples of which are shown for (9, 1), (10, 0), and (11, −1).

is an object with non-zero curvature, that curvature must be encoded in some way into the subunit. We accomplish this goal by considering subunits that form a specified dihedral angle between their edges when they bind. Therefore, each triangle has associated with it three dihedral angles, one for each side. We consider a case where the interactions are specific and all three sides of the triangle are distinguishable. The simplest set of interactions that prescribe the assembly of a tubule is side 1 binds to side 1, side 2 to side 2, and side 3 to side 3. The specificity of these interactions is important in order to preserve the same local curvature everywhere across the assembly and to ensure that the pattern is deterministic. A schematic of such specific interactions and dihedral angles are shown in figure 1(b).

Any tubule formed from triangular subunits can be classified uniquely by a pair of indices. Consider a tubule as a rolled-up triangular lattice or a sheet with periodic boundary conditions along two parallel lines (figure 1(c)). Within this conceptual framework, a tubule can be constructed by choosing any two points on this plane and enforcing them to be periodic with one another. In other words, these two points overlap when the sheet is rolled up into a tubule, as in figure 1(c). To classify the tubule that forms, we count how many lattice

edges need to be traversed between the periodic vertices. Taking the unit vectors of this tiling to be **m** and **n**, as shown in figure 1(c), we can go between the vertices in *m* and *n* steps in the respective lattice directions, giving a total displacement of $\mathbf{w} = m\mathbf{m} + n\mathbf{n}$. The corresponding tubule would be identified as (m, n) and have a width of $w = \sqrt{m^2 + n^2 + mn}$. Here, the width w refers to the width of the strip on the triangular lattice (figure 1(c)), which corresponds to the circumference of the closed tubule.

For any given tubule, there is a unique set of dihedral angles between adjacent triangular edges that will yield that desired structure. These values can be obtained by constraining the vertices of the lattice to lie on the surface of a cylinder and finding the angle between adjacent faces. Note that in this construction, the specified dihedral angles are constant along a given lattice direction, denoted by indices 1, 2, 3 in figure 1(c). This constraint imposes an orientational order for the triangular subunits, which we will return to later.

2.2. Computational methods

We explore the assembly outcomes using grand canonical Monte Carlo simulations. Specifically, we use the event-driven Monte Carlo algorithm developed by Rotskoff and Geissler [39] and Li *et al* [40], and adapted to tubules in reference [41], in which an assembled structure exchanges subunits with a bath at fixed chemical potential. Each triangular subunit is modeled by three vertices and three straight edges connecting the vertices. The Hamiltonian of the system is given by,

$$H = \frac{1}{2} \sum_{\text{Bound Edges}} E_{\text{B}} + \sum_{\text{Edges}} \frac{1}{2} S(l - l_0)^2$$
$$+ \sum_{\text{Adjacent Faces}} \frac{1}{4} B(\theta - \theta_0)^2. \tag{1}$$

The first term is the binding energy. E_{B} is the energy difference between a pair of bound and unbound edges, and is set to be the same for all favorable interactions. The second term is the stretching energy. S is the stretching modulus of the edge, l is the instantaneous length of the edge, and l_0 is the stress-free length of the edge. The third term is the bending energy. B is the bending modulus of the edge pairs and is again set to be the same for all edges. θ is the instantaneous dihedral angle between two subunits and θ_0 is the preferred dihedral angle for a given tubule structure and type of edge-pair. See supporting information (https://stacks.iop.org/JPCM/34/134003/mmedia) (SI) section 1 for a detailed description of our computational methods.

For a given set of input parameters—the bending modulus, the lattice numbers *m* and *n* of the target, and the number of unique species—we perform one thousand independent simulations and analyze the distribution of tubule types that form. We prescribe the equilibrium dihedral angles, θ_0, to favor $(m, 0)$ and enforce the binding rules specified previously. We tune the various energies in the Hamiltonian to keep the supersaturation low enough that the structure can nucleate, grow, and close near to equilibrium in a reasonable time scale. The simulation starts with a single subunit and grows by adding

subunits onto the pre-existing structure, ending once the tubule has a length roughly three times its circumference. Finally, we determine the tubule type of the end state and compute the distribution of tubule structures for each condition.

We only examine defect-free tubules in the following analysis. We consider a tubule to be defect-free if it has the same tubule structure, characterized by the same pair of indices (m, n) along its entire length. While it is possible for defective tubules to form in our simulations, we find that the likelihood of forming a defective tubule does not depend on the number of subunit species used for assembly. However, the fraction of defect-free tubules does vary between 95%–63% depending on the targeted tubule width (see SI figure S3). We hypothesize that defect creation is related to the degree of supersaturation and the kinetics of monomer addition near closure. This hypothesis is consistent with the type of defective tubules that we observe, which tend to have a central region that does not close properly, resulting in a tubule with a heterogeneous tubule type along its length. See SI section 2 for a description of the algorithm that we use to determine the tubule type and for details on the defect rates.

3. Results and discussion

3.1. The origin of off-target assemblies

The results of our simulations show that assembly yields a broad distribution of tubule types. Figure 2(a) shows an example of such a distribution. The bevel angles for this example were chosen to target a $(10, 0)$ structure. While we do see that the majority of tubules formed the target state, there is a broad distribution of adjacent states that have also formed, extending one to two lattice steps in both **m** and **n** directions. The probability for assembling off-target states falls off further from the target state as expected, since the dihedral angle differences, and hence the elastic energy cost, become more significant.

To understand the origin of this distribution, we consider the process by which assembly occurs and hypothesize that the distribution is dictated by the mechanism of tubule closure. Assembly begins when subunits come together to nucleate a curved sheet. Once the extent of the sheet is large enough, it can close to form a short tubule, which then grows by the addition of subunits to its free ends. In the pre-closure state, thermal fluctuations can cause the sheet to close at larger or smaller widths around the target state, sampling the energies of different closed states. We hypothesize that once the tubule closes, it is highly unlikely to open back into a sheet because doing so would require rupturing multiple edge–edge interactions simultaneously. Therefore, once the assembly has gone through closure and begun to extend, the tubule type is essentially fixed, even if it is not the global free-energy minimum. While the energy difference between off-target states and the target structure continues to increase as the tubule grows longer, without the ability to open again, the system cannot equilibrate to the designed global energy minimum. Indeed, in simulations, we find that the tubule type does not change after a structure has closed (see figure S2(c)).
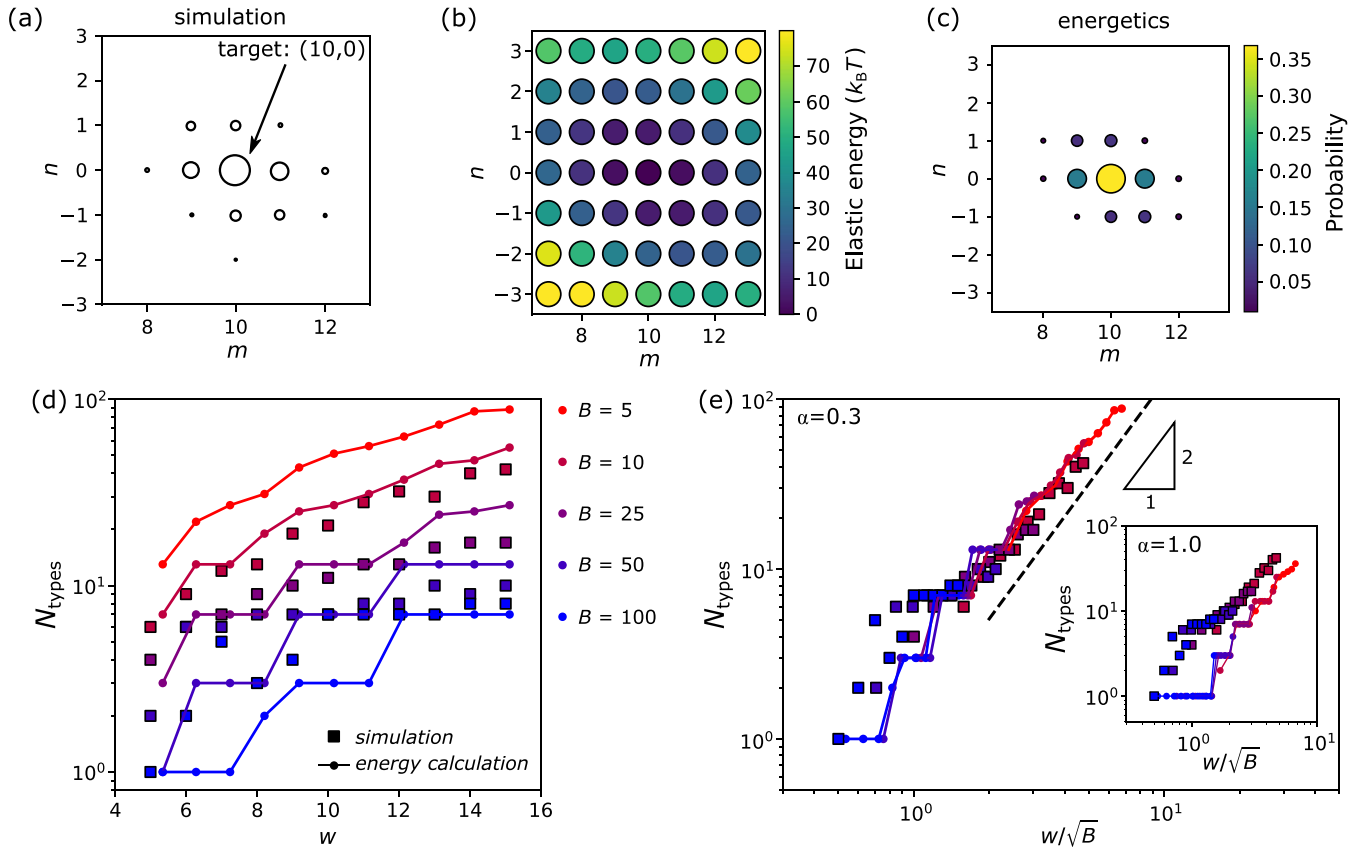
Figure 2. *Reduction of target yield due to fluctuations.* (a) Distribution of tubules from a simulation with a target structure of $(10, 0)$ with $B = 20\ k_{\mathrm{B}}T/\mathrm{rad}^2$. The area of each circle is proportional to the number of defect-free structures that formed. (b) Theoretical elastic energy cost of forming different tubule types with dihedral angles that are designed to form a $(10, 0)$ tubule with $B = 20\ k_{\mathrm{B}}T/\mathrm{rad}^2$. (c) Estimated probability for different assembled tubule types; circle size is proportional to the probability. The probability is estimated by $P(s) \sim \mathrm{e}^{-\alpha E(s)/k_{\mathrm{B}}T}$, where s is the state, $E(s)$ is the elastic energy of that state, and α is a factor adjusting the free-energy barrier described in the text. For clarity, only states with probability greater than 0.005 are shown. (d) Number of tubule types N_{types} that form for varying bending rigidity, B (units of $k_{\mathrm{B}}T/\mathrm{rad}^2$), and target width, w. The points with connected lines are from the elastic energy calculations; the squares are the results from simulations. (e) Number of tubule types versus a rescaled width of $w/\sqrt{B}$. The dashed line is a guide to the eye and shows a power law with slope 2; the inset shows the same data, but with $\alpha = 1.0$.

In addition to running simulations, we also estimate the distribution of structures that may arise during assembly from calculations of the elastic energy at closure. The penalty for forming off-target states comes from the elastic energy cost of not abiding by the prescribed dihedral angles. Recall that the bending energy, E_θ, for any binding site is given by $E_\theta = \frac{1}{2}B(\theta - \theta_0)^2$. For the case of our triangular subunits, the elastic energy per subunit is half the sum over all three binding sites. When closure happens, it occurs for a finite number of subunits in the assembly, N_{C}, all of which contribute to the elastic energy. The full elastic energy cost at closure is therefore

$$E_\theta = \frac{1}{4}N_{\mathrm{C}} \sum_{i=1,2,3} B(\theta_i - \theta_{i,0})^2, \qquad (2)$$

where the extra factor of $\frac{1}{2}$ comes from the fact that a single binding site is shared between two subunits. To estimate the closure size, we assume that subunit addition occurs isotropically, forming a circular disk with a diameter of the tubule width, w. Taking the ratio of the disk area to the area of a

subunit gives $N_{\mathrm{C}} = \pi w^2/\sqrt{3}$. To get a more accurate accounting of the free-energy cost of forming misassembled structures, one would need to include the surface energies and the various entropic contributions, but we find that just considering the bulk elastic terms reproduces the scaling that we find in simulation.

We see the elastic energy at closure is lowest at the target state and increases by a few $k_{\mathrm{B}}T$ at the nearest off-target states, suggesting that these states are likely accessible in a system at finite temperature (figure 2(b)). To get an estimate of the tubule type distribution, we compute the probability of a structure s according to $P(s) = \mathrm{e}^{-\alpha E(s)/k_{\mathrm{B}}T}/Z$, where Z is the partition function and α is a factor that adjusts the height of the free-energy barrier between the pre-closure and the closed states to account for missing contributions to the closure rate (e.g. deviations of pre-closed tubules from the assumed circular shape and other entropic effects). Supplementary simulations found a value of $\alpha = 0.3$ and are described in SI section 3. Without this additional factor the energetics predict narrower distributions. Figure 2(c) shows the probabilities according to the energy

landscape in figure 2(b), which bears a similar shape and extent to the simulated distribution in figure 2(a).

Estimates of the number of accessible states follow power-law scalings with the tubule width and bending rigidity. From the energetics calculation, we make an estimate of the number of different structures, N_{types}, by counting how many states have a probability greater than 0.25%. From simulations, we look at the final tubule structures that form in four hundred different runs. We find that the number of different structures that assemble is sensitive to both the target width and the bending rigidity: the number of accessible states increases with increasing width and decreases with increasing bending rigidity (figure 2(d)). We find that these same data collapse to a single curve when rescaled by $w/\sqrt{B}$ (figure 2(e)). In the inset of figure 2(e) we show the same data but using $\alpha = 1.0$ for the energetics, finding that the number of types predicted is lower than what we see in simulation. This implies that α is capturing some aspects of the kinetic processes during closure.

The scaling for the breadth of the tubule type distribution comes from a balance between the closure size and the fluctuations of the curvature of the sheet before closure. We consider the Helfrich energy of a curved sheet [42], $E = \frac{1}{2}BA(\Delta\kappa)^2$, where A is the area of the sheet and $\Delta\kappa$ is the deviation of its curvature away from the ideal curvature. We can approximate this curvature as $(\Delta w/w^2)$, where Δw is the fluctuations of the width and w^2 is the area of the sheet. Looking at the size of these fluctuations on an energy scale of $k_B T$ shows that $\Delta w \sim w/\sqrt{B}$. Therefore, the scaling that we find in figure 2(e), $N_{types} \sim \Delta w^2$, arises from the fact that thermal fluctuations populate a region of vertices around the target vertex with an area of Δw^2.

These results illustrate a fundamental hurdle for self-limited assembly: in a thermal system, it is difficult to achieve specificity of a target state when the self-limited length scale is large compared to the subunit size. Small fluctuations of the dihedral angles between subunits become amplified as the number of subunits in the self-limited length scale increases. Even though the rigidity of individual dihedral angles may remain the same, the fluctuations of the self-limited length scale grow proportionally to the self-limited lengthscale itself, as we saw from the Helfrich energy. Compounding this effect, the process of irreversible closure prevents the assembly from visiting different states at later times to further relax. Therefore, the breadth of the distribution is driven by a kinetic process, which yields a larger variety of states than would be expected in equilibrium. It is this bottleneck to high-yield of the specific target that must be engineered around, either by making ever stiffer subunits, circumventing closure-control by seeding nucleation with specific geometries, or by altering the energy landscape near closure. Here, we will explore this last direction by considering how multiple types of interacting subunits limit the accessible states at closure, and thereby prune off-target states from the distribution of final structures.

3.2. Allowed tilings with multiple species

To proceed, we extend our framework to allow for multiple species of triangles, where a species of triangle is a subunit
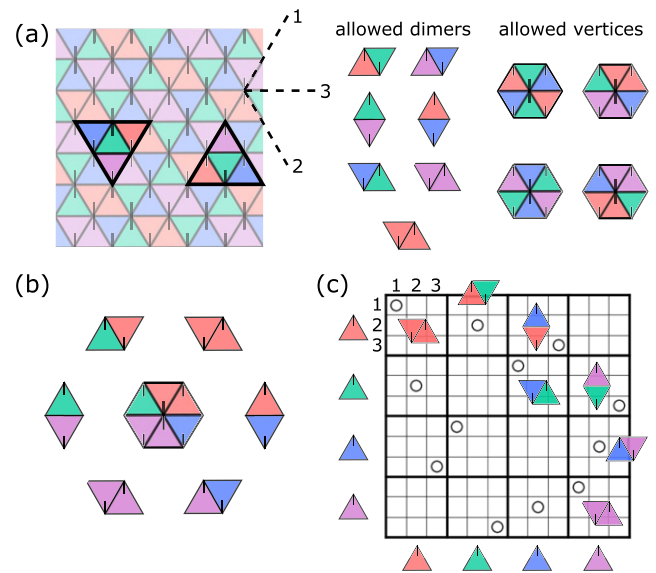
Figure 3. *Rules for allowed tilings.* (a) An allowed four color tiling that can be wrapped into a tubule (pattern 4-4). Dashed lines labeled 1, 2, 3 show the three lattice directions that denote the specific sides of the triangles. Two green triangles and their neighbors are outlined to highlight their interactions. Tick marks on triangles are placed opposite of side 3 to illustrate the triangle orientation. Note that triangles may only appear in orientations with 180-degree rotations. Each specific interaction between two sides of triangles forms a unique dimer. To the right of the pattern are the only allowed dimers and the six-particle vertices that can be made from them. (b) For tilings, we require that each vertex be made up of six valid dimers, i.e. we do not allow non-bonded edges to appear in the tiling. (c) Each pattern can be represented by a symmetric interaction matrix. This matrix corresponds to the pattern shown in (a). Each circle shows a favorable interaction and has its associated dimer next to it. Colored triangles appear next to the columns and rows associated with each color of triangle. Each interior 3×3 block shows the interactions for the three sides of one triangle species.

with a distinct set of specific interactions encoded in its edges. The first task is to identify allowed patterns of multiple species that still satisfy the requirements imposed by the tubule geometry. This challenge involves finding periodic patterns in a triangular lattice with multiple colors of triangles. There are many ways that one can imagine periodically coloring a triangular lattice, but not all of these tilings will necessarily preserve the physical rules for tubule assembly. Here, we introduce three rules for multispecies tubule assembly. First, to have a deterministic assembly, there need to be unique interactions between subunits: each side can only bind to one other side within the mixture of many subunit species. In the context of a tiling, this constraint means that once the colors of tiles adjacent to any subunit have been specified, there cannot be different neighbors at any other location in the tiling. Second, to impose constant dihedral angles along specific lattice directions, only rotations of 180 degrees for a given subunit type are allowed throughout the tiling (figure 3(a)). Third, if we view all interactions between subunits as creating allowed dimers with specific orientations, then we also must be able to construct closed vertices on the plane using either three or six of our specified interactions (figure 3(b)); this last constraint will
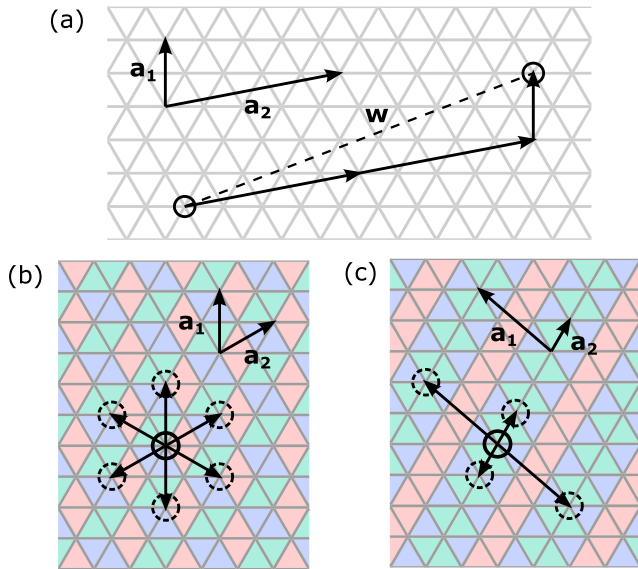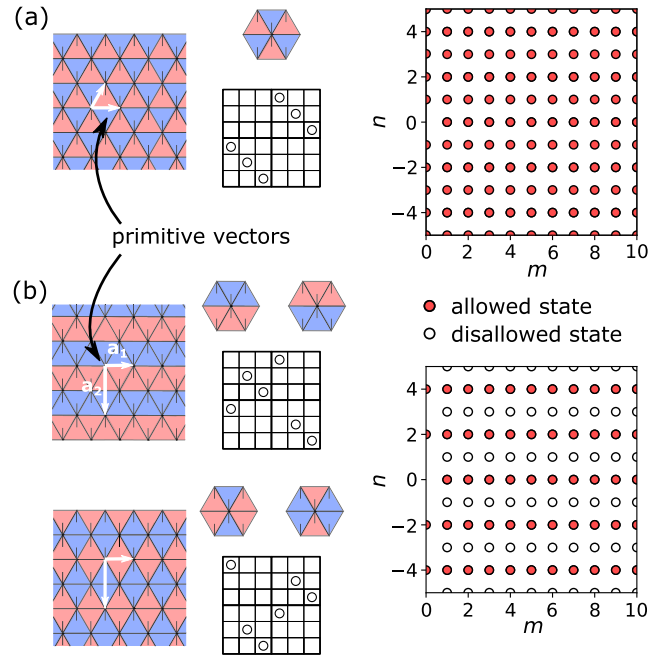
Figure 4. *Restriction of allowed tubule states from multispecies patterns.* (a) Construction of a target tubule type with displacement vector, **w**, from a set of primitive vectors **a₁** and **a₂**. The primitive vectors will differ depending on the specific tilings, but any allowed **w** must be a linear combination of primitive vectors. (b) and (c) Show two different tilings for three species with different primitive vectors. The full circles show a potential target vertex for a tubule. The dashed circles show the nearest vertices accessible using only the primitive vectors of the tiling. (b) Shows uniform distances between nearest vertices (pattern 3-1) while (c) has a large anisotropy between the **a₁** and **a₂** directions (pattern 3-0).

Figure 5. *Two-color tilings and their allowed tubule types.* (a) A two-color tiling that results in no additional specificity (SI pattern 2-2). The arrows on the tiling denote the primitive vectors. For this pattern the single vertex is shown as well as its interaction matrix. On the right plot, filled red points show allowed (m, n) tubules resulting from this tiling. (b) Two other two-color tilings with their respective unique vertices, interaction matrices, and primitive vectors (top: SI pattern 2-0; bottom: SI pattern 2-1). Note that the primitive vectors have the same lengths and orthogonality, resulting in the same set of allowed tubules, shown in the plot to the right. Open points show disallowed states.

force the system to assemble into a deterministic pattern that is the same everywhere in the tiling.

The concept of an interaction matrix helps to make these rules for subunit interactions more concrete. Elements of an interaction matrix will either be non-zero, prescribing an allowed attraction between specific sides of triangles, or zero, meaning no binding is allowed (figure 3(c)). In SI section 4, we describe how the restrictions mentioned above can be translated to allowed constructions of interaction matrices. Following this construction, we enumerated all allowed tilings of a triangular sheet for up to ten different species of subunits. Figure 3 shows one allowed tiling for four species, which are illustrated as four different colors.

Looking at the different tilings that can be wrapped into tubules, we notice that some of them will be more useful for restricting neighboring states than others. The most important feature of a tiling in this regard is its set of primitive vectors (figure 5). For any tiling we can define a pair of vectors, **a₁** and **a₂**, that go between similar vertices in the tiling, have a minimal length, and are maximally orthogonal. These primitive vectors can then be used to identify which tubule types can be formed from a certain pattern. Recall that for tubule type (m, n) there is an associated displacement vector $\mathbf{w} = m\mathbf{m} + n\mathbf{n}$. If **w** can be made from a linear combination of **a₁** and **a₂**, then (m, n) is an allowed state for that tiling (figure 4(a)). See the SI for patterns and interaction matrices.

When the two primitive vectors are large, the similar vertices are farther apart, leading to greater distances between

allowed states. Furthermore, the closer in magnitude the primitive vectors are to one another, the more uniform the restriction of states will be around the target state. If there is a difference in the length of the primitive vectors then the tiling will be anisotropic and there will be an additional orientation dependence of the restricted states with respect to the orientation of the triangles. Two patterns made from three species of triangles are shown in figures 4(b) and (c). From these two patterns we can see that figure 4(b) shows isotropic distances to nearby vertices while figure 4(c) shows anisotropic distances. Depending upon how **w** is aligned with respect to the anisotropic primitive vectors, the assembly outcome changes. For instance, if **w** is aligned along a shorter primitive vector, the density of states in that direction will be larger and will result in larger fluctuations of the widths of tubules that form. In the other case, when **w** is aligned along the longer primitive vector, there will be larger fluctuations in the chirality of tubules that form. To exemplify some of these points, we will look in detail at the patterns formed from two colors.

3.3. Example of tilings with two colors

To illustrate how multispecies tilings can limit the number of accessible off-target states, we first consider the allowed tilings composed of two species. We find that there are only three unique patterns that satisfy the restrictions for assembling
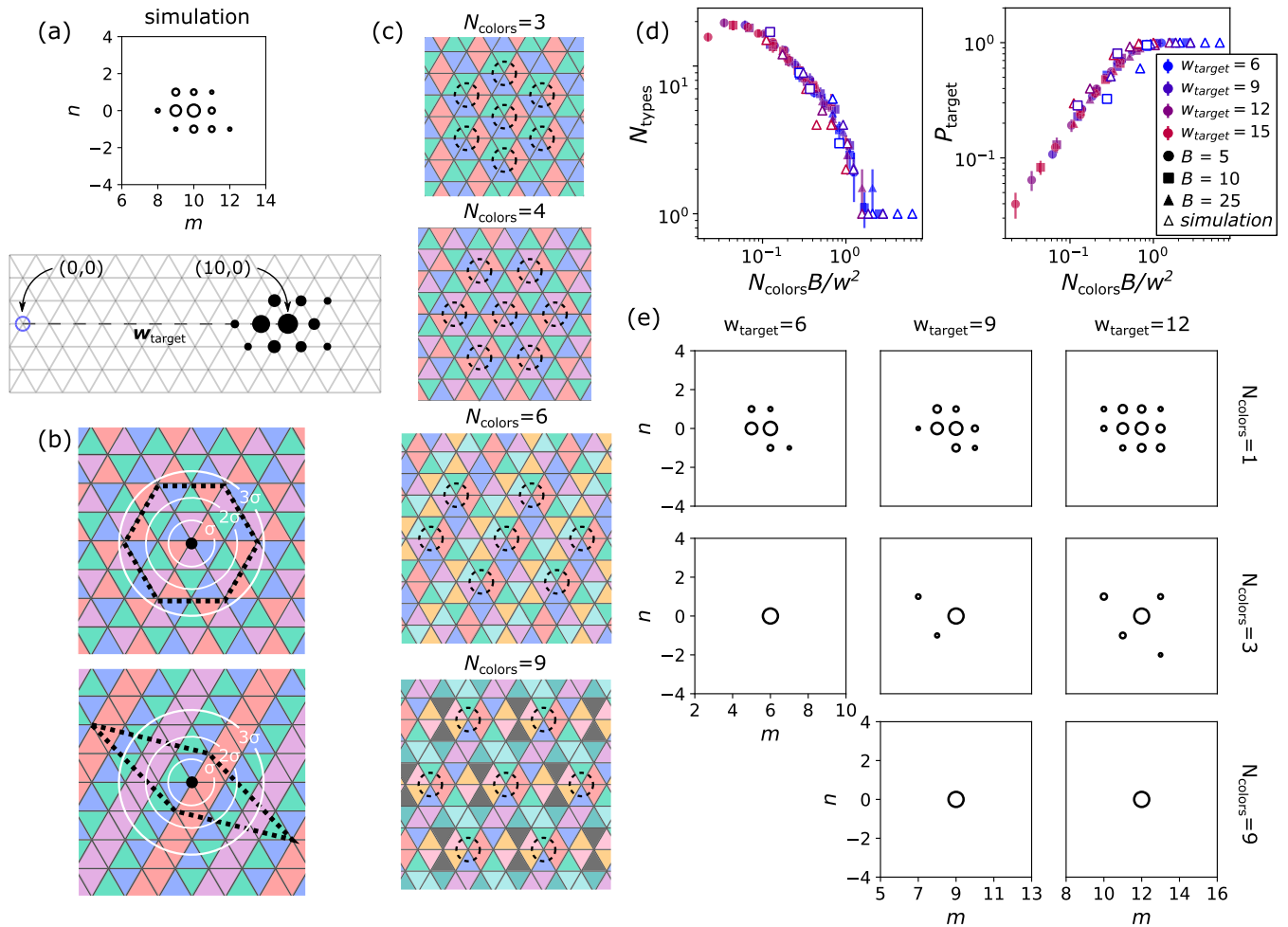
Figure 6. *Increasing the target yield using multiple species.* (a) Tubule type distribution computed from simulations for a (10, 0) target, represented in the *m, n* space (top) and the same distribution overlaid on a triangular lattice (bottom). The diameter of each circles is proportional to the probability of observing that state. On the triangular lattice, we observe that the spatial distribution of states around the target vertex is roughly isotropic. (b) Examples of two four-color tilings (top: SI pattern 4-6; bottom: SI pattern 4-0). In both images the dotted line shows the region around a target vertex where no similar vertex lies. The white circles show the accessible area for closure sites of an assembly due to fluctuations of the pre-closure disk; higher standard deviation (σ) contours denote lower probability regions. The anisotropic tiling excludes fewer of the fluctuations. (c) Patterns used for calculating the reduction of assembly states as we increase the number of species, N_{colors} (from top to bottom SI patterns are 3-1, 4-6, 6-3, 9-2). All but the $N_{\text{colors}} = 6$ pattern are considered isotropic patterns. (d) Number of accessible tubule types, N_{types}, and probability of forming the target state, P_{target}, versus the rescaled number of colors, $N_{\text{colors}}B/w^2$. Results from energetics calculations are the solid points; results from simulations are the open points. We observe a nearly perfect yield of the target state when $N_{\text{colors}}B/w^2 \gtrsim 1$. (e) Examples of simulated tubule distributions for target states (6, 0), (9, 0), and (12, 0) with $N_{\text{colors}} = 1, 3, 9$. Larger target structures have broader initial distributions, but all exhibit a high probability of the target state with a small increase in the number of species for the tiling.

tubules (figure 5), with each pattern having a unique interaction matrix required for assembly. For each pattern, we also compute the allowed tubule geometries. Interestingly, we find that one of the three patterns (figure 5(a)) does not give additional restrictions compared to a single-species tiling. In contrast, the other two tilings reduce the available states by half, albeit with the same restrictions as one another (figure 5(b)). Note that even though some patterns have the same restrictions, all three patterns have different interaction matrices, and thus there is only one designed ground-state tiling for each pattern.

As mentioned above, we rationalize the observation that some tilings restrict the allowed states while others do not

by looking at the primitive vectors of the tilings. For the pattern in figure 5(a), we see that every vertex of the tiling is the same, meaning that all points in *m, n*-space are allowed tubule types. This occurs because the length of its primitive vectors both match the subunit length. In contrast, the two patterns in figure 5(b) have two distinct vertices that cannot overlap with one another on a tubule. Because the two vertex types appear with equal frequency, each vertex can only bind with half of the total vertices in the pattern, thereby reducing the number of accessible states by a factor of two. Additionally, since the primitive vectors for the patterns in figure 5(b) are not equal length, the restrictions imposed on the allowed states are anisotropic. For this pattern, any value of *m* is allowed, while

only even values of n are permitted. Depending on whether or not one wants to more tightly restrict the chirality of the structure or the available widths, one can adjust the orientation of the target displacement vector, **w**, with respect to the two primitive vectors.

A subtle point about the two patterns in figure 5(b) is that even though they produce the same restrictions on allowed states, they require different numbers of specific interactions to encode their patterns. For example, we see that there are five and four unique matrix elements for the top and bottom patterns in figure 5(b), respectively. We hypothesize that reducing the number of interaction types to restrict a greater number of states will be an important design criteria for multiple-species experiments, since there will inevitably be a finite capacity for specific interactions beyond which undesired crosstalk between edges becomes non-negligible.

3.4. Finding the minimal number of species required for high-yield assembly

Now that we understand how a tiling with multiple species can change the number of accessible states, we explore how to design a system that targets a single state with a high yield. As we saw in figure 2(d), the number of off-target states, N_{types}, increases as the area of fluctuations around the target vertex, Δw^2, increases. To compensate for this effect, we expect that the required primitive vector length of a tiling should be comparable to the fluctuations of the self-limited length scale to achieve high yield of the target state. Figure 6(a) shows how a distribution of tubule types relates to the spatial distribution of allowed states around the target vertex. We see that the accessible vertices are nearly uniformly distributed around the target vertex, with a slightly larger variation in the **w** direction. Therefore, we expect that tilings that have primitive vectors of similar lengths will eliminate off-target states most effectively. Figure 6(b) shows examples of two tilings and their respective excluded regions. The first pattern shows an isotropic tiling with the nearest accessible vertices shown connected by the dotted line. The second pattern shows an extreme case in which the excluded region is highly anisotropic. In the short direction, smaller fluctuations would be needed to access another allowed vertex. Going forward, we restrict ourselves to the use of isotropic or near-isotropic tilings; the patterns used for the different numbers of species are shown in figure 6(c).

To see how increasing the number of species, N_{colors} affects the assembly specificity, we perform both simulations and energy calculations as before. Specifically, we calculate the number of accessible states for tubules of different target widths, w_{target}, and bending rigidity, which are assembled from different numbers of subunit species.

We find that the assembly yield of the target state increases as the number of species increases. Figure 6(d) shows the number of tubule types that form, N_{types} (left), as well as the probability of forming the target state, P_{target} (right), as we change the number of species, rescaled by w^2/B. Recalling that the number of tubule types for a single species, N_{types}, grows as w^2/B (figure 2(e)), we consider the quantity w^2/B as the area of the fluctuations around the target vertex. For isotropic tilings the

primitive vectors have a length that grows as $\sqrt{N_{\text{colors}}}$, meaning the excluded area around a target vertex will grow linearly with the number of species. When $N_{\text{colors}}B/w^2 \approx 1$ the area of the disallowed region and the area of fluctuations become comparable. At this point we observe full specificity in both N_{types} and P_{target}. Both the energetics and the simulation show the same scaling.

To clearly illustrate how adding additional species impacts the tubule distributions, we show examples of three different target widths for $N_{\text{colors}} = 1, 3, 9$ in figure 6(e). As the number of species is increased there is a reduction in the number of types of structures that form, with a corresponding increase in the fraction that form the target state. Note that even though the specificity of the target is increased, the extent of the off-target states is not impacted, as seen most clearly in the three-species case. These results further illustrate that the minimum number of subunit species to achieve full specificity is proportional to the size of the fluctuations of the system.

4. Conclusions

In this report, we have shown that multiple species with specific interactions can be used to reduce the assembly of off-target structures for cylindrical tubules. As more species are added to the tiling, the distance between similar vertices increases, corresponding to an area of disallowed states around that target vertex. Full specificity of the target can be achieved when the area of the disallowed region encompasses all off-target states that would have been accessible due to thermal fluctuations. In using multiple species, there is a trade-off between the increase in complexity of the system and the benefit of greater specificity. Therefore, design rules, like the one we develop here, are essential to program the assembly of self-limiting structures in as economical a way as possible. An important note is that while this strategy successfully reduces nearby states for specific targets, the use of multiple species also limits the available target states that can be designed. This idea for how to engineer the free-energy landscape near a target structure should hold for other types of self-limited assembly as well.

An aspect of assembly that has not been discussed in this work is the dynamics of assembly. As the number of subunit species increases, we expect the time for nucleation and growth to increase significantly. This slowing-down of the dynamics of assembly places further emphasis on the need to design the number of unique subunits in as economical a way as possible. Beyond economical design, other strategies might also be combined with multispecies assembly to help overcome kinetic bottlenecks. Nature offers one strategy to improve the target yield without dramatically slowing the assembly kinetics: seeded nucleation. For example, *in vivo*, microtubules—cylindrical cytoskeletal filaments formed from proteins—assemble with a narrow distribution of diameters with the help of seeds [43]. Without such a mechanism to control the initial diameter of the filaments, an array of different tubule structures with varying width and chirality can be seen [44, 45]. Having seeds allows the system to target

a specific structure without sacrificing the kinetics. However, this approach would offer its own set of challenges for synthetic self-assembly with respect to the creation and purity of templates that are used to seed nucleation.

Aside from controlling the assembly specificity, there are other interesting directions that can be explored with the tilings that we have identified. Foremost is the possibility for increasing the addressability of an assembled structure with a tunable length-scale that is not limited to the particle size or the self-limited length scale. In the case of subunits made from DNA origami, we can imagine creating unique addressable sites for conjugating molecules or other small particles to specific subunits within a complex triangular lattice [10, 46, 47]. This strategy could be a way of patterning structures with receptors with certain biological functions or with nanoparticles to create materials with unique photonic responses. Similarly, by leaving out certain species, one could create a user-prescribed pattern of holes in the final structure, enabling the assembly of structures with tunable porosity.

Lastly, we have only explored the role of multiple species in the context of geometrically identical subunits with specific interactions. In terms of self-limited assemblies, there are other closed structures that have variable curvature throughout, such as toroids or helicoids. By engineering a set of allowed interactions between multiple subunits geometries, we could envision changing curvature or edge length for each component as well. Specifically, the linear-tilings that we have found (see SI figure S11 pattern 10-4 as an example) could be used to construct more complex manifolds. Manifolds that have different topologies from a tubule may require finding new tiling patterns and interaction matrices that come from different design rules for that surface.

The recent work by Sigl *et al* [10] provides a promising pathway towards realizing the assembly strategies presented in this report. In their work, they made triangular subunits using DNA origami and used them to assemble icosahedral shells from mixtures containing as many as three distinct subunit species and five unique interactions. One could imagine extending the space of shapes used to facilitate their lock-and-key interactions, as well as adding additional specificity through DNA hybridization. Using these approaches, making experimental systems containing tens of unique subunits should be feasible. An important aspect to consider going forward is how cross-talk between these numerous interactions would impact the fidelity of the assembled structures.

Acknowledgments

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Thomas E Videbæk ⓘ https://orcid.org/0000-0001-9158-2824
Huang Fang ⓘ https://orcid.org/0000-0001-5802-9330
Daichi Hayakawa ⓘ https://orcid.org/0000-0002-2783-5943
Botond Tyukodi ⓘ https://orcid.org/0000-0001-6868-1234
Michael F Hagan ⓘ https://orcid.org/0000-0002-9211-2434
W Benjamin Rogers ⓘ https://orcid.org/0000-0001-8587-8215

References

[1] Kraft D J, Groenewold J and Kegel W K 2009 *Soft Matter* **5** 3823–6
[2] Sacanna S, Korpics M, Rodriguez K, Colón-Meléndez L, Kim S H, Pine D J and Yi G R 2013 *Nat. Commun.* **4** 1688
[3] Yi G-R, Pine D J and Sacanna S 2013 *J. Phys.: Condens. Matter* **25** 193101
[4] Wang Y, Wang Y, Zheng X, Yi G-R, Sacanna S, Pine D J and Weck M 2014 *J. Am. Chem. Soc.* **136** 6866–9
[5] Diaz A J A, Oh J S, Yi G-R and Pine D J 2020 *Proc. Natl Acad. Sci.* **117** 10645–53
[6] Li W, Palis H, Mérindol R, Majimel J, Ravaine S and Duguet E 2020 *Chem. Soc. Rev.* **49** 1955–76
[7] He M, Gales J P, Shen X, Kim M J and Pine D J 2021 *Langmuir* **37** 7246–53
[8] Gerling T, Wagenbauer K F, Neuner A M and Dietz H 2015 *Science* **347** 1446–52
[9] Wagenbauer K F, Sigl C and Dietz H 2017 *Nature* **552** 78–83
[10] Sigl C *et al* 2021 *Nat. Mater.* **20** 1281–9
[11] Chen Q, Bae S C and Granick S 2011 *Nature* **469** 381–4
[12] He M, Gales J P, Ducrot É, Gong Z, Yi G-R, Sacanna S and Pine D J 2020 *Nature* **585** 524–9
[13] Zhang Z and Glotzer S C 2004 *Nano Lett.* **4** 1407–13
[14] Glotzer S C and Solomon M J 2007 *Nat. Mater.* **6** 557–62
[15] Sciortino F, Bianchi E, Douglas J F and Tartaglia P 2007 *J. Chem. Phys.* **126** 194903
[16] Zerrouki D, Baudry J, Pine D, Chaikin P and Bibette J 2008 *Nature* **455** 380–2
[17] Sacanna S, Irvine W T M, Chaikin P M and Pine D J 2010 *Nature* **464** 575–8
[18] Chen Q, Whitmer J K, Jiang S, Bae S C, Luijten E and Granick S 2011 *Science* **331** 199–202
[19] Yan J, Chaudhary K, Bae S C, Lewis J A and Granick S 2013 *Nat. Commun.* **4** 1516
[20] Grünwald M and Geissler P L 2014 *ACS Nano* **8** 5891–7
[21] Bharti B and Velev O D 2015 *Langmuir* **31** 7897–908
[22] Mahynski N A and Panagiotopoulos A Z 2015 *J. Chem. Phys.* **142** 074901
[23] Whitelam S, Tamblyn I, Garrahan J P and Beton P H 2015 *Phys. Rev. Lett.* **114** 115702

[24] Wolters J R, Avvisati G, Hagemans F, Vissers T, Kraft D J, Dijkstra M and Kegel W K 2015 *Soft Matter* **11** 1067–77

[25] Whitelam S 2016 *Phys. Rev. Lett.* **117** 228003

[26] Preisler Z, Saccà B and Whitelam S 2017 *Soft Matter* **13** 8894–902

[27] Morphew D, Shaw J, Avins C and Chakrabarti D 2018 *ACS Nano* **12** 2355–64

[28] Tikhomirov G, Petersen P and Qian L 2018 *J. Am. Chem. Soc.* **140** 17361–4

[29] Oh J S, Lee S, Glotzer S C, Yi G R and Pine D J 2019 *Nat. Commun.* **10** 3936

[30] Ben-Sasson A J *et al* 2021 *Nature* **589** 468–73

[31] Mahynski N A and Shen V K 2021 *Soft Matter* **17** 7853–66

[32] Hagan M F and Grason G M 2021 *Rev. Mod. Phys.* **93** 025008

[33] VanBuren V, Cassimeris L and Odde D J 2005 *Biophys. J.* **89** 2911–26

[34] Caspar D L D and Klug A 1962 Physical principles in the construction of regular viruses *Cold Spring Harbor Symp. Quantitative Biology* vol 27 (New York: Cold Spring Harbor Laboratory Press) pp 1–24

[35] Klinge S and Woolford J L 2019 *Nat. Rev. Mol. Cell Biol.* **20** 116–31

[36] Cheng S and Stevens M J 2014 *Soft Matter* **10** 510–8

[37] Zeravcic Z, Manoharan V N and Brenner M P 2014 *Proc. Natl Acad. Sci. USA* **111** 15918–23

[38] Murugan A, Zeravcic Z, Brenner M P and Leibler S 2015 *Proc. Natl Acad. Sci. USA* **112** 54–9

[39] Rotskoff G M and Geissler P L 2018 *Proc. Natl Acad. Sci. USA* **115** 6341–6

[40] Li S, Roy P, Travesset A and Zandi R 2018 *Proc. Natl Acad. Sci. USA* **115** 10971–6

[41] Tyukodi B, Mohajerani F, Hall D M, Grason G M and Hagan M F 2021 arXiv:2109.01174

[42] Helfrich W and Prost J 1988 *Phys. Rev.* A **38** 3065

[43] Roostalu J and Surrey T 2017 *Nat. Rev. Mol. Cell Biol.* **18** 702–10

[44] Chrétien D and Wade R H 1991 *Biol. Cell* **71** 161–74

[45] Sui H and Downing K H 2010 *Structure* **18** 1022–31

[46] Kuzyk A, Jungmann R, Acuna G P and Liu N 2018 *ACS Photon.* **5** 1151–63

[47] Dey S *et al* 2021 *Nat. Rev. Methods Primers* **1** 1–24