# TEACHING NANOSTRUCTURES TO BUILD THEMSELVES

Gregory Grason, W. Benjamin Rogers, and Michael Hagan

Gregory Grason is a professor of polymer science and engineering at the University of Massachusetts Amherst. **W. Benjamin Rogers** and **Michael Hagan** are faculty of the Martin A. Fisher School of Physics at Brandeis University in Waltham, Massachusetts. The three authors are members of the NSF-funded Bioinspired Soft Materials Center.

Bottom-up self-assembly is a powerful approach to engineering at small scales. Special strategies are needed to formulate components that assemble into predetermined shapes with precise sizes.

(Design by Masie Chong with artwork adapted from Bella Aizenberg, Thomas Videbæk, and iStock.com artists Xansa Green and bgblue.)
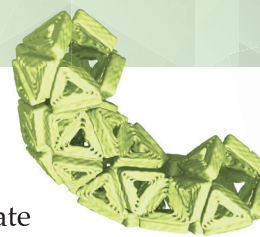
Self-assembly is a widespread phenomenon in which multiple initially disparate building blocks spontaneously arrange into collective multiunit structures. To the materials scientist, self-assembly offers an attractive paradigm to build useful superstructures at especially small size scales purely through the careful design and synthesis of the building blocks. Bottom-up, molecule-by-molecule fabrication enables structural complexity and functional utility that can otherwise be achieved only at larger scales through top-down fabrication methods, like lithography and 3D printing. The approach has already been used to build simple yet useful nanostructures, like spherical shells that can engulf pathogenic viruses. As the ability to define and control the final structure matures, one can envision future applications of self-assembly to build far more complex architectures, like 3D integrated circuitry.

Nature uses self-assembly to build the functional elements of cells and tissues of living organisms in the form of multiprotein superstructures. Those structures include shells, plates, filaments, tubules, ropelike fibers, and nanoporous material networks;[1] two such structures are shown in figure 1. For many biological assemblies, their tightly controlled, finite size is critical to their function. For example, a virus's protein shell, or capsid, must be sufficiently large to enclose the virus's genome but small enough to gain entry to the cells that the virus infects. Photonic nanostructures in the feathers and scales of ani-

mals are composed of multiprotein domains whose sizes directly control what colors they give off.

Self-limiting assemblies in biology exhibit well-controlled structures that are much larger than their components. Proteins that are only a few nanometers in size cooperatively assemble to build structures 10–100 times as large, and they somehow achieve the same finite size from structure to structure, notwithstanding the strong effects of randomness at the nanoscale.

Though self-assembly processes that autonomously terminate at a finite, predetermined size are especially
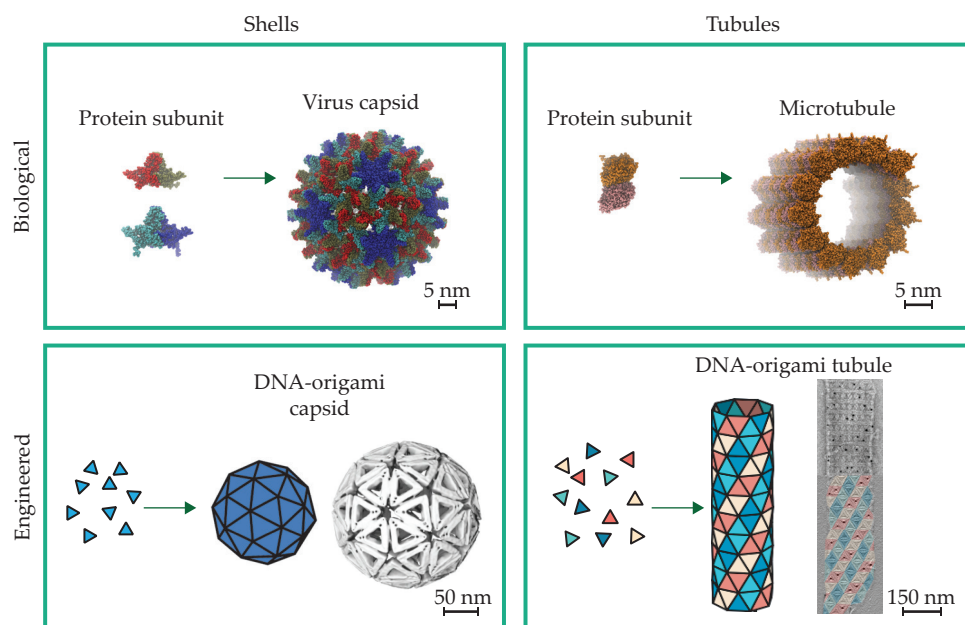

FIGURE 1. SELF-LIMITING ASSEMBLIES observed in biological systems have served as inspiration for synthetically engineered nanostructures. Triangles of folded DNA are particularly useful for building engineered structures because their edges can be designed to form specific bonds that shape the design of the final structure. For the examples of engineered nanostructures, models of self-assembly components and target structures are shown next to micrograph images of self-assembled 3D nanostructures. (Figure adapted from ref. 10 and ref. 17, with contributions from Thomas Videbæk and Layne Frechette.)

attractive to materials scientists, in synthetic self-assembly they are the exception rather than the rule.[2] That's because self-assembly of nanoscopic subunits is driven toward thermal equilibrium, and for most engineered systems, the cohesive forces that bring subunits together also favor forming spatially unbounded assemblies. As depicted in figure 2, unbounded sizes minimize the energy cost of open boundaries at the exterior of assemblies, where the subunits have fewer or weaker cohesive bonds than in the interior.

Because such equilibrium assemblies are driven to arbitrarily large sizes, many synthetic approaches for making size-controlled structures, such as uniformly sized nanoparticles and colloids, rely on nonequilibrium processes.[3] Those approaches might be viewed as supervised assembly because they require the adjustment of parameters like temperature and concentration in a particular temporal sequence to turn assembly processes on and off and trap the aggregating structures at finite and well-defined sizes. Of course, those strategies can't be used where the assembly environment is not under global control, such as in the biological milieu.

For unsupervised self-assembly, in contrast, all the information needed to dictate the assembly pathways and terminate growth must be encoded in the subunits themselves. But what physical mechanisms can tell unsupervised self-assembly when to stop at a particular size—that is, to self-limit? And how much bigger can a self-limiting assembly be than the subunits from which it forms? In this article, we summarize recent efforts to understand the mechanisms of self-limiting assembly, principles for encoding information about finite-size assembly into subunits, and ways in which those principles can inform efforts to engineer new classes of synthetic, biologically inspired nanomaterials with programmable self-limiting structures.

Costs of fitting in

In some scenarios, the size of an assembly is intrinsically limited by the size of its constituents. A classic example is soap. Soap molecules are composed of oily chains with hydrophilic chemical groups at their ends. Toss them into water and they ball up into a sphere, called a micelle, with their hydrophobic ends buried in the middle, shrouded by the water-loving ends of the molecules.[4] The radius of the sphere is limited by the length of the soap molecules. And just as there is room for only a certain number of players' hands in the middle of a sports team huddle, there is room for only a certain number of molecules to reach the center. As the micelle reaches capacity, attempts to
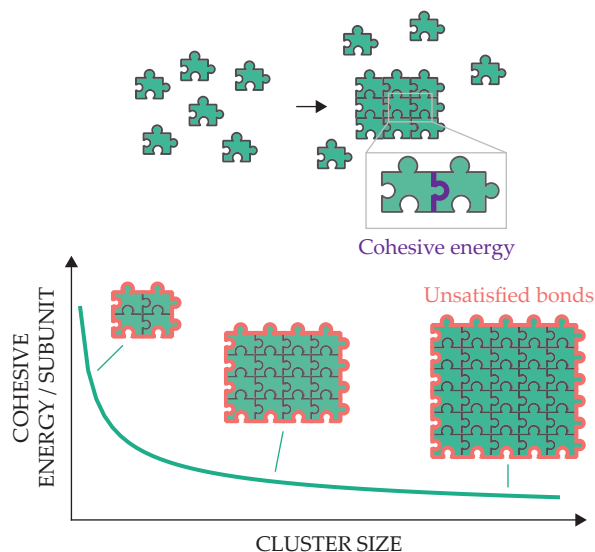

Cohesive energy

Unsatisfied bonds

COHESIVE ENERGY / SUBUNIT

CLUSTER SIZE

FIGURE 2. UNLIMITED SELF-ASSEMBLY is driven by interparticle physical attraction, such as electrostatic forces, van der Waals forces, or hydrogen bonding. Starting from a disassembled state, a system of particles (each puzzle piece represents a particle) moves to a lower energy state by forming cohesive interparticle bonds (represented as fitted edges) and reducing the proportion of unsatisfied bonds. The energetic benefit of increasing the cluster size makes arbitrarily large clusters thermodynamically favored.

squeeze in more molecules or extend the radius strain the components, leading to an added energy cost.

Such excess energy costs depend superextensively on size—the energy cost per subunit increases with additional subunits—and they push back against the generic tendency of cohesion to drive assemblies to unlimited size. Under certain conditions, those superextensive energy costs can result in finite, self-limiting assembly. As shown in figure 3, the increase of excess cost with size can counterbalance cohesive forces to produce a specific assemblage size that, being at an energy minimum, is thermodynamically favored.

But can an assembly limit its finite size at length scales that exceed subunits and the range of their interunit forces? And if so, how could a subunit at the open boundary of such an assembly sense the finite size and determine whether to bind to it? Answering those questions would open up attractive possibilities of designing and synthesizing mixtures of nanoscopic self-assembling building blocks that can be programmed to sense and limit their finite assembly size at microscopic or even millimeter dimensions.

Such seemingly paradoxical thermodynamic action at a distance is possible via geometrically frustrated assembly. Geometric frustration broadly refers to scenarios in which a locally favored packing arrangement cannot be extended uniformly in an ordered system. In the context of assembly, the scenario is illustrated by self-assembling building blocks that are misshapen, like the warped puzzle pieces in figure 3, such that formation of interunit bonds requires distortion of the subunits.

If the multiunit assembly is soft enough, the degree of misfit and its energetic costs will accumulate with size, leading to the type of superextensive costs that restrain self-limiting assembly. The thermodynamic sensitivity at large, multiunit aggregate sizes derives from self-organized strain gradients that extend through the structure: The degree of misfit is higher for the highly constrained and stressed subunits at the core of clusters than for those at the open boundary, where they have fewer bound neighbors and are relatively freer.

That basic paradigm of frustration-limited assembly has been used to explain self-limiting dimensions of a range of biological and synthetic systems,[5] including twisted bundles of protein filaments, liquid-crystalline membranes, and 2D crystals growing on spherical su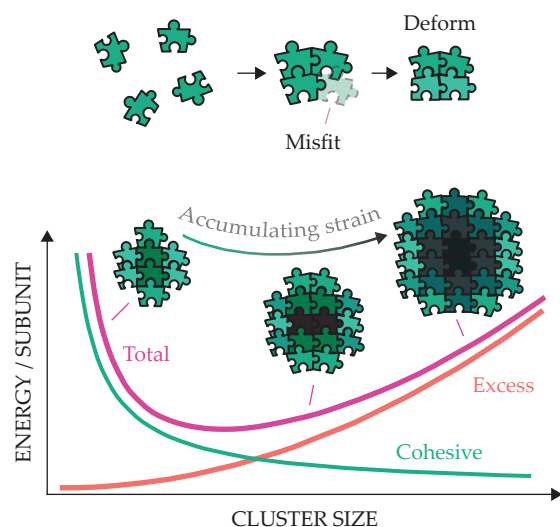rfaces. An emerging and open area of research is exploring avenues to rationally design and engineer frustrated building blocks for the purposes of programming their self-limiting size.[6] Those efforts rely on state-of-the-art techniques for synthesizing colloids or nanoparticles with defined anisotropic shapes and programmable flexibility and interparticle interactions,[7] as well as on critical breakthroughs that use AI for design of *de novo* proteins tailored for an increasing number of assembly shapes.[8]

The ability to fabricate particles of essentially arbitrary 3D shapes raises some basic unsolved questions about the physical principles underlying frustrated self-assembly. Chief among those questions is, Given an arbitrarily misfitting particle design, is it possible to predict the thermodynamic accumulation of misfit cost at the multi-subunit scale? What are the fundamental limits for the maximum size range a frustrated assembly can sense? And under what circumstances do systems escape frustration by forming defective assemblies of unbounded size?
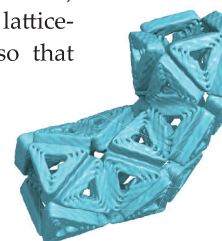
Closing the loop

Another strategy for restraining the effects of open boundaries in cohesive assemblies is to eliminate open boundaries from the target structure. That is the approach used for self-closing assemblies, in which interactions between neighboring subunits are slightly rotated from one to the next. Interunit rotation produces curvature with a preferred radius and can eventually lead the assembly to close upon itself with a finite size larger than the single subunit, as shown in figure 4.

Self-closing assembly is widely employed in multiprotein structures in biology, including cylindrical tubules, which are closed in one dimension, and spherical shells and capsules, which are doubly closed.[9] The ability to target a particular self-closing size relies on careful control of the relative angles between bound subunits, which biology achieves via precise taper of protein subunits.

Recent advances in the design of shape-programmed particles make use of a technique known as DNA origami to achieve an analogous level of curvature-controlled assembly.[10,11] DNA origami exploits the genetic code to fold long loops of single-stranded DNA into essentially arbitrary 3D shapes. (See, for example, Physics Today's 2021 article "DNA assembles nano-objects," by Oleg Gang.) To target a range of self-closing assemblies, researchers begin with triangular DNA particles (those forming the white capsid in figure 1 are 50 nm on a side) whose attractive interedge binding promotes a 2D lattice-like assembly. The subunit edges are beveled so that neighbors bind at controlled angles, thereby driving the assemblies to roll up in specific directions.

The simplest example of that type of curvature control comes from triangular DNA

FIGURE 3. GEOMETRICALLY FRUSTRATED ASSEMBLIES accumulate strain as cluster size increases. The excess energy cost of assembly limits the cluster size. Particles (depicted as warped puzzle pieces) must elastically distort (indicated by darker shading) to form interparticle bonds. In addition to geometric frustration, excess energies may also arise from long-range interactions, such as electrostatic or magnetic repulsion.
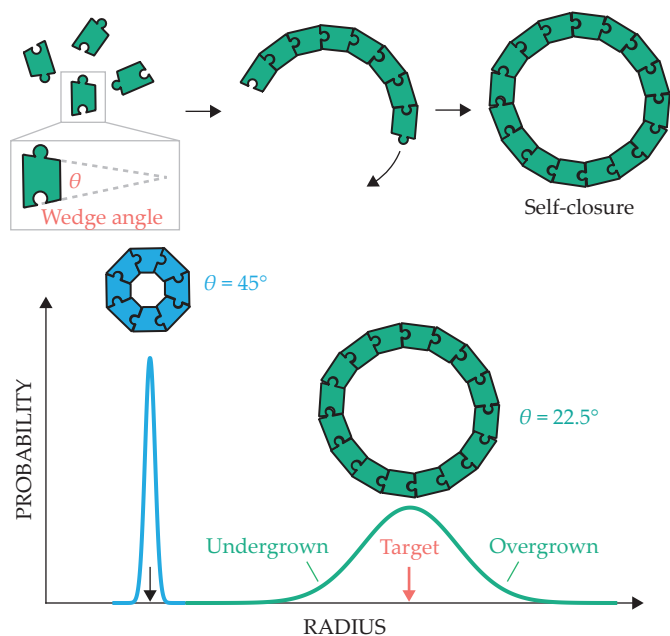
FIGURE 4. SELF-CLOSING ASSEMBLY of a ring is controlled by the wedge angle of its constituents. The radius of the designed target structure depends on the wedge angle. Structures built from more subunits are susceptible to greater deviation from a target size because of propagation of a larger number of small errors in the angles between neighbors.

subunits that curve in one direction to make self-closing cylindrical tubules, as shown in the bottom right of figure 1. In principle, one can target tubules of increasingly large but still finite diameter by redesigning the intertriangle bevel angles to be progressively shallower. In practice, curvature-controlled assembly faces a fundamental limitation because of inevitable fluctuations in the interparticle angle θ. While the particle shape sets a preferred angle, thermal fluctuations generically lead to finite angular deviations away from that target value. Since the angle between particles determines the preferred radius, imprecision in that angle is directly correlated with imprecision in the radius.

No matter how precisely angles between neighbors are defined, the more subunits a structure contains, the more likely imprecision in the self-closing radius becomes (a concept shown schematically in figure 4). A preferred curvature's diminishing power to target ever-larger self-closing assemblies also generalizes to other shapes, such as shells; it was first theoretically predicted by Wolfgang Helfrich in the context of spherical vesicles and microemulsions.[12] That effect of imprecision increasing with size was recently directly measured for assemblies of self-closed tubules of DNA-origami particles.[11] As we describe in the next section, however, one can rescue target fidelity from the increasing imprecision of curvature fluctuations at large scales by increasing the design complexity of self-closing assemblies.

Every subunit in its place

In the examples above, the assemblies can be formed from many copies of a single species of subunits. A different strategy for targeting finite-size equilibrium structures, often called addressable assembly,[13] relies on multiple species of subunits combined with species-specific interactions. The mechanism of assembly takes inspiration from multiprotein complexes in biology, in which several distinct protein subunits bind together in a specific arrangement to form a 3D functional structure, such as a ribosome. Materials scientists have developed several ways to mimic the approach synthetically.

One example is coating colloidal spheres with dangling single-stranded DNA molecules, whose sequence-specific interactions drive self-assembly.[14] In that approach, an assembly may be designed on the basis of multiple particle species A, B, C, and so on. Interactions can be designed so that specific pairs of particle species bind to one another, as illustrated in figure 5. For example, A is not attracted to A but strongly binds to B, while B may bind to both A and itself. The scheme can be described by an interaction matrix. With a sufficiently large number of possible interactions, one can design a multispecies mixture that assembles into a specified structure. If the interaction types are sufficiently complex, there is ideally only a single 3D structure that can self-assemble from that mixture. Since only favorable bonds are formed in the target, it is preferred in thermodynamic equilibrium.

Taking advantage of recent advances in DNA nanotechnology, researchers have been remarkably successful in using that approach to design synthetic assemblies with previously unimaginable complexity, including nanomachine elements and nanobeams of DNA with size-controlled cross sections.[15] In the context of self-limiting assembly, however, addressable assemblies face a particular
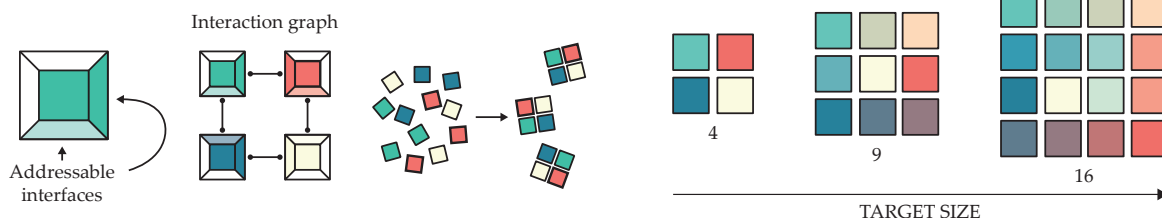
FIGURE 5. PROGRAMMABLE BONDING SITES enable detailed control of self-assembled structures in a technique known as addressable assembly. The technique relies on the ability to configure the subunits' binding sites to interact only with other specific subunit species. The interaction graph shows what the preferred interactions are for this model system; that set of interactions yields specific subunit arrangements. Though the technique provides exquisite control over a final product, the number of unique components needed to build a specific structure increases rapidly as the size of the target structure increases.

challenge: The number of distinct species and interaction types needed to target a structure often grows very rapidly as the desired finite size increases.[14] The larger the finite-size structure is, the higher the costs become in terms of both the expense of synthesizing the species and the time to form the assembly. The time it takes each particle to find its correct position grows with the size of the target structure.

The costs of complexity associated with increasing target size raise general questions about strategies for optimizing the economy of design in addressable assembly: Given a target assembly, what is the minimal complexity of a multispecies mixture needed to robustly achieve it? And how does the minimal complexity grow with increasing finite size?

One approach to answering those questions is to look to biological assemblies, which are subject to evolutionary pressure toward optimized assembly strategies. The quasispherical capsids (see figure 1) that enclose many viruses arguably constitute a case of optimization.

As structural biologists Donald Caspar and Aaron Klug describe, a large family of viral capsids adopt icosahedral symmetry, with protein subunits in a triangular arrangement that is related to the spherical fullerene structure.[16] They argue that the high-symmetry icosahedral arrangement minimizes the number of inequivalent protein conformations needed to form the shell. Therefore, shells composed of a large number of proteins can be built from multiple copies of a much smaller number (by a factor of 60) of distinct species. The genome that is enclosed in the capsid thus only needs to encode for one or a few distinct capsid proteins.

Combining self-closure and high symmetry with ad-dressable assembly provides at least one strategy to mitigate the growth of complexity cost with size. The design paradigm from Caspar and Klug has been directly implemented using DNA origami,[10] which combines both the careful design of bevels that program the spherical curvature and the selective interactions between distinct edges of triangles. Through that strategy, Christian Sigl and coworkers achieved high-yield shell assemblies with modular sizes of 20 to 180 particles per shell (see the example in the bottom left of figure 1).[10]

Though symmetry is a blessing in terms of the economy of addressable assemblies, it can also be curse for self-closing assemblies. That is evidenced by the ability of tubules composed of a single particle type to close at off-target radii: The edge-matching rules of a single species enable the assembling sheet to roll up into many different tubules with similar bending energies. Those results reveal a generic trade-off in self-assembly—increasing assembly economy tends to reduce fidelity.

Thomas Videbæk and coworkers recently explored the optimal balance between maximal economy (one subunit species) and complete specificity (full addressability) in the context of self-closing tubules like that in the bottom right of figure 1.[17] Those experiments indicate that there is a minimum amount of complexity needed to prune away off-target, misassembled states that are accessible through thermal fluctuations and to guarantee nearly 100% yield of the target tubule geometry. Such examples of self-closing assemblies with just enough addressability illustrate that optimal strategies for achieving self-limiting assemblies, a central concern for synthetic efforts, likely require a measured blend of different physical size-control paradigms.

Beyond equilibrium

In many cases, equilibrium processes of assembly are not necessarily sufficient, nor are they the best way to achieve size-regulating nanostructures. Self-assembly is notoriously prone to getting stuck in kinetic traps: Misassembled states rapidly form and then take prohibitively long times to rearrange into their proper target geometries. A key focus of research in the engineering of synthetic assembly systems is the development of nonequilibrium strategies to coax assembling systems toward self-limited structures or other desired end states.

One strategy takes advantage of the programmability of subunits in order to devise time-varying environmental conditions that steer interactions toward target shapes more robustly than can be achieved in passive self-assembly. Another frontier direction is inspired by subunits found in biology that actively consume chemical fuel during assembly and thereby break the rules of equilibrium physics to enable exotic size-regulated assemblies with persistent motion. In such treadmilling assemblies, like cytoskeletal filaments, subunits are constantly binding to one end of a structure and simultaneously unbinding from the other to maintain a fixed and controllable average size.[18]

Though many of the basic paradigms for those classes of nonequilibrium assembly have been studied theoretically, a frontier question remains of how to encode and engineer the necessarily dynamic features into the subunits to achieve assemblies that are far out of equilibrium but remain under control.

REFERENCES

1. R. Phillips et al., *Physical Biology of the Cell*, 2nd ed., Garland Science (2012).
2. M. F. Hagan, G. M. Grason, *Rev. Mod. Phys.* **93**, 025008 (2021).
3. J. P. Rao, K. E. Geckeler, *Prog. Polym. Sci.* **36**, 887 (2011).
4. J. N. Israelachvili, *Intermolecular and Surface Forces*, Academic Press (2011).
5. G. M. Grason, *J. Chem. Phys.* **145**, 110901 (2016).
6. J. F. Berengut et al., *ACS Nano* **14**, 17428 (2020).
7. Z. Su et al., *Prog. Polym. Sci.* **103**, 101230 (2020).
8. Q. M. Dowling et al., *Nature* **638**, 553 (2025).
9. F. Oosawa, S. Asakura, *Thermodynamics of the Polymerization of Protein*, Academic Press (1975).
10. C. Sigl et al., *Nat. Mater.* **20**, 1281 (2021).
11. D. Hayakawa et al., *Proc. Natl. Acad. Sci. USA* **119**, e2207902119 (2022).
12. W. Helfrich, *J. Phys. Fr.* **47**, 321 (1986).
13. W. M. Jacobs, A. Reinhardt, D. Frenkel, *Proc. Natl. Acad. Sci. USA* **112**, 6313 (2015).
14. Z. Zeravcic, V. N. Manoharan, M. P. Brenner, *Rev. Mod. Phys.* **89**, 031001 (2017).
15. Y. Ke et al., *Nat. Chem.* **6**, 994 (2014).
16. D. L. D. Caspar, A. Klug, *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1 (1962).
17. T. E. Videbæk et al., *Sci. Adv.* **10**, eado5979 (2024).
18. L. Mohapatra et al., *Annu. Rev. Biophys.* **45**, 85 (2016).

PT