Expanding protein nanocages through designed symmetry-breaking

Sangmin Lee[1,2,3,4], Ryan D. Kibler[1,2], Quinton Dowling[2,5], Yang Hsia[1,2], Neil P. King[1,2,*], David Baker[1,2,3,*]

[1]Department of Biochemistry, University of Washington, Seattle, WA, USA
[2]Institute for Protein Design, University of Washington, Seattle, WA, USA
[3]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA
[3]Department of Chemical Engineering, Pohang University of Science and Technology (POSTECH), Pohang, Republic of Korea
[5]Department of Bioengineering, University of Washington, Seattle, WA, USA

*Correspondence to: neilking@uw.edu, dabaker@uw.edu

Abstract

Polyhedral protein nanocages have had considerable success as vaccine platforms (*1*–*3*) and are promising vehicles for biologics delivery (*4*–*7*). Hence there is considerable interest in designing larger and more complex structures capable of displaying larger numbers of antigens or packaging larger cargos. However, the regular polyhedra are the largest closed structures in which all subunits have identical local environments (*8*–*11*), and thus accessing larger and more complex closed structures requires breaking local symmetry. Viruses solve this problem by placing chemically distinct but structurally similar chains in unique environments (pseudosymmetry) (*12*) or utilizing identical subunits that adopt different conformations in different environments (quasisymmetry) (*13*–*15*) to access higher triangulation (T) number (*13*) structures with larger numbers of subunits and interior volumes. A promising route to designing larger and more complex nanocages is to start from regular polyhedral nanocages (T=1) constructed from a symmetric homotrimeric building block, isolate cyclic arrangements of these building blocks by substituting in pseudosymmetric heterotrimers, and then build T=4 and larger structures by combining these with additional homo- and heterotrimers. Here we provide a high-level geometric overview of this design approach to illustrate how tradeoffs between design diversity and design economy can be used to achieve different design outcomes, as demonstrated experimentally in two accompanying papers, Lee et al (*16*) and Dowling et al (*17*).

Main Text

A T=1 cage with icosahedral point group symmetry can be built from 20 homotrimers aligned along the icosahedral three-fold symmetry axes (Figure 1a and 1b). The centers of the homotrimers can be visualized as the vertices of 12 pentagonal faces that form the sides of a dodecahedron (Figure 1b). Higher T-number cages conserve the number of pentagons (pentons), which are always aligned along the 12 icosahedral five-fold symmetry axes. The general routes to higher T-number cages described here extract pentons from T=1 cages and place additional trimers between them to introduce hexagons bridging the 12 pentons. For

instance, T=4 structures can be built from the pentons forming a T=1 cage by inserting homotrimers along the icosahedral three-fold symmetry axes connecting three pentons (Figure 1c-e). Other T-number cages can be built by introducing additional trimers between the 12 pentons.

Extracting pentons from T=1 cages requires that one of the three interfaces joining each trimer with its nearest neighbors be eliminated. This can be accomplished by substituting the homotrimers with geometrically identical (i.e., pseudosymmetric) heterotrimers composed of two or three distinct chains with different amino acid sequences. Local symmetry-breaking allows the interface facing outwards from the penton to be specifically eliminated, isolating the penton as a free-standing structure that can be used to build larger structures. This hierarchical approach simplifies the design of high T-number cages through the re-use of trimeric components and designed interfaces that are compatible with the target high T-number architecture. As illustrated in Figure 2 and described below, using 3-chain ABC heterotrimers has the advantages of generating greater structural diversity and precision, while breaking symmetry using 2-chain AAB (and ABB) heterotrimers has the advantage of design simplicity and efficiency.

Maximizing diversity and precision by symmetry-breaking using ABC heterotrimers

The first approach, reported by Lee et al. (*16*), isolates pentons from T=1 structures by substituting ABC-type pseudo-symmetric heterotrimers consisting of three distinct chains in place of the homotrimers (Figure 2c, approach 1). The A_{out}-B_{out} interface (see the interface notation shown in Figure 2a) has the same structure as in the T=1 cage but with heterodimeric sequences, and the C_{out} interface is replaced by polar residues to block further assembly. This extracts a cyclic arrangement of five heterotrimers with C5 symmetry (penton) from the T=1 cage. Next, the C subunit of the pentons is docked with DDD homotrimers (generating a C_{out}-D_{out} interface) to form a closed structure, where the pentons and homotrimers are aligned along the icosahedral five-fold and three-fold symmetry axes, respectively. This results in a T=4 structure (class 1 by the definition of Goldberg (*18*)) consisting of 4 unique chains (A, B, C, and D) and 6 distinct interfaces (A_{in}-B_{in}, A_{in}-C_{in}, B_{in}-C_{in}, D_{in}-D_{in}, A_{out}-B_{out}, and C_{out}-D_{out}). Similarly, T=7 cages (class 3) can be built using two different ABC-type heterotrimers and one homotrimer (GGG), which requires 11 distinct interfaces (Figure 2e). T=9 cages (class 1) can be generated using three different ABC-type heterotrimers and 14 distinct interfaces (Figure 2f). Thus, in this design approach, the required number of unique chains and distinct interfaces linearly increases as the T-number increases (Figure 2d-2f). Although this scaling makes it more challenging to design high (T≥9) T-number cages compared to the strategy described below, the approach has considerable advantages. First, since each designed interface is unique, monodisperse preparations of precisely the target assembly are obtained; alternative structures or mixtures of assemblies are not possible. Second, the structures of each trimeric component and interface can be independently controlled, which allows the design of any T-number cage in any class (Figure 2d-2f) and variation of the detailed geometry within a given cage architecture (Figure 2f and 2g).

Maximizing simplicity and efficiency by symmetry-breaking using AAB heterotrimers

The second approach, reported by Dowling et al. (*17*), isolates pentons from T=1 structures by substituting homotrimers with AAB-type pseudo-symmetric heterotrimers consisting of two distinct chains (Figure 2c, approach 2). The five heterotrimers forming a penton interact with each other through A_{out}-A_{out} interfaces (see the interface notation shown in Figure 2h), and the B_{out} interfaces pointing outward from the penton are redesigned. Although the A_{out}-A_{out} interfaces of such two-chain heterotrimers could theoretically associate with each other in unproductive configurations, the highly cooperative assembly of hierarchically structured cages we have observed experimentally (*17*, *19*) results in such transient associations having no practical effect. Similar to approach 1, docking the pentons to CCC homotrimers produces T=4 structures (Figure 2i) comprising 3 unique chains (A, B, and C) and 5 distinct interfaces (A_{in}-A_{in}, A_{in}-B_{in}, C_{in}-C_{in}, A_{out}-A_{out}, and B_{out}-C_{out}). The next T number that can be designed using this approach is T=9 (class 1), which only requires one more type of heterotrimer derived from two of the existing chains (ABB). This requires the design of one new interface (B_{in}-B_{in}), but other than that, all other interfaces in T=4 cage are reused. Thus, only 3 unique chains and 6 distinct interfaces are needed for T=9 cages. Compared to approach 1, this approach—like virus capsids—combines pseudosymmetry and quasisymmetry to more efficiently use genetic information to expand the T number from T=4 to T=9. This efficiency grows remarkably at higher T numbers: by introducing only one more homotrimeric component from the existing B chain (BBB), the T number can scale infinitely from T=16 onwards to generate extremely large assemblies with no additional design (Figure 2k). However, this approach in turn has its limitations. First, because assemblies with T numbers greater than 16 are constructed from the same set of four oligomeric components, mixtures of different species are obtained. As shown in Dowling et al., this can be partially, although imperfectly, controlled by varying the stoichiometry of the components during assembly. Second, the reuse of chains and interfaces in quasisymmetric environments places certain geometric restrictions on the diversity of structures that can be obtained. Nevertheless, cages with a wide variety of shapes and sizes can be constructed from a minimal number of unique chains and interfaces using this approach.

Outlook

The symmetry-breaking strategies described above can be used to generate a wide variety of protein nanomaterials beyond the constraints of strict point group symmetry. Regular polyhedral nanostructures serve as testing grounds from which symmetric substructures such as pentons can be extracted and combined with other building blocks to generate novel structures. Moving forward, the rich conformational and chemical diversity of proteins together with the control afforded by computational protein design should enable the creation of a remarkable variety of sophisticated functional protein nanomaterials.

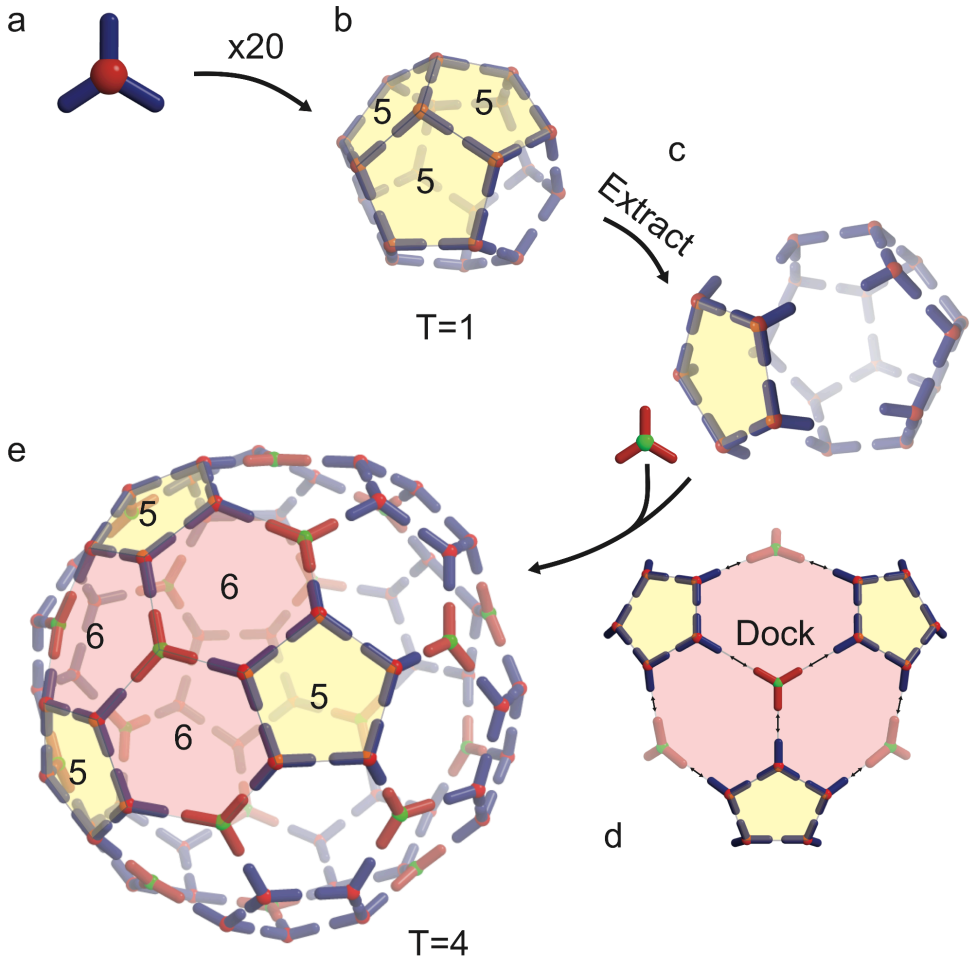

Figure 1. General design route to T=4 icosahedral cages using substructures extracted from T=1 cage. Twelve pentagonal substructures (pentons) are docked to twenty homotrimers to form a closed cage structure, which creates hexagonal local structures (transparent red hexagons) placed between pentons (transparent yellow pentagons).

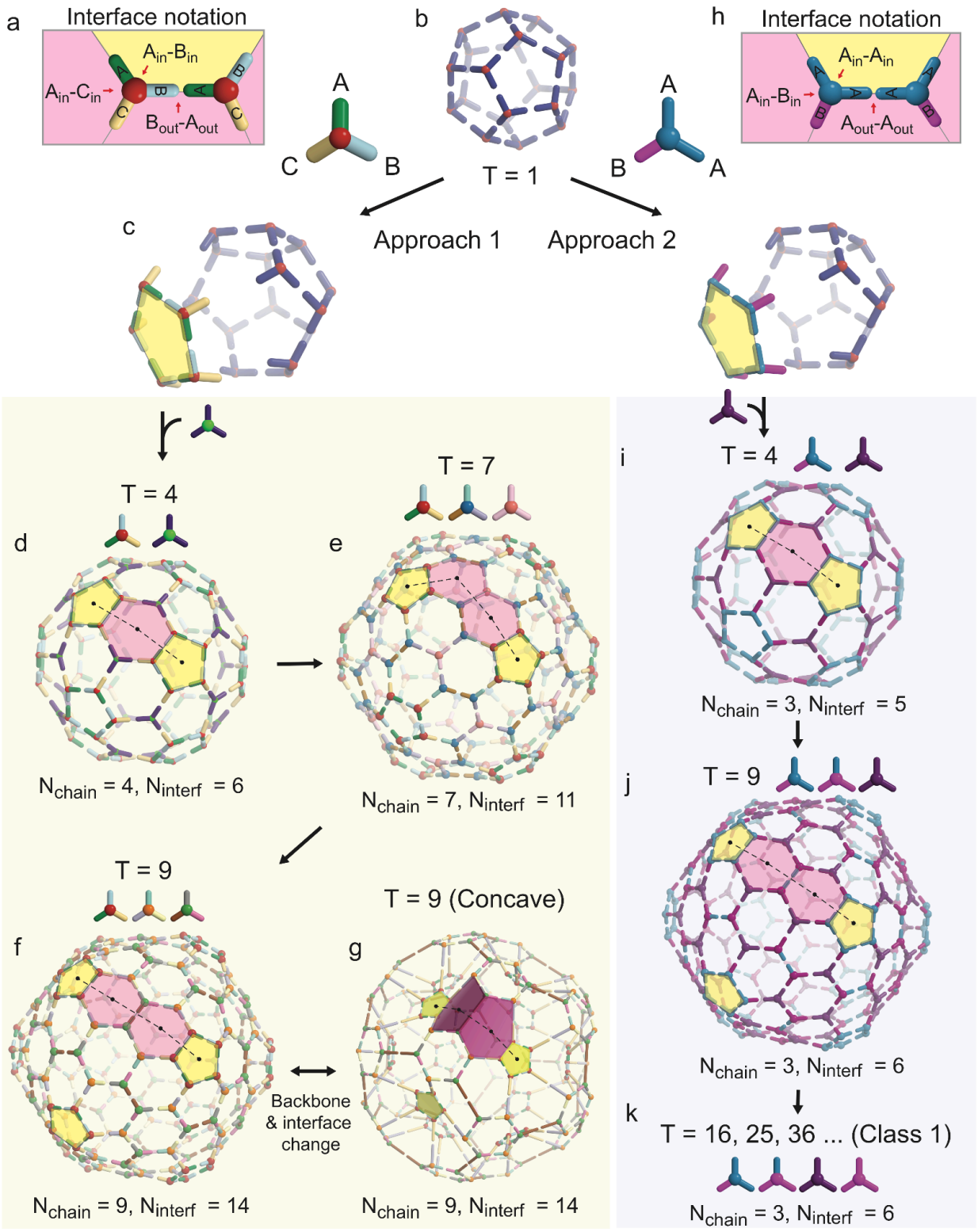

Figure 2. Two design approaches for generating high T-number icosahedral cages using pseudosymmetric building blocks. (**a, h**) Interface notation used in the main text. (**b**) Icosahedral T=1 cage consisting of 20 homotrimers. (**c**) There are two different approaches to extract a

cyclic arrangement of trimers from a T=1 cage. (left) One is to replace the homotrimers with ABC-type heterotrimers, and (right) the other is to replace the homotrimers with AAB-type heterotrimers. (**b, d**-**g**) Design approach 1 (diversity and precision) for constructing high T-number cages using ABC-type heterotrimers. (**b, i**-**k**) Design approach 2 (simplicity and efficiency) for constructing high T-number cages using AAB- and ABB-type heterotrimers. N_{chain} and N_{interf} are the number of unique components and interfaces required for each target architecture, respectively.

References

1. J. Marcandalli, B. Fiala, S. Ols, M. Perotti, W. de van der Schueren, J. Snijder, E. Hodge, M. Benhaim, R. Ravichandran, L. Carter, W. Sheffler, L. Brunner, M. Lawrenz, P. Dubois, A. Lanzavecchia, F. Sallusto, K. K. Lee, D. Veesler, C. E. Correnti, L. J. Stewart, D. Baker, K. Loré, L. Perez, N. P. King, Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell* **176**, 1420-1431.e17 (2019).
2. A. C. Walls, M. C. Miranda, A. Schäfer, M. N. Pham, A. Greaney, P. S. Arunachalam, M.-J. Navarro, M. A. Tortorici, K. Rogers, M. A. O'Connor, L. Shirreff, D. E. Ferrell, J. Bowen, N. Brunette, E. Kepl, S. K. Zepeda, T. Starr, C.-L. Hsieh, B. Fiala, S. Wrenn, D. Pettie, C. Sydeman, K. R. Sprouse, M. Johnson, A. Blackstone, R. Ravichandran, C. Ogohara, L. Carter, S. W. Tilles, R. Rappuoli, S. R. Leist, D. R. Martinez, M. Clark, R. Tisch, D. T. O'Hagan, R. Van Der Most, W. C. Van Voorhis, D. Corti, J. S. McLellan, H. Kleanthous, T. P. Sheahan, K. D. Smith, D. H. Fuller, F. Villinger, J. Bloom, B. Pulendran, R. S. Baric, N. P. King, D. Veesler, Elicitation of broadly protective sarbecovirus immunity by receptor-binding domain nanoparticle vaccines. *Cell* **184**, 5432-5447.e16 (2021).
3. S. Boyoglu-Barnum, D. Ellis, R. A. Gillespie, G. B. Hutchinson, Y.-J. Park, S. M. Moin, O. J. Acton, R. Ravichandran, M. Murphy, D. Pettie, N. Matheson, L. Carter, A. Creanga, M. J. Watson, S. Kephart, S. Ataca, J. R. Vaile, G. Ueda, M. C. Crank, L. Stewart, K. K. Lee, M. Guttman, D. Baker, J. R. Mascola, D. Veesler, B. S. Graham, N. P. King, M. Kanekiyo, Quadrivalent influenza nanoparticle vaccines induce broad protection. *Nature* **592**, 623–628 (2021).
4. K. A. Cannon, J. M. Ochoa, T. O. Yeates, High-symmetry protein assemblies: patterns and emerging applications. *Curr. Opin. Struct. Biol.* **55**, 77–84 (2019).
5. T. G. W. Edwardson, S. Tetter, D. Hilvert, Two-tier supramolecular encapsulation of small molecules in a protein cage. *Nat. Commun.* **11**, 5410 (2020).
6. T. G. W. Edwardson, T. Mori, D. Hilvert, Rational Engineering of a Designed Protein Cage for siRNA Delivery. *J. Am. Chem. Soc.* **140**, 10439–10442 (2018).
7. M. A. Bartel, J. R. Weinstein, D. V. Schaffer, Directed evolution of novel adeno-associated viruses for therapeutic gene delivery. *Gene Ther.* **19**, 694–700 (2012).
8. Y. Hsia, J. B. Bale, S. Gonen, D. Shi, W. Sheffler, K. K. Fong, U. Nattermann, C. Xu, P.-S. Huang, R. Ravichandran, S. Yi, T. N. Davis, T. Gonen, N. P. King, D. Baker, Design of a hyperstable 60-subunit protein icosahedron. *Nature* **535**, 136–139 (2016).
9. Y.-T. Lai, D. Cascio, T. O. Yeates, Structure of a 16-nm Cage Designed by Using Protein Oligomers. *Science* **336**, 1129–1129 (2012).
10. J. B. Bale, S. Gonen, Y. Liu, W. Sheffler, D. Ellis, C. Thomas, D. Cascio, T. O. Yeates, T. Gonen, N. P. King, D. Baker, Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science* **353**, 389–394 (2016).
11. N. P. King, W. Sheffler, M. R. Sawaya, B. S. Vollmar, J. P. Sumida, I. André, T. Gonen, T. O. Yeates, D. Baker, Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science* **336**, 1171–1174 (2012).
12. L. De Colibus, E. Roine, T. S. Walter, S. L. Ilca, X. Wang, N. Wang, A. M. Roseman, D. Bamford, J. T. Huiskonen, D. I. Stuart, Assembly of complex viruses exemplified by a halophilic euryarchaeal virus. *Nat. Commun.* **10**, 1456 (2019).
13. D. L. D. Caspar, A. Klug, Physical Principles in the Construction of Regular Viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1–24 (1962).
14. I. Rayment, T. S. Baker, D. L. D. Caspar, W. T. Murakami, Polyoma virus capsid structure at 22.5 Å resolution. *Nature* **295**, 110–115 (1982).
15. B. V. V. Prasad, M. E. Hardy, T. Dokland, J. Bella, M. G. Rossmann, M. K. Estes, X-ray

Crystallographic Structure of the Norwalk Virus Capsid. *Science* **286**, 287–290 (1999).
16. S. Lee, R. D. Kibler, Y. Hsia, A. J. Borst, A. Philomin, M. A. Kennedy, B. Stoddard, D. Baker, Design of four component T=4 tetrahedral, octahedral, and icosahedral protein nanocages through programmed symmetry breaking. *bioRxiv*, doi: 10.1101/2023.06.16.545341 (2023).
17. Q. M. Dowling, Y.-J. Park, N. Gerstenmaier, E. C. Yang, A. Wargacki, Y. Hsia, C. N. Fries, R. Ravichandran, C. Walkey, A. Burrell, D. Veesler, D. Baker, N. P. King, Hierarchical design of pseudosymmetric protein nanoparticles. *bioRxiv*, doi: 10.1101/2023.06.16.545393 (2023).
18. G. Michael, A class of multi-symmetric polyhedra. *Tohoku Math. J. First Ser.* **43**, 104–108 (1937).
19. A. J. Wargacki, T. P. Wörner, M. van de Waterbeemd, D. Ellis, A. J. R. Heck, N. P. King, Complete and cooperative in vitro assembly of computationally designed self-assembling protein nanomaterials. *Nat. Commun.* **12**, 883 (2021).