

9. A. Botev *et al.*, *Biochemistry* **50**, 828 (2011).
10. K. D. Nadezhdin, O. V. Bocharova, E. V. Bocharov, A. S. Arseniev, *Acta Naturae* **2010**, 69 (2011).
11. A. J. Beel *et al.*, *Biochemistry* **47**, 9428 (2008).
12. T. Sato *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1421 (2009).
13. P. Kienlen-Campard *et al.*, *J. Biol. Chem.* **283**, 7733 (2008).
14. L. M. Munter *et al.*, *EMBO J.* **26**, 1702 (2007).
15. N. Miyashita, J. E. Straub, D. Thirumalai, *J. Am. Chem. Soc.* **131**, 17843 (2009).
16. P. Osenkowski *et al.*, *J. Mol. Biol.* **385**, 642 (2009).
17. F. Renzi *et al.*, *J. Biol. Chem.* **286**, 21440 (2011).
18. S. Urban, *Biochem. J.* **425**, 501 (2010).
19. M. S. Wolfe, *Chem. Rev.* **109**, 1599 (2009).
20. J. Hardy, R. Crook, Alzheimer Research Forum, [www.alzforum.org/res/com/mut/app/](http://www.alzforum.org/res/com/mut/app/) (2012).
21. G. van Meer, D. R. Voelker, G. W. Feigenson, *Nat. Rev. Mol. Cell Biol.* **9**, 112 (2008).
22. R. M. Epan, *Biochim. Biophys. Acta* **1778**, 1576 (2008).
23. M. A. Lemmon *et al.*, *J. Biol. Chem.* **267**, 7683 (1992).
24. K. R. MacKenzie, J. H. Prestegard, D. M. Engelman, *Science* **276**, 131 (1997).
25. S. Kim *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 14278 (2005).
26. T. L. Kukar *et al.*, *Nature* **453**, 925 (2008).
27. S. A. Sagi *et al.*, *J. Biol. Chem.* **286**, 39794 (2011).
28. R. Ehehalt, P. Keller, C. Haass, C. Thiele, K. Simons, *J. Cell Biol.* **160**, 113 (2003).
29. S. J. Lee *et al.*, *Nat. Med.* **4**, 730 (1998).
30. P. Osenkowski, W. Ye, R. Wang, M. S. Wolfe, D. J. Selkoe, *J. Biol. Chem.* **283**, 22529 (2008).
31. S. Bodovitz, W. L. Klein, *J. Biol. Chem.* **271**, 4436 (1996).
32. E. Kojro, G. Gimpl, S. Lammich, W. Marz, F. Fahrenholz, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5815 (2001).
33. K. Yanagisawa, *Subcell. Biochem.* **38**, 179 (2005).

**Acknowledgments:** This work was supported by Alzheimer's Association grant IIRG-07-59379, NIH grants P01 GM080513 (C.R.S. and E.J.H.) and T32 GM08320 (P.J.B.), and by an American Heart Association grant AHA10POST4210008 (W.D.V.H.). We thank S. Howell for technical assistance and F. Sönnichsen, K. Mittendorf, B. Carter, E. Stanley, and C. Deatherage for comments on this manuscript. Coordinates for the structure of C99 have been deposited in the Protein Data Bank (accession code 2LP1).

#### Supplementary Materials

[www.sciencemag.org/cgi/content/full/336/6085/1168/DC1](http://www.sciencemag.org/cgi/content/full/336/6085/1168/DC1)  
Materials and Methods  
Figs. S1 to S9  
Table S1  
References (34–44)

2 February 2012; accepted 13 April 2012  
10.1126/science.1219988

# Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy

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

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. We used trimeric protein building blocks to design a 24-subunit, 13-nm diameter complex with octahedral symmetry and a 12-subunit, 11-nm diameter complex with tetrahedral symmetry. The designed proteins assembled to the desired oligomeric states in solution, and the 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 pat-

terns (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 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 nonbiological materials in conjunction with

computational interface design (17), or genetic fusion of multiple protein domains or fragments that 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 because of 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 that consists of two steps: (i) symmetrical docking of protein building blocks in a target symmetric architecture, followed by (ii) 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, because 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

<sup>1</sup>Department of Biochemistry, University of Washington, Seattle, WA 98195, USA. <sup>2</sup>Howard Hughes Medical Institute, University of California at Los Angeles, and U.S. Department of Energy (UCLA-DOE) Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA. <sup>3</sup>Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA. <sup>4</sup>Department of Medicinal Chemistry, University of Washington, Seattle, WA 98177, USA. <sup>5</sup>Department of Biochemistry and Structural Biology, Lund University, SE-221 00 Lund, Sweden. <sup>6</sup>Department of Chemistry and Biochemistry, University of California at Los Angeles, Los Angeles, CA 90095, USA. <sup>7</sup>UCLA-DOE Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA. <sup>8</sup>Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA.

\*To whom correspondence should be addressed. E-mail: [dabaker@uw.edu](mailto:dabaker@uw.edu)

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), which allows 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. In this study, 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). Eight T and 33 O designs derived from 15 distinct natural trimeric proteins, containing on average nine mutations per monomer, were selected for experimental characterization (table S1).

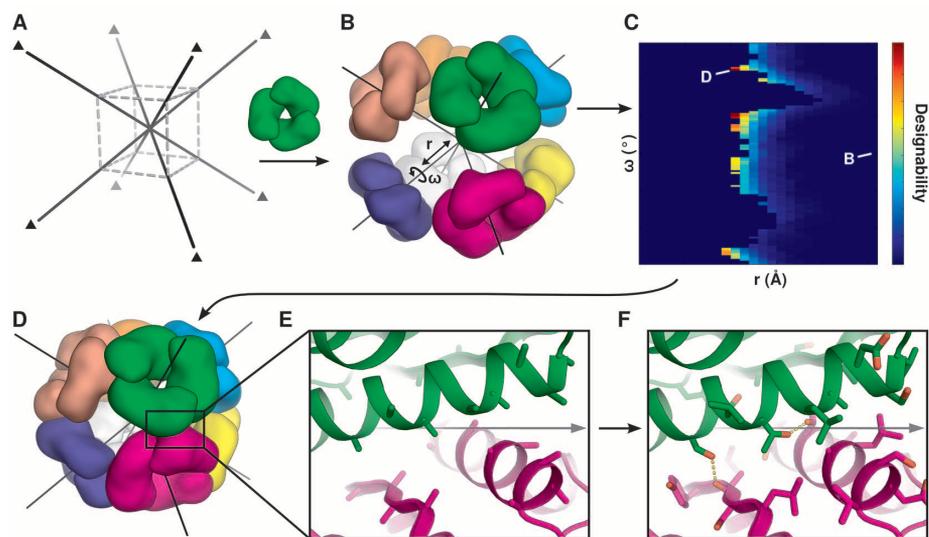
Genes encoding the designed proteins and the corresponding wild-type trimers were constructed and cloned into an expression vector that appended an 11-residue peptide substrate for fluorescent modification by the *Escherichia coli* acyl-carrier protein synthase AcpS (33). *E. coli* cells expressing the 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 polyacrylamide gel electrophoresis (PAGE) under nonreducing (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”;

nine mutations from the wild-type protein), the tetrahedral design (“T3-08”; eight mutations), and the corresponding wild-type trimeric proteins were then subcloned into an expression vector that appended C-terminal (His)<sub>6</sub> 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 [Protein Data Bank (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, which suggests 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 two-fold, three-fold, or four-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, which verified that the protein assembles in solution as designed (Fig. 3C and fig. 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 root mean square deviation (RMSD) over all 24 chains is 1.07 Å and is lower if calculated by using only the residues at the interface (0.85 Å). The high resolution of the structure allowed

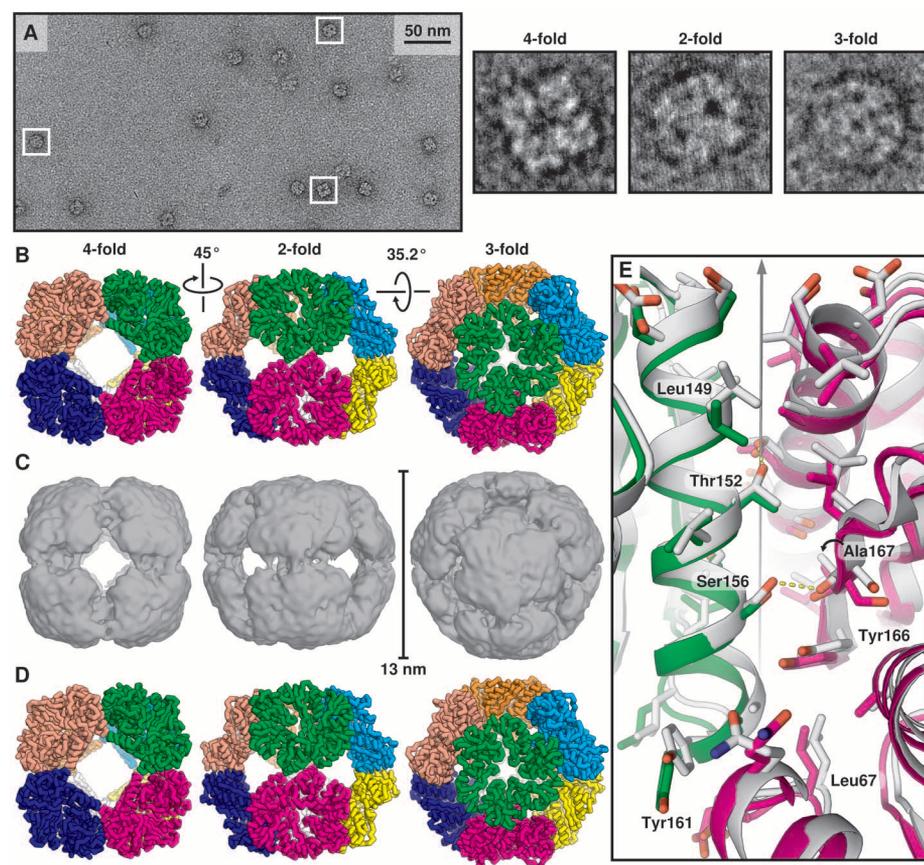
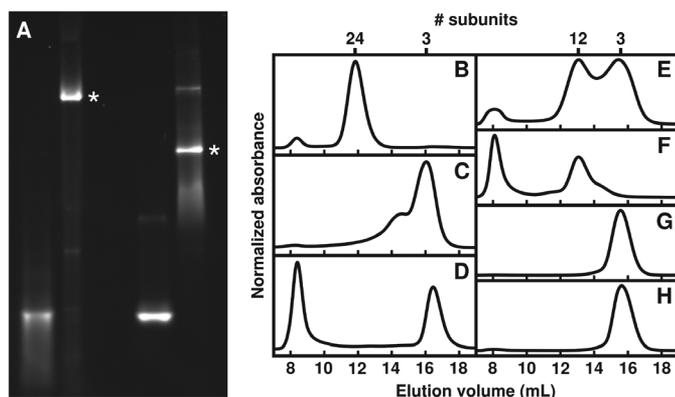


**Fig. 1.** General approach to designing self-assembling protein nanomaterials. (A) First, a target symmetric architecture is chosen. Octahedral point group symmetry is used in this example; the three-fold rotational axes are marked here by triangles and shown as black lines throughout. The dashed cube 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 preexisting 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 ( $\omega$ ), 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 (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 gray 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.

confident determination of the side-chain configurations at the designed interface, which revealed 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

**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 chromatograms 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.



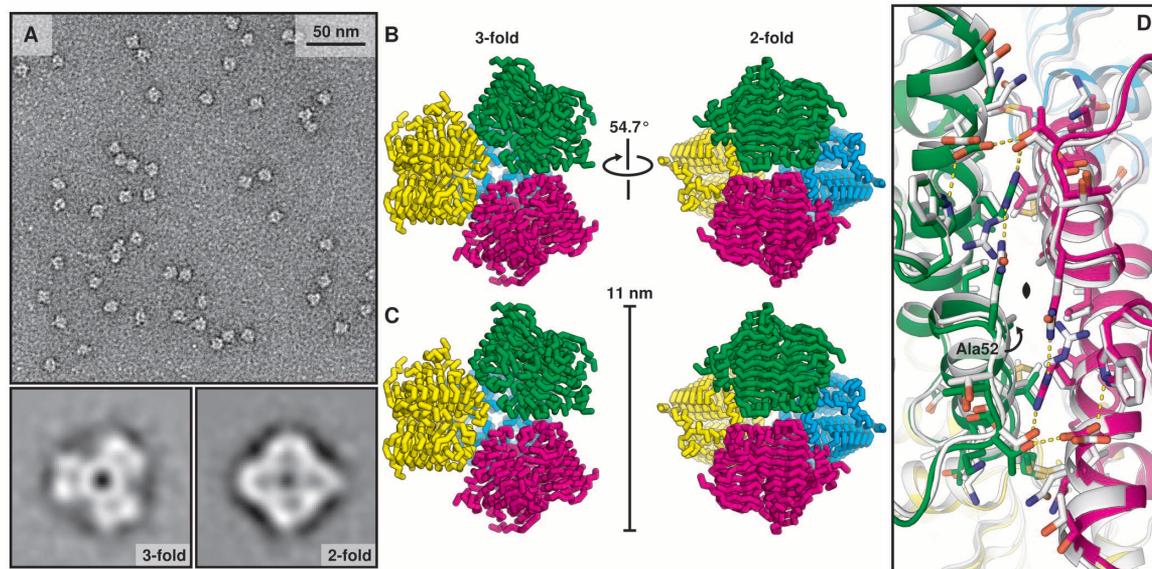
**Fig. 3.** Structural characterization of O3-33. **(A)** A representative negative-stain electron micrograph of O3-33. Selected particles (boxed in white) that resemble views of the design model along its four-fold, two-fold, and three-fold rotational axes, shown in **(B)**, are enlarged at right. **(B)** 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 two-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.

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 2827 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 3 and ~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, which subtly alters the atomic contacts at the designed interface relative to the design model and results in a backbone RMSD of 2.66 Å over all 12 subunits (fig. S5).

We designed two additional variants of T3-08 (table S1) 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 three 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 tetrahedral 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 two-fold and three-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.

**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 two-fold and three-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 two-fold rotational axis. Alanine 52 is labeled; when mutated to glutamine in T3-08, it disrupts assembly of the designed 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, cagelike 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.

#### References and Notes

- N. C. Seeman, *Annu. Rev. Biochem.* **79**, 65 (2010).
- P. W. Rothmund, *Nature* **440**, 297 (2006).
- E. S. Andersen *et al.*, *Nature* **459**, 73 (2009).
- J. Zheng *et al.*, *Nature* **461**, 74 (2009).
- S. Zhang, *Nat. Biotechnol.* **21**, 1171 (2003).
- T. Douglas, M. Young, *Science* **312**, 873 (2006).
- B. Lovejoy *et al.*, *Science* **259**, 1288 (1993).
- P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, *Science* **282**, 1462 (1998).
- C. Gribbon *et al.*, *Biochemistry* **47**, 10365 (2008).
- N. R. Zaccai *et al.*, *Nat. Chem. Biol.* **7**, 935 (2011).
- R. L. Koder *et al.*, *Nature* **458**, 305 (2009).
- E. R. Ballister, A. H. Lai, R. N. Zuckermann, Y. Cheng, J. D. Mougous, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 3733 (2008).
- K. Usui *et al.*, *Protein Sci.* **18**, 960 (2009).
- P. Ringle, G. E. Schulz, *Science* **302**, 106 (2003).
- E. N. Salgado, J. Faraone-Mennella, F. A. Tezcan, *J. Am. Chem. Soc.* **129**, 13374 (2007).
- E. N. Salgado, R. J. Radford, F. A. Tezcan, *Acc. Chem. Res.* **43**, 661 (2010).
- G. Grigoryan *et al.*, *Science* **332**, 1071 (2011).
- J. E. Padilla, C. Colovos, T. O. Yeates, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2217 (2001).
- J. C. Sinclair, K. M. Davies, C. Vénien-Bryan, M. E. Noble, *Nat. Nanotechnol.* **6**, 558 (2011).
- J. Janin, R. P. Bahadur, P. Chakrabarti, *Q. Rev. Biophys.* **41**, 133 (2008).
- D. Grueninger *et al.*, *Science* **319**, 206 (2008).
- T. Kortemme *et al.*, *Nat. Struct. Mol. Biol.* **11**, 371 (2004).
- P. S. Huang, J. J. Love, S. L. Mayo, *Protein Sci.* **16**, 2770 (2007).
- R. K. Jha *et al.*, *J. Mol. Biol.* **400**, 257 (2010).
- J. Karanikolas *et al.*, *Mol. Cell* **42**, 250 (2011).
- P. B. Stranges, M. Machius, M. J. Miley, A. Tripathy, B. Kuhlman, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20562 (2011).
- S. J. Fleishman *et al.*, *Science* **332**, 816 (2011).
- S. J. Fleishman *et al.*, *J. Mol. Biol.* **413**, 1047 (2011).
- Materials and methods are available as supplementary materials on Science Online.
- B. Kuhlman, D. Baker, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10383 (2000).
- F. DiMaio, A. Leaver-Fay, P. Bradley, D. Baker, I. André, *PLoS ONE* **6**, e20450 (2011).
- S. Cooper *et al.*, *Nature* **466**, 756 (2010).
- Z. Zhou *et al.*, *ACS Chem. Biol.* **2**, 337 (2007).

**Acknowledgments:** We thank J. Navarro for assistance with protein crystallization at the UCLA crystallization core facility, which is supported by DOE Biological and Environmental Research (DOE-BER) grant DE-FC03-02ER63421, and D. Cascio and the 24-ID-C beamline staff for their assistance in data collection. This work is based on 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, NIH. Use of the Advanced Photon Source, operated for the Office of Science, DOE, by Argonne National Laboratory, was supported by the DOE under contract no. DE-AC02-06CH11357. Analytical ultracentrifugation 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 also thank M. Iadanza for EM analysis of T3-10; Y. Cheng and X. Li (UCSF) for giving us access to their electron cryomicroscope for data collection, for helpful discussions, and for sharing scripts with us; and N. Grigorieff (Brandeis) for helpful discussions. The Gonen laboratory is supported by Howard Hughes Medical Institute (HHMI). Work by N.P.K., W.S., and D.B. 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). N.P.K., W.S., T.O.Y., and D.B. have filed a provisional patent application, U.S. 61/622,889, on the described method for designing self-assembling protein materials.

#### Supplementary Materials

www.sciencemag.org/cgi/content/full/336/6085/1171/DC1  
Materials and Methods  
Figs. S1 to S6  
Tables S1 and S2  
References (34–61)

19 January 2012; accepted 13 April 2012  
10.1126/science.1219364