Published as: *Science*. 2012 June 1; 336(6085): 1171–1174.

Computational design of self-assembling protein nanomaterials with atomic level accuracy

Neil P. King[1], **William Sheffler**[1], **Michael R. Sawaya**[2], **Breanna S. Vollmar**[3], **John P. Sumida**[4], **Ingemar André**[5], **Tamir Gonen**[3], **Todd O. Yeates**[6,7], and **David Baker**[1,8,*]

[1]Department of Biochemistry, University of Washington, Seattle, WA 98195, USA

[2]Howard Hughes Medical Institute, UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA

[3]Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA

[4]Department of Medicinal Chemistry, University of Washington, Seattle, WA 98177, USA

[5]Department of Biochemistry and Structural Biology, Lund University, Lund, Sweden

[6]UCLA Department of Chemistry and Biochemistry, Los Angeles, CA 90095, USA

[7]UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA

[8]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA

Abstract

We describe a general computational method for designing proteins that self-assemble to a desired symmetric architecture. Protein building blocks are docked together symmetrically to identify complementary packing arrangements, and low-energy protein-protein interfaces are then designed between the building blocks in order to drive self-assembly. Here we use trimeric protein building blocks to design a 24-subunit, 13 nm diameter complex with octahedral symmetry and two related variants of a 12-subunit, 11 nm diameter complex with tetrahedral symmetry. The designed proteins assembled to the desired oligomeric states in solution, and crystal structures of the complexes revealed that the resulting materials closely match the design models. The method can be used to design a wide variety of self-assembling protein nanomaterials.

Molecular self-assembly is an elegant and powerful approach to patterning matter on the atomic scale. Recent years have seen advances in the development of self-assembling biomaterials, particularly those composed of nucleic acids (1). DNA has been used to create, for example, nanoscale shapes and patterns (2), molecular containers (3), and three-dimensional macroscopic crystals (4). Methods for designing self-assembling proteins have progressed more slowly, yet the functional and physical properties of proteins make them attractive as building blocks for the development of advanced functional materials (5, 6). The sophisticated protein-based molecular machines observed in natural systems—which

*To whom correspondence should be addressed. dabaker@uw.edu.

often require self-assembly to function as, for example, cellular motors, pumps, or scaffolds —provide a suggestion of the practical potential of designed protein materials.

In any self-assembling structure, interactions between the subunits are required to drive assembly. Previous approaches to designing self-assembling proteins have satisfied this requirement in various ways, including the use of relatively simple and well-understood coiled-coil and helical bundle interactions (7–11), engineered disulfide bonds (12, 13), chemical cross-links (14), metal-mediated interactions (15, 16), templating by non-biological materials in conjunction with computational interface design (17), or genetic fusion of multiple protein domains or fragments which naturally self-associate (18, 19). In contrast, natural protein assemblies are most often held together by many weak, noncovalent interactions which together form large, highly complementary, low energy protein-protein interfaces (20). Such interfaces spontaneously self-assemble and allow precise definition of the orientation of subunits relative to one another, which is critical for obtaining the desired material with high accuracy (18). Designing assemblies with these properties has been difficult due to the complexities of modeling protein structures and energetics. For instance, a pioneering study used interface design by visual inspection to design new oligomeric structures, yet the experimentally determined dimeric interfaces were largely unanticipated (21). However, recent advances (22–26), including the *de novo* design of a heterodimeric protein interface with atomic level accuracy (27, 28), suggest that our ability to computationally model and design protein-protein interactions is rapidly maturing.

We describe a general computational method for designing self-assembling protein materials which consists of two steps: 1) symmetrical docking of protein building blocks in a target symmetric architecture, followed by 2) design of low energy protein-protein interfaces between the building blocks to drive self-assembly. Here we use as building blocks oligomeric proteins that share an element of symmetry with the target architecture. This reduces by one the number of new protein-protein interfaces that must be designed, since the interface within the oligomer contributes to the self-assembly of the subunits to the target material. Furthermore, the energetic contribution of each designed interaction is multiplied by the symmetry of the building block, which reduces the number of distinct new interactions required to overcome the entropic cost of self-assembly (21).

We used the method to design cage-like protein nanomaterials with either tetrahedral (T) or octahedral (O) point group symmetry (Fig. 1). An assembly with symmetry T requires 12 copies of a protein molecule arranged in 12 symmetry-related orientations, whereas symmetry O requires 24 molecules. Both point groups can be generated from sets of three-fold rotational symmetry axes (Fig. 1A), allowing the use of protein trimers with C3 symmetry as building blocks; in each case, only a single new interface between the trimeric building blocks is required for self-assembly. 271 naturally trimeric protein structures (29) were docked symmetrically in both the T and O target architectures by aligning the three-fold axis of each building block with the three-fold axes in the target architecture and then systematically sampling the two remaining rigid body degrees of freedom, radial displacement and axial rotation, in increments of 1 Å and 1°, respectively (Fig. 1, B and C). For each docked configuration in which no clashes between the backbone and beta carbon atoms of adjacent building blocks were present, a simple proxy for interface size and

complementarity was computed to gauge the "designability" of the configuration (Fig. 1, C and D) (29). Around each of the 10 (O) or 20 (T) most designable configurations for each building block, a set of input structures for design was generated by sampling the radial displacement and axial rotation of the subunits more finely (0.1 Å, 0.5°). For each of these input structures, symmetric RosettaDesign calculations (30, 31) were used to design a new amino acid sequence for the protein that resulted in low-energy, symmetric protein-protein interactions between the trimeric building blocks (Fig. 1, E and F). Designs with the lowest predicted binding energies and geometrically complementary interfaces of sufficient size were further optimized using RosettaDesign and interactive design in foldit (32). 8 T and 33 O designs derived from 15 distinct natural trimeric proteins, containing on average 9 mutations per monomer, were selected for experimental characterization (Table S1).

Genes encoding the designed proteins and the corresponding wild-type trimers were constructed and cloned into an expression vector that appended an 11-residue peptide substrate for fluorescent modification by the *E. coli* acyl-carrier protein synthase AcpS (33). *E. coli* cells expressing the proteins were lysed, the proteins were fluorescently labeled in the clarified lysates by the addition of AcpS and the CoA-488 fluorophore, and the apparent size of each protein was visualized by subjecting the labeled lysates to PAGE under non-denaturing (native) conditions. Out of 7 T and 17 O designs that expressed solubly (Table S1), one designed protein of each architecture revealed a shift in apparent size relative to the corresponding wild-type trimer that suggested self-assembly to the desired material (Fig. 2A). Size exclusion chromatography (SEC) of the labeled lysates confirmed the change in apparent molecular weight for the two designs (Fig. S1). Genes encoding the octahedral design ("O3-33"; 9 mutations from the wild-type protein), the tetrahedral design ("T3-08"; 8 mutations), and the corresponding wild-type trimeric proteins were then subcloned into an expression vector that appended C-terminal $(His)_6$ tags, after which the proteins were expressed and purified by nickel affinity chromatography and SEC.

The designed protein O3-33 eluted from the SEC column as a single peak with an apparent size of about 24 subunits (Fig. 2B). The wild-type protein from which O3-33 was derived (PDB ID 3N79) did not assemble to a higher order structure; it eluted from the column mostly as trimers, with a small peak corresponding to a dimer of trimers (Fig. 2C). Analytical ultracentrifugation revealed that the designed protein sedimented as a single discrete species with a Stokes radius of 7.3 nm, in close agreement with the radius of the designed 24-subunit assembly (Fig. S2). A point mutation (Ala167Arg) that introduced unfavorable steric clashes at the designed interface disrupted the material, suggesting that the observed self-assembly is due to the designed interface (Fig. 2D). Negative stain electron microscopy (EM) of O3-33 revealed fields of monodisperse particles of the expected size (~13 nm), many of which strikingly resembled projections of the design model along its 2-fold, 3-fold, or 4-fold symmetry axes (Fig. 3A). A single particle reconstruction of O3-33 obtained by EM analysis under cryogenic conditions clearly recapitulated the architecture of the design model, verifying that the protein assembles in solution as designed (Figs. 3C and S3).

We solved crystal structures of O3-33 to evaluate the accuracy of our design protocol at high resolution. Structures from two different crystal forms confirmed that the designed material

adopts the target architecture and that the designed interface is responsible for driving self-assembly; the higher resolution (2.35 Å) crystal form is shown in Fig. 3. The structure proved remarkably similar to the design model: the backbone RMSD over all 24 chains is 1.07 Å, and is lower if calculated using only the residues at the interface (0.85 Å). The high resolution of the structure allowed confident determination of the side chain configurations at the designed interface, revealing that the atomic contacts closely match those in the design model (Fig. 3E). The asymmetric unit of the designed interface consists of one alpha helix packing against a beta strand, a loop, and the symmetrically related helix in a neighboring building block. Several ordered water molecules were resolved at the designed interface that contribute bridging hydrogen-bonding interactions between neighboring building blocks. Truncation of designed interface residues to alanine disrupted octahedral self-assembly (Fig. S4). For example, the Ser156Ala mutation, which alters O3-33 by the removal of only two atoms out of 2,827 total atoms in the subunit, significantly impaired assembly. This result underscores the importance of both the detailed atomic contacts designed by our protocol and the multiplicative effect of the symmetry of the system: the Ser156Ala mutation results in the loss of 24 interface hydrogen bonds in the fully assembled material.

The designed protein T3-08 appeared by SEC to be in a slow equilibrium between two states comprising three and approximately 12 subunits (Fig. 2E). The corresponding wild-type trimeric protein (PDB ID 3FTT) eluted from the column as trimer only (Fig. 2G). Disruption of the designed interface by a point mutation, Ala52Gln, again suggested that the designed interface is responsible for the observed self-assembly (Fig. 2H). A crystal structure of T3-08 revealed that the protein assembles to the desired tetrahedral architecture, but the trimeric building blocks are slightly rotated about the shared trimeric/tetrahedral three-fold rotational axes, subtly altering the atomic contacts at the designed interface relative to the design model and resulting in a backbone RMSD of 2.66 Å over all 12 subunits (Fig. S5).

We designed two additional variants of T3-08 (Table S1) in order to determine whether we could preferentially stabilize the designed configuration over the unanticipated configuration observed in the T3-08 crystal structure. One of the variants, T3-10, which contained 3 mutations relative to T3-08 intended to provide better hydrophobic packing near the tetrahedral three-fold interface (Fig. S6), was purified by nickel affinity chromatography and appeared by SEC to self-assemble efficiently to the tetrahedal state, yielding little detectable trimer (Fig. 2F). Negative stain EM images of T3-10 revealed monodisperse particles of the expected size (~11 nm), averages of which closely resembled projections of the design model along its 2-fold and 3-fold symmetry axes (Fig. 4, A and B). A crystal structure of T3-10 verified that the original designed configuration was stabilized as intended; the backbone RMSD between the T3-10 crystal structure and the T3-08/T3-10 design models is 0.62 Å (Fig. 4, B and C). As observed for O3-33, the atomic contacts at the designed interface, which consists of two alpha helices and two short loops, closely match those in the design model (Fig. 4D). This result illustrates how small alterations to the protein sequence at the designed interface may allow fine control over the structure of the resulting material.

Our results establish a method by which self-assembling protein materials may be designed with high accuracy. The design strategy, combining symmetrical docking with interface design, is conceptually simple and generally applicable to the design of a broad range of

symmetric materials. In addition to the finite, cage-like materials described here, unbounded materials in one, two, or three dimensions (i.e., fibers/helices, layers, or crystals) may be designed by choosing an appropriate target symmetric architecture. Although in the present study we used naturally occurring oligomeric proteins as building blocks, novel oligomeric building blocks could first be designed from monomers and after structural validation, used in the design of higher order assemblies with the attendant advantages of hierarchical assembly, or, with improvements in our symmetrical docking protocol, larger self-assembling systems could be designed directly from monomeric building blocks. The atomic-level accuracy of our designed materials demonstrates that using designed protein-protein interfaces to drive self-assembly results in highly ordered materials with superior rigidity and monodispersity. With further development, designed self-assembling protein materials similar to those described here could form the basis of advanced functional materials and custom-designed molecular machines with wide-ranging applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Jason Navarro for assistance with protein crystallization at the UCLA crystallization core facility, which is supported by DOE-BER Grant Number DE-FC03-02ER63421. We are grateful to Duilio Cascio and the 24-ID-C beamline staff for their assistance in data collection. This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under contract no. DE-AC02-06CH11357. AUC was performed in the Bioanalytical Pharmacy Core at the University of Washington, which is supported by the Washington State Life Sciences Discovery Fund and the Center for the Intracellular Delivery of Biologics. We are grateful to Matt Iadanza for EM analysis of T3-10. We would like to thank Yifan Cheng and Xueming Li (UCSF) for giving us access to their electron cryomicroscope for data collection, for helpful discussions, and for sharing scripts with us. We also thank Nikolas Grigorieff (Brandeis) for helpful discussions. The Gonen laboratory is supported by HHMI. Work by NPK, WS, and DB was supported by DOE, HHMI, and the International Aids Vaccine Initiative. Coordinates and structure factors were deposited in the Protein Data Bank with the accession codes 3VCD (O3-33, R32 crystal form), 4DDF (O3-33, P4 crystal form), 4DCL (T3-08), and 4EGG (T3-10).

References

1. Seeman NC. Nanomaterials based on DNA. Annu Rev Biochem. 2010; 79:65–87. [PubMed: 20222824]
2. Rothemund PW. Folding DNA to create nanoscale shapes and patterns. Nature. 2006; 440:297–302. [PubMed: 16541064]
3. Andersen ES, et al. Self-assembly of a nanoscale DNA box with a controllable lid. Nature. 2009; 459:73–6. [PubMed: 19424153]
4. Zheng J, et al. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. Nature. 2009; 461:74–7. [PubMed: 19727196]
5. Zhang S. Fabrication of novel biomaterials through molecular self-assembly. Nat Biotechnol. 2003; 21:1171–8. [PubMed: 14520402]
6. Douglas T, Young M. Viruses: making friends with old foes. Science. 2006; 312:873–5. [PubMed: 16690856]
7. Lovejoy B, et al. Crystal structure of a synthetic triple-stranded alpha-helical bundle. Science. 1993; 259:1288–93. [PubMed: 8446897]

8. Harbury PB, Plecs JJ, Tidor B, Alber T, Kim PS. High-resolution protein design with backbone freedom. Science. 1998; 282:1462–7. [PubMed: 9822371]
9. Gribbon C, et al. MagicWand: a single, designed peptide that assembles to stable, ordered alpha-helical fibers. Biochemistry. 2008; 47:10365–71. [PubMed: 18767812]
10. Zaccai NR, et al. A de novo peptide hexamer with a mutable channel. Nat Chem Biol. 2011; 7:935–41. [PubMed: 22037471]
11. Koder RL, et al. Design and engineering of an O(2) transport protein. Nature. 2009; 458:305–9. [PubMed: 19295603]
12. Ballister ER, Lai AH, Zuckermann RN, Cheng Y, Mougous JD. In vitro self-assembly of tailorable nanotubes from a simple protein building block. Proc Natl Acad Sci USA. 2008; 105:3733–8. [PubMed: 18310321]
13. Usui K, et al. Nanoscale elongating control of the self-assembled protein filament with the cysteine-introduced building blocks. Protein Sci. 2009; 18:960–9. [PubMed: 19384998]
14. Ringler P, Schulz GE. Self-assembly of proteins into designed networks. Science. 2003; 302:106–9. [PubMed: 14526081]
15. Salgado EN, Faraone-Mennella J, Tezcan FA. Controlling protein-protein interactions through metal coordination: assembly of a 16-helix bundle protein. J Am Chem Soc. 2007; 129:13374–5. [PubMed: 17929927]
16. Salgado EN, Radford RJ, Tezcan FA. Metal-directed protein self-assembly. Acc Chem Res. 2010; 43:661–72. [PubMed: 20192262]
17. Grigoryan G, et al. Computational design of virus-like protein assemblies on carbon nanotube surfaces. Science. 2011; 332:1071–6. [PubMed: 21617073]
18. Padilla JE, Colovos C, Yeates TO. Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments. Proc Natl Acad Sci USA. 2001; 98:2217–21. [PubMed: 11226219]
19. Sinclair JC, Davies KM, Vénien-Bryan C, Noble ME. Generation of protein lattices by fusing proteins with matching rotational symmetry. Nat Nanotechnol. 2011; 6:558–62. [PubMed: 21804552]
20. Janin J, Bahadur RP, Chakrabarti P. Protein-protein interaction and quaternary structure. Q Rev Biophys. 2008; 41:133–80. [PubMed: 18812015]
21. Grueninger D, et al. Designed protein-protein association. Science. 2008; 319:206–9. [PubMed: 18187656]
22. Kortemme T, et al. Computational redesign of protein-protein interaction specificity. Nat Struct Mol Biol. 2004; 11:371–9. [PubMed: 15034550]
23. Huang PS, Love JJ, Mayo SL. A de novo designed protein protein interface. Protein Sci. 2007; 16:2770–4. [PubMed: 18029425]
24. Jha RK, et al. Computational design of a PAK1 binding protein. J Mol Biol. 2010; 400:257–70. [PubMed: 20460129]
25. Karanicolas J, et al. A de novo protein binding pair by computational design and directed evolution. Mol Cell. 2011; 42:250–60. [PubMed: 21458342]
26. Stranges PB, Machius M, Miley MJ, Tripathy A, Kuhlman B. Computational design of a symmetric homodimer using β-strand assembly. Proc Natl Acad Sci USA. 2011; 108:20562–7. [PubMed: 22143762]
27. Fleishman SJ, et al. Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. Science. 2011; 332:816–21. [PubMed: 21566186]
28. Fleishman SJ, et al. Hotspot-centric de novo design of protein binders. J Mol Biol. 2011; 413:1047–62. [PubMed: 21945116]
29. Information on materials and methods is available on Science Online.
30. Kuhlman B, Baker D. Native protein sequences are close to optimal for their structures. Proc Natl Acad Sci USA. 2000; 97:10383–8. [PubMed: 10984534]
31. DiMaio F, Leaver-Fay A, Bradley P, Baker D, André I. Modeling symmetric macromolecular structures in Rosetta3. PLoS ONE. 2011; 6:e20450. [PubMed: 21731614]

32. Cooper SF, et al. Predicting protein structures with a multiplayer online game. Nature. 2010; 466:756–60. [PubMed: 20686574]
33. Zhou Z, et al. Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases. ACS Chem Biol. 2007; 2:337–46. [PubMed: 17465518]
34. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007; 372:774–97. [PubMed: 17681537]
35. Levy ED, Pereira-Leal JB, Chothia C, Teichmann SA. 3D complex: a structural classification of protein complexes. PLoS Comput Biol. 2006; 2:e155. [PubMed: 17112313]
36. Fleishman SJ, et al. Community-wide assessment of protein-interface modeling suggests improvements to design methodology. J Mol Biol. 2011; 414:289–302. [PubMed: 22001016]
37. Lawrence MC, Colman PM. Shape complementarity at protein/protein interfaces. J Mol Biol. 1993; 234:946–50. [PubMed: 8263940]
38. Sheffler W, Baker D. RosettaHoles2: a volumetric packing measure for protein structure refinement and validation. Protein Sci. 2010; 19:1991–5. [PubMed: 20665689]
39. Fleishman SJ, Khare SD, Koga N, Baker D. Restricted sidechain plasticity in the structures of native proteins and complexes. Protein Sci. 2011; 20:753–7. [PubMed: 21432939]
40. Prodromou C, Pearl LH. Recursive PCR: a novel technique for total gene synthesis. Protein Eng. 1992; 5:827–9. [PubMed: 1287665]
41. Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J. 2000; 78:1606–19. [PubMed: 10692345]
42. Ohi M, Li Y, Cheng Y, Walz T. Negative staining and image classification - powerful tools in modern electron microscopy. Biol Proced Online. 2004; 6:23–34. [PubMed: 15103397]
43. Smith J. XIMDISP—A visualization tool to aid structure determination from electron microscope images. J Struct Biol. 1999; 125:223–228. [PubMed: 10222278]
44. van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M. A new generation of the IMAGIC image processing system. J Struct Biol. 1996; 116:17–24. [PubMed: 8742718]
45. Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. J Struct Biol. 2003; 142:334–47. [PubMed: 12781660]
46. Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. J Struct Biol. 1999; 128:82–97. [PubMed: 10600563]
47. Grigorieff N. FREALIGN: high-resolution refinement of single particle structures. J Struct Biol. 2007; 157:117–25. [PubMed: 16828314]
48. Pettersen EF, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004; 25:1605–12. [PubMed: 15264254]
49. Tsai Y, Sawaya MR, Yeates TO. Analysis of lattice-translocation disorder in the layered hexagonal structure of carboxysome shell protein CsoS1C. Acta Crystallogr D Biol Crystallogr. 2009; 65:980–8. [PubMed: 19690376]
50. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. In: Carter, CW., Jr; Sweet, RM., editors. Methods in Enzymology 276, Macromolecular Crystallography, part A. Academic Press; New York: 1997. p. 307-326.
51. Kabsch W. XDS. Acta Cryst D. 2010; 66:125–132. [PubMed: 20124692]
52. McCoy AJ, et al. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
53. Crowley CS, et al. Structural insight into the mechanisms of transport across the Salmonella enterica Pdu microcompartment shell. J Biol Chem. 2010; 285:37838–46. [PubMed: 20870711]
54. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–55. [PubMed: 15299926]
55. Blanc E, et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallogr D Biol Crystallogr. 2004; 60:2210–21. [PubMed: 15572774]
56. Winn MD, Murshudov GN, Papiz MZ. Macromolecular TLS refinement in REFMAC at moderate resolutions. Meth Enzymol. 2003; 374:300–21. [PubMed: 14696379]

57. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010; 66:486–501. [PubMed: 20383002]

58. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst. 1993; 26:283–291.

59. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993; 2:1511–9. [PubMed: 8401235]

60. Lüthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature. 1992; 356:83–5. [PubMed: 1538787]

61. The PyMOL Molecular Graphics System, Version 1.4. Schrödinger, LLC; (http://www.pymol.org)

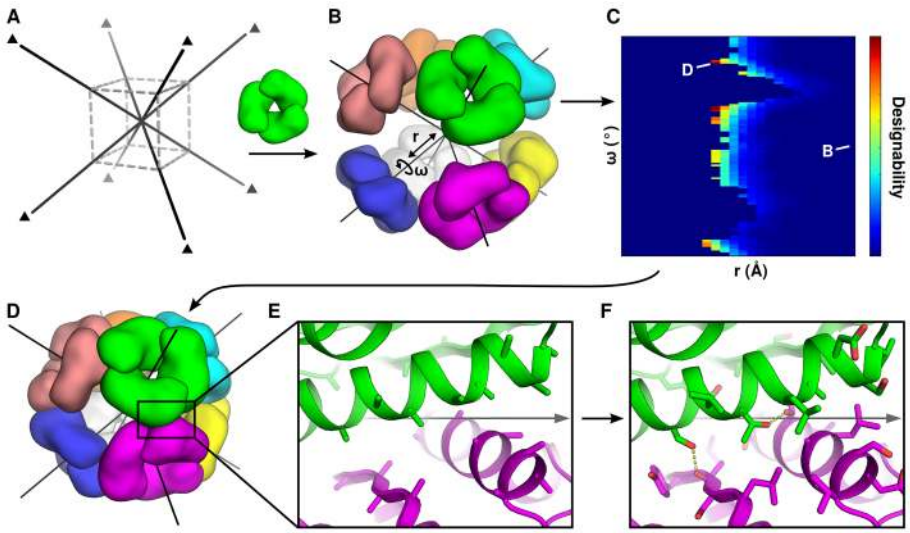

Fig. 1.

General approach to designing self-assembling protein nanomaterials. (**A**) First, a target symmetric architecture is chosen. Octahedral point group symmetry is used in this example; the threefold rotational axes are marked here by triangles and shown as black lines throughout. The dashed cube is shown to orient the viewer. A symmetric oligomer which shares an element of symmetry with the target architecture, here a C3 symmetric trimer (green), is selected as a building block. (**B**) Multiple copies of the building block are symmetrically arranged in the target architecture by aligning their shared symmetry axes. The pre-existing organization of the oligomeric building block fixes several (in this case four) rigid body degrees of freedom (DOFs). The two remaining DOFs, radial displacement (r) and axial rotation (ω), are indicated. (**C**) Symmetrical docking is performed by systematically varying the two DOFs (moves are applied symmetrically to all subunits) and computing the suitability of each configuration for interface design (red: more suitable; blue: less suitable). Points corresponding to the docked configurations in panels (B), in which the building blocks are not in contact, and (**D**), a highly complementary interface, are indicated. (**E**) Closer view of the interface in (D). The interface lies on an octahedral two-fold symmetry axis shown as a grey line. In all steps before interface design, only backbone (shown in cartoon) and carbon beta (shown in sticks) atoms are considered. (**F**) Sequence design calculations are used to create low-energy protein-protein interfaces that drive self-assembly of the desired material. Designed hydrogen bonds across the interface are indicated by dashed lines.

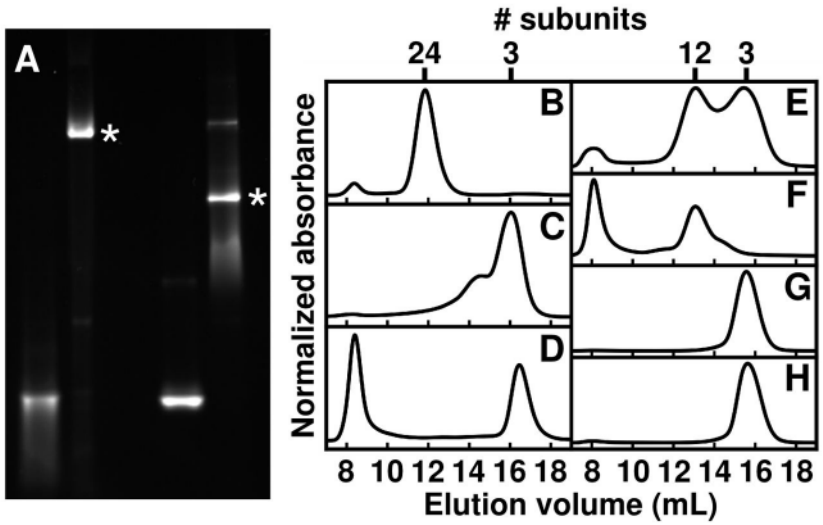

Fig. 2.

Experimental characterization of O3-33, T3-08, and T3-10. (**A**) Native PAGE of fluorescently labeled (from left) 3n79-wt, O3-33, 3ftt-wt, and T3-08 in lysates. Bands corresponding to the designed octahedral (O3-33) and tetrahedral (T3-08) assemblies are indicated with asterisks. SEC chromatgrams of nickel-purified (**B**) O3-33, (**C**) 3n79-wt, (**D**) O3-33(Ala167Arg), (**E**) T3-08, (**F**) T3-10, (**G**) 3ftt-wt, and (**H**) T3-08(Ala52Gln) collectively demonstrate that the assembly of the designed proteins is a result of the designed interfaces.

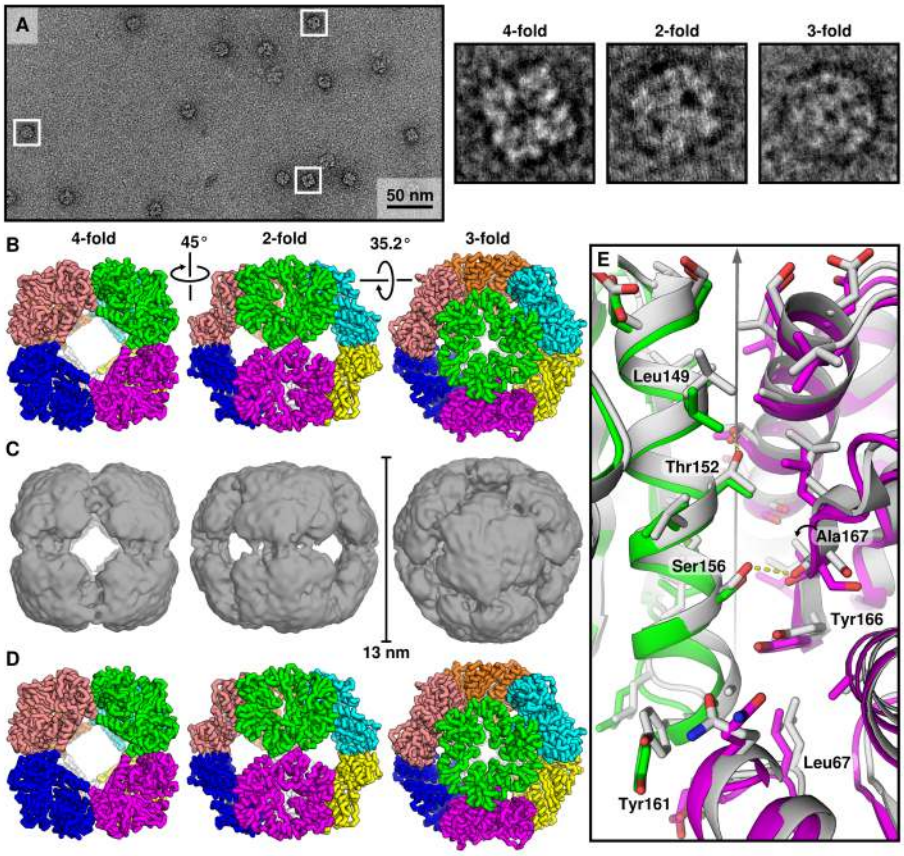

Fig. 3.

Structural characterization of O3-33. (**A**) A representative negative stain electron micrograph of O3-33. Selected particles (boxed in white) that resemble views of the design model along its 4-fold, 2-fold, and 3-fold rotational axes, shown in (B), are enlarged at right. (**B**) The O3-33 design model, depicted in ribbon format. Each trimeric building block is shown in a different color. (**C**) The density map from a 20 Å resolution cryo-EM reconstruction of O3-33 clearly recapitulates the architecture of the design model. (**D**) The crystal structure of O3-33 (R32 crystal form). Images in (B) to (D) are shown to scale along the three types of symmetry axes present in point group O. (**E**) The designed interface in O3-33, highlighting the close agreement between the crystal structure (green and magenta) and the design model (white). Oxygen atoms are red; nitrogens, blue. Hydrogen bonds between the building blocks are shown as yellow dashes, and an octahedral 2-fold rotational axis that passes through the interface is shown as a gray line. Residues in which substitution disrupted self-assembly (see Fig. S4) are labeled.

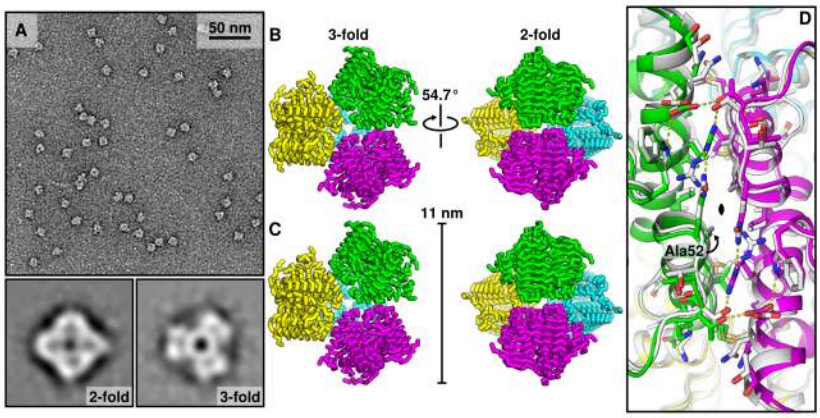

Fig. 4.

Structural characterization of T3-10. (**A**) A representative negative stain electron micrograph of T3-10. At bottom, averages of the particles resemble views of the design model along its 2-fold and 3-fold rotational axes, shown in (B). (**B**) Backbone representation T3-08/T3-10 design model, depicted as in Fig. 3B. (**C**) The T3-10 crystal structure. Images in (B) and (C) are shown to scale along the two types of symmetry axes present in point group T. (**D**) The designed interface in T3-10, revealing the close agreement of the crystal structure (green and magenta) to the design model (white). A network of polar interactions observed in the crystal structure at the designed interface is indicated by yellow dashes. The interface is viewed along an indicated tetrahedral 2-fold rotational axis. Alanine 52, which when mutated to glutamine in T3-08 disrupts assembly of the designed material, is labeled.