

published DEER study on GDP-bound  $G\alpha_1$  in the absence of the  $\beta\gamma$  subunit also suggest a minority population with separated domains (23).

Our simulations suggest that the minority domain-separated population in GDP-bound G proteins arises due to rapid fluctuations between closed and open conformations and that this spontaneous opening plays an essential role in nucleotide exchange. This implies that constraining domain opening would substantially slow basal nucleotide exchange and, in particular, GDP release. To test this prediction, we engineered a  $G_i$  variant to restrict domain opening. In this construct, the N terminus of the  $\gamma$  subunit was fused to a peptide fragment designed to bind the helical domain without impinging on either the nucleotide-binding site or the Ras domain (fig. S16). Binding kinetics measured by fluorescence quenching showed that this helical domain tether slowed basal nucleotide exchange by a factor of 20, under conditions in which GDP release is rate-limiting (Fig. 4B).

Our nucleotide-release mechanism is consistent with earlier mutagenesis studies. Point mutations to the Ras domain  $\beta 6$ - $\alpha 5$  loop (24) accelerate nucleotide exchange in the absence of a receptor substantially more than mutations that weaken contacts between the Ras and helical domains (25), suggesting that weakening interactions between  $\beta 6$ - $\alpha 5$  and the GDP guanine ring facilitates nucleotide release to a greater extent than does increasing domain separation. Mutations to  $\alpha 5$  that energetically favor the distal conformation increase both receptor-catalyzed and basal nucleotide-exchange rates, whereas those disfavoring that conformation decrease nucleotide-exchange rates (21) (fig. S10D).

Several caveats are in order. First, because we did not simulate the complete process of receptor-G protein association, we have not determined the sequence of steps by which a receptor couples to a G protein, nor have we addressed the question of whether a G protein might associate with a receptor before receptor activation (26–28). Second, although our simulations are orders of magnitude longer than previous atomistic G protein simulations, they still lack sufficient length, and perhaps sufficient accuracy, to reliably determine equilibrium populations of the various conformations. However, our simulations strongly imply the existence of certain conformations and dynamical interchange among them. We cannot rule out the possibility that additional conformational changes to the G protein would manifest themselves on longer time scales. Thus, the GPCR might also induce GDP release, in part through other mechanisms, such as displacement of the  $\beta 1$  strand of  $G\alpha$  (7). Third, because crystal structures of nucleotide-bound and receptor-bound heterotrimers are not available for the same G protein, our analysis combines data from different G proteins, under the common assumption that their high level of structural homology implies similar functional mechanisms (1, 2).

Why might heterotrimeric G proteins have evolved to fluctuate spontaneously between open

and closed conformations? Tight apposition of the Ras and helical domains appears to be essential for efficient hydrolysis of GTP to GDP (29). In the closed conformation, the helical domain plays a role similar to that of the GTPase activating proteins (GAPs) required by small G proteins—which contain only a Ras domain—for efficient catalysis (18). Conversely, our results suggest that rapid GDP release requires an open conformation. Spontaneous fluctuation of the helical domain position thus provides an elegant solution to the conflicting needs of catalysis and nucleotide release.

#### REFERENCES AND NOTES

1. W. M. Oldham, H. E. Hamm, *Nat. Rev. Mol. Cell Biol.* **9**, 60–71 (2008).
2. S. R. Sprang, *Annu. Rev. Biochem.* **66**, 639–678 (1997).
3. D. M. Rosenbaum, S. G. F. Rasmussen, B. K. Kobilka, *Nature* **459**, 356–363 (2009).
4. S. G. F. Rasmussen *et al.*, *Nature* **477**, 549–555 (2011).
5. A. J. Venkatakrishnan *et al.*, *Nature* **494**, 185–194 (2013).
6. N. Van Eps *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9420–9424 (2011).
7. K. Y. Chung *et al.*, *Nature* **477**, 611–615 (2011).
8. G. H. Westfield *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16086–16091 (2011).
9. H. G. Dohlman, J. C. Jones, *Sci. Signal.* **5**, re2 (2012).
10. M. A. Wall *et al.*, *Cell* **83**, 1047–1058 (1995).
11. D. G. Lambright *et al.*, *Nature* **379**, 311–319 (1996).
12. Materials and methods are available as supplementary materials on Science Online.
13. L. V. Mello, D. M. van Aalten, J. B. Findlay, *Biochemistry* **37**, 3137–3142 (1998).
14. M. A. Ceruso, X. Periole, H. Weinstein, *J. Mol. Biol.* **338**, 469–481 (2004).
15. K. Khafizov, G. Lattanzi, P. Carloni, *Proteins* **75**, 919–930 (2009).
16. M. Louet, D. Perahia, J. Martinez, N. Floquet, *J. Mol. Biol.* **411**, 298–312 (2011).
17. J. C. Jones, A. M. Jones, B. R. Temple, H. G. Dohlman, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7275–7279 (2012).
18. D. W. Markby, R. Onrust, H. R. Bourne, *Science* **262**, 1895–1901 (1993).
19. N. S. Alexander *et al.*, *Nat. Struct. Mol. Biol.* **21**, 56–63 (2014).
20. R. Onrust *et al.*, *Science* **275**, 381–384 (1997).
21. E. P. Marin, A. G. Krishna, T. P. Sakmar, *J. Biol. Chem.* **276**, 27400–27405 (2001).

22. W. M. Oldham, N. Van Eps, A. M. Preininger, W. L. Hubbell, H. E. Hamm, *Nat. Struct. Mol. Biol.* **13**, 772–777 (2006).
23. N. Van Eps, C. J. Thomas, W. L. Hubbell, S. R. Sprang, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 1404–1409 (2015).
24. T. Iiri, P. Herzmark, J. M. Nakamoto, C. van Dop, H. R. Bourne, *Nature* **371**, 164–168 (1994).
25. E. P. Marin *et al.*, *J. Biol. Chem.* **276**, 23873–23880 (2001).
26. R. Herrmann, M. Heck, P. Henklein, K. P. Hofmann, O. P. Ernst, *J. Biol. Chem.* **281**, 30234–30241 (2006).
27. P. Scheerer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10660–10665 (2009).
28. M. Elgeti *et al.*, *J. Am. Chem. Soc.* **135**, 12305–12312 (2013).
29. T. Shnerb, N. Lin, A. Shurki, *Biochemistry* **46**, 10875–10885 (2007).

#### ACKNOWLEDGMENTS

We thank A. Pan for assistance with TAMM, K. Palmo for help with force field parameterization, J. Valcourt and H. Green for advice on figures, and M. Kirk for editorial assistance. Portions of this work were funded by NIH grants R01GM083118 (B.K.K. and R.K.S.) and R01EY05216 (W.L.H.), the Jules Stein Professor Endowment (W.L.H.), National Eye Institute Core Grant P30EY00331 (W.L.H.), and a Research to Prevent Blindness unrestricted grant (to the Jules Stein Eye Institute). D.H. was supported by the German Academic Exchange Service (DAAD). Author contributions: R.O.D. conceived this study and, with D.E.S., oversaw MD simulations and analysis. R.O.D., T.J.M., and D.H.A. designed MD simulations. T.J.M. performed MD simulations and, with D.W.B., modeled proteins based on crystal structures. T.J.M. analyzed MD simulations, with guidance from R.O.D. T.J.M. and A.P. performed molecular visualization. D.H. developed a minimal cysteine construct for  $G_s$ , prepared spin-labeled protein, and assisted Z.Y. and M.T.L. in obtaining and analyzing DEER data. A.M. and D.H. designed the HD-tether construct, and A.M. performed the kinetics experiments. W.L.H. and B.K.K. supervised experimental work. R.O.D., T.J.M., N.V., and R.K.S. interpreted and synthesized data. R.O.D., T.J.M., D.W.B., and D.E.S. wrote the manuscript, with input from D.H., A.M., B.K.K., and R.K.S. R.O.D. and D.E.S. supervised the overall research.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6241/1361/suppl/DC1  
Materials and Methods  
Figs. S1 to S17  
Table S1  
References (30–58)

18 December 2014; accepted 15 May 2015  
10.1126/science.aaa5264

#### PROTEIN DESIGN

## Design of ordered two-dimensional arrays mediated by noncovalent protein-protein interfaces

Shane Gonen,<sup>1,2,3,4</sup> Frank DiMaio,<sup>2,3</sup> Tamir Gonen,<sup>1\*</sup> David Baker<sup>2,3,4,\*</sup>

We describe a general approach to designing two-dimensional (2D) protein arrays mediated by noncovalent protein-protein interfaces. Protein homo-oligomers are placed into one of the seventeen 2D layer groups, the degrees of freedom of the lattice are sampled to identify configurations with shape-complementary interacting surfaces, and the interaction energy is minimized using sequence design calculations. We used the method to design proteins that self-assemble into layer groups P 3 2 1, P 4 2<sub>1</sub> 2, and P 6. Projection maps of micrometer-scale arrays, assembled both in vitro and in vivo, are consistent with the design models and display the target layer group symmetry. Such programmable 2D protein lattices should enable new approaches to structure determination, sensing, and nanomaterial engineering.

Programmed self-assembly provides a route to patterning matter at the atomic scale. DNA origami methods (1, 2) have been used to generate a wide variety of ordered structures, but progress in designing pro-

tein assemblies has been slower owing to the greater complexity of protein-protein interactions. Biology provides a number of examples of ordered two-dimensional (2D) protein arrays: Bacterial S-layer proteins assemble into

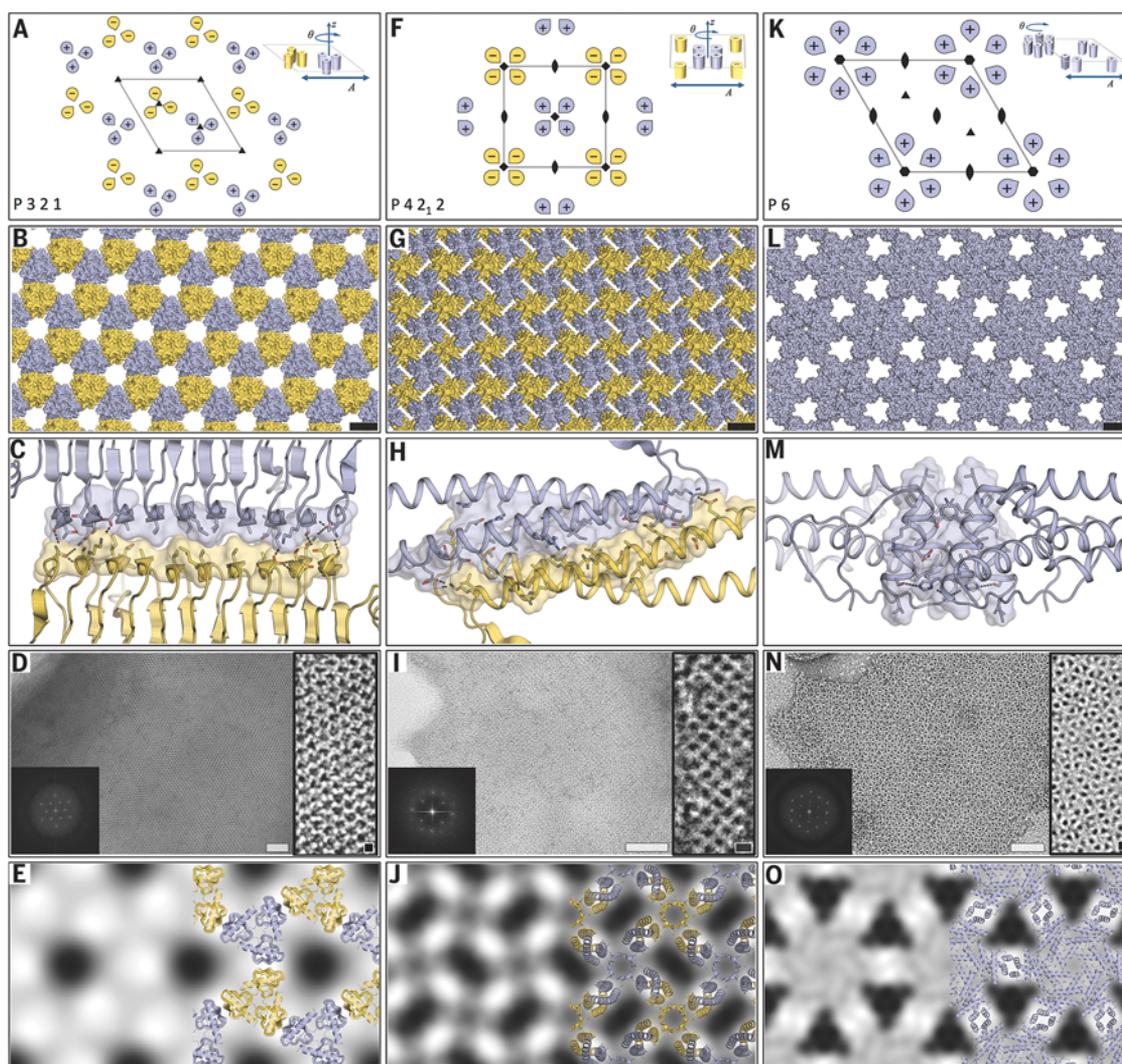
oblique, square, or hexagonal planar symmetry (3); gap-junction plaques, abundant in muscle and heart tissue, display hexagonal planar symmetry (4); and water channels display square planar symmetry (5). Although proteins that form ordered 3D crystals have been designed (6) and

2D lattices have been generated by genetically fusing or chemically cross-linking oligomers with appropriate point symmetric groups (7–10), there has been little success in designing self-assembling 2D lattices with order sufficient to diffract electrons or x-rays below 15 Å resolution (7). Naturally occurring 2D arrays and assemblies are stabilized by extensive noncovalent interactions between protein subunits (10, 11), and this principle has been used to design self-assembling tetrahedral and octahedral cages (12, 13).

We sought to design ordered 2D arrays mediated by designed protein-protein interfaces sta-

bilized by extensive noncovalent interactions. We focused on symmetric arrays, as symmetry reduces the number of distinct protein interfaces required to stabilize the lattice (14, 15). There are 17 distinct ways (layer groups) in which 3D objects can come together to form periodic 2D layers (16). In some layer groups, there are only two unique interfaces between identical subunits, in others, three or four (17). To simplify the design challenge, we focused on the layer groups that involve only two unique interfaces and building blocks with internal point symmetry (which already contain one of the two required interfaces), which leaves only one unique interface to be

<sup>1</sup>Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA. <sup>2</sup>Department of Biochemistry, University of Washington, Seattle, WA 98195, USA. <sup>3</sup>Institute for Protein Design, University of Washington, Seattle, WA 98195, USA. <sup>4</sup>Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA.  
\*Corresponding author. E-mail: gonent@janelia.hhmi.org (T.G.); dabaker@uw.edu (D.B.)



**Fig. 1. Computational design strategy and experimental analysis of designed arrays.** (A) The P 3 2 1 unit cell with threefold axes represented by triangles. Yellow (–) and purple (+) C3 objects have opposite orientations along the z axis. (Inset) The three degrees of freedom of the lattice. (B) p3Z\_42 2D array. (C) p3Z\_42 designed interface with “zipper-like” hydrophobic packing and peripheral hydrogen bonds. (D) Large (>1 μm) *E. coli*-grown array (middle), higher magnification view with lattice spacing as in (B) (right), and Fourier transform (amplitudes) of the large array (left). (E) (left) Projection map at 15 Å calculated from a large array. (Right) overlay of the p3Z\_42 design model on the projection map. (F) The P 4 2 1 2 lattice. Ovals represent twofold axes

and squares, fourfold axes. (G) p4Z\_9 array. (H) p4Z\_9 designed interface. (I) Negatively stained *E. coli*-grown array (main panel), an in vitro refolded lattice at higher magnification (inset), and Fourier transform of the main panel (left). (J) Projection map at 14 Å calculated from an *E. coli* array as in (I) without (left) and with (right) p4Z\_9 design model. (K) The P 6 lattice has two degrees of freedom (A, θ) (inset) available for sampling. Sixfolds are represented by hexagons. (L) p6\_9H array. (M) p6\_9H designed interface. (N) p6\_9H lattice grown in vivo with Fourier transform at left and higher magnification view at right. (O) Projection map at 14 Å of p6\_9H from *E. coli*-grown arrays as in (N) and cartoon overlay (right). All scale bars: black, 5 nm; white, 50 nm.

designed to form the 2D array. Of the 17 layer groups, 11 have two unique interfaces; we focused here on 6 of these 11 groups involving cyclic rather than dihedral point groups because there are considerably more cyclic oligomers than dihedral oligomers in the Protein Data Bank that can serve as building blocks. The six layer groups with two unique interfaces that can be built from cyclic oligomers are P 2 2<sub>1</sub> 2<sub>1</sub> (from C2 building blocks), P 3 and P 3 2 1 (from C3 building blocks), P 4 and P 4 2<sub>1</sub> 2 (from C4 building blocks), and P 6 (from C6 building blocks). The different groups have different numbers of degrees of freedom describing the placement of an object with cyclic symmetry in the lattice, for example, for P 3 2 1 (Fig. 1A) and P 4 2<sub>1</sub> 2 (Fig. 1F), there are three degrees of freedom, whereas for P 6 (Fig. 1K) there are only two.

We used symmetric docking in Rosetta (14, 18, 19) to search for placements of cyclic oligomers into each of the six layer groups with shape-complementary (20) interfaces between different oligomer copies. The docking scoring function consisted of a soft sphere model of steric interactions and a simple measure of the designable interface area: the number of interface C $\beta$ s within 7 Å. For each cyclic oligomer in each layer group, ~20 independent Monte Carlo docking trajectories were carried out that started from placements of six to nine copies of the oligomer with its symmetry axis aligned with the corresponding symmetry axes of the layer group (for example, trimers were placed on the threefold symmetry axes indicated by the triangles in Fig. 1A, tetramers on the fourfold symmetry axes indicated by squares in Fig. 1F, and hexamers on the sixfold symmetry axes indicated by hexagons in Fig. 1K). In the Monte Carlo docking simulations, the degrees of freedom sampled were those compatible with the layer group

[Fig. 1, A, F, and K (right)], and hence, the layer group symmetry was preserved throughout the calculations.

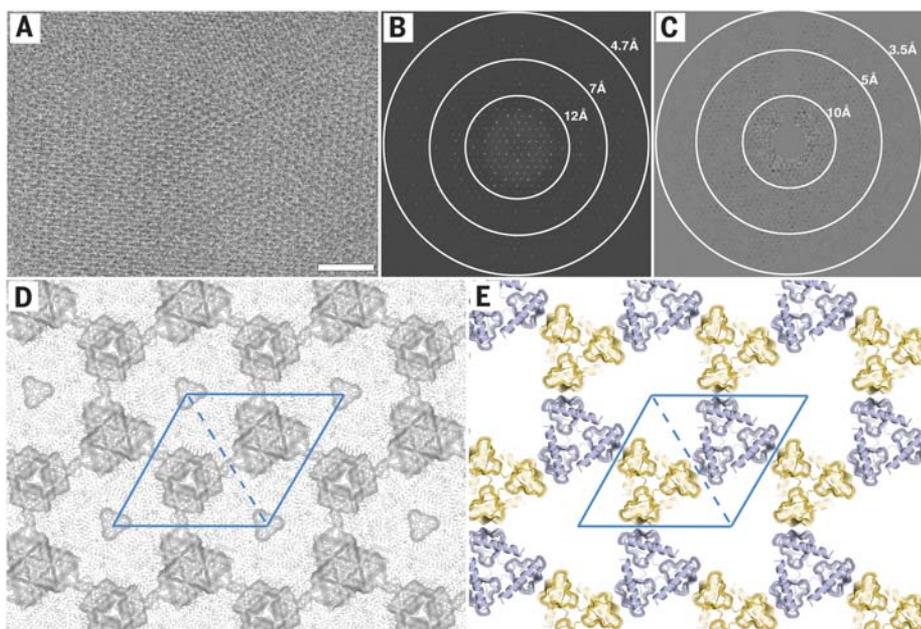
We then selected the most shape-complementary (largest number of contacting residues with fewest clashes) solutions from the trajectories and carried out Rosetta sequence design calculations to generate well-packed low-energy interfaces between oligomers. Monte Carlo searches were carried out over all amino acid identities and side-chain rotamer states for residues near the newly formed interface between oligomers, while optimizing the Rosetta all-atom energy of the entire complex (22, 13, 21). After this sequence design step, the energy was further minimized with respect to the side-chain torsion angles of residues near the interface and the symmetric degrees of freedom of the layer group. Finally, the resulting lattice models were filtered on the basis of the shape-complementarity of the designed interface (>0.5), surface area of the designed interface (>400 Å per monomer), buried unsatisfied hydrogen bonds introduced at the new interface (<4 using a 1.4 Å solvent accessibility probe) (22), and predicted relative free energy (23) of complex formation ( $\leq 10$  Rosetta energy units per subunit) (sample Rosetta script files accompany the supplementary material). After further sequence optimization (13, 24), models passing the filters were manually inspected, and 62 designs were selected for experimental characterization; 16 for P 2 2<sub>1</sub> 2<sub>1</sub>, 2 for P 3, 10 for P 3 2 1, 16 for P 4, 3 for P 4 2<sub>1</sub> 2, and 15 for P 6.

Synthetic genes were obtained for the 62 designs, and the proteins were expressed in the *Escherichia coli* cytoplasm by using a standard T7-based expression vector. Of the 62 designs, 43 expressed; of these, 18 had protein in the supernatant after clearing the lysate at

12,000g for 30 min, whereas all 43 had protein in the pellet. To investigate the degree of order in the pelleted material, we examined negatively stained samples by electron microscopy (EM). Regular lattices were observed for four of the designs: One formed only stacked 2D layers (fig. S1), whereas three formed planar arrays. The latter are described in the following sections.

Design p3Z\_42 is in layer group P 3 2 1. The rigid body arrangement of the constituent  $\beta$ -helix trimers in the lattice was identified by Monte Carlo search over the three degrees of freedom of the lattice: the rotation of the trimer around its axis  $\theta$ , the lattice spacing  $A$ , and the  $z$  offset of the trimer from the lattice plane (Fig. 1A). In the lattice identified in the Monte Carlo docking calculations, the oligomeric building blocks pack into a dense array (Fig. 1B; the yellow and purple copies are inverted with respect to each other) stabilized by a large contact surface between adjacent copies with close complementary side-chain packing (Fig. 1C) generated in the sequence design calculations.

p3Z\_42 formed large and very well ordered 2D crystals (Fig. 1D). Most of the protein expressed in *E. coli* appeared to assemble in these 2D crystals, as there was very little present in the soluble fraction (fig. S3). At low (16°C) expression temperatures, 2D sheets were obtained (Fig. 1D), whereas at 37°C, where larger amounts of proteins are produced, large 2D sheets stacked mainly into thick 3D crystals. Higher magnification (Fig. 1D, inset) showed a trigonal lattice similar to that of the design model [compare Fig. 1D (right) with Fig. 1B]. Fourier transformation of the lattice [Fig. 1D (left)] yielded peaks out to 15 Å resolution; the order in the unstained lattice is probably markedly higher, as the negative stain likely limits the observed resolution. A 15 Å



**Fig. 2. Cryo-EM analysis of design p3Z\_42.** (A) Cryo-EM micrograph of *E. coli*-grown p3Z\_42 recorded from nonpurified, resuspended insoluble material. (B) Fourier transform calculated from motion-corrected movies taken from samples like those in (A). (C) Electron diffraction of a crystal as in (A). (D) Projection map at 4 Å calculated from motion-corrected movies from material as in (A) showing a linked repeat-protein arrangement similar to the p3Z\_42 design model. The unit cell is shown in blue and contains two alternating trimeric units. Triangular density at the corners of the unit cell is likely an averaging artifact. (E) p3Z\_42 design model in a similar view as in (D). Scale bar, 50 nm.

projection map (Fig. 1E) back-computed from the Fourier components followed the contour of the designed lattice [Fig. 1E (right)] (unit cell dimensions  $a = b = 85 \text{ \AA}$ ,  $\gamma = 120^\circ$ ). It is notable that planar crystals of such large size can grow without support within the confines (and with the many cellular obstacles) of an *E. coli* cell. Cell-free expression of this design yielded large, ordered 2D crystals similar to those formed in *E. coli* (fig. S4A).

Design p4Z\_9 is in layer group P 4 2<sub>1</sub> 2. Search over the three degrees of freedom of the layer group [the rotation around the internal C<sub>4</sub> axis, the lattice spacing, and the  $z$  offset between adjacent inverted tetramers (Fig. 1F)] yielded the close-packed arrangement shown in Fig. 1G (side view in fig. S2B). The designed interface is composed of hydrophobic residues nestled between two  $\alpha$  helices surrounded by polar residues (Fig. 1H).

p4Z\_9 formed crystals up to 1  $\mu\text{m}$  in width (Fig. 1I) with little of the protein present in the soluble fraction (fig. S3). Incubation of the pellet material with 6 M guanidine and subsequent purification and refolding (by dialysis or fast dilution) yielded crystalline 2D arrays and fibers with the same square packing (fig. S4, B and C). Fourier transformation of the negatively stained large 2D lattices generated in vivo yielded peaks out to 14  $\text{\AA}$  resolution [Fig. 1I (left)]. The 14  $\text{\AA}$  projection map produced by back-transformation had distinctive rectangular voids in alternating directions, which closely matched the design model [Fig. 1J and 1J (right)] (unit cell dimensions  $a = b = 56 \text{ \AA}$ ,  $\gamma = 90^\circ$ ).

Design p6\_9 is built from  $\alpha$ -helical hexamers in layer group P 6. In this case, all oligomers are in the same orientation along the  $z$  axis (perpendicular to the plane in Fig. 1K), and hence, there are only two degrees of freedom—the rotation around the sixfold axis and the lattice spacing [Fig. 1K (right)]. The shape-complementary docking solution (Fig. 1L and side view fig. S2C) is composed of four closely associating  $\alpha$  helices along the twofold axis of the lattice (Fig. 1M) with two interacting phenylalanines. We also tested a variant, p6\_9H, which introduces a hydrogen bond network across the interface (Fig. 1M).

Design p6\_9 expressed in *E. coli* was found in both the supernatant and pellet (fig. S3). EM investigation revealed that the pellet contained highly ordered single-layer 2D hexagonal arrays, whereas the supernatant did not. p6\_9H formed even larger arrays (Fig. 1N, fig. S5, and table S1). The 2D layers in the pellet were highly ordered with clearly evident hexagonal packing [Fig. 1N and 1N (inset)]. Fourier transformation of the negatively stained arrays [Fig. 1N (left)] yielded peaks out to 14  $\text{\AA}$  resolution; and the back-computed 14  $\text{\AA}$  map was again closely consistent with the design model of the array [Fig. 1O and 1O (right)] (unit cell dimensions:  $a = b = 120 \text{ \AA}$ ,  $\gamma = 120^\circ$ ). Large arrays were also formed in vitro after concentration of soluble p6\_9H purified from the supernatant after lysis of *E. coli* (fig. S4, D and E).

To achieve higher resolution than possible with negatively stained samples, we analyzed designs without stain by electron cryomicroscopy (cryo-EM). Analysis of p3Z\_42 crystals by cryo-EM (Fig. 2, A and B) and electron diffraction yielded data to 3.5  $\text{\AA}$  resolution (Fig. 2C). The vast majority of crystals diffracted to this resolution in the cryo preparations, indicating high long-range order. Movie micrographs of the resulting crystals were also collected, motion corrected, and processed in *2dx* (25) to yield a projection map at 4  $\text{\AA}$  resolution in agreement with the design model (Fig. 2, compare D and E). To our knowledge, this is the highest order observed to date for a designed macromolecular 2D lattice.

Our designed planar protein arrays form large planar 2D crystals both in vivo and in vitro that are closely consistent with the design models. Two of the three successes were with layer groups with adjacent building blocks in opposite orientations along the  $z$  axis; these have the advantages that (i) there is an additional degree of freedom (the  $z$  offset) providing more possible packing arrangements for a given oligomeric building block; (ii) the interfaces are antiparallel rather than parallel so that, in the design calculations, opposing residues can have different identities; and (iii) inaccuracies in the design calculations that result in deviation from planarity effectively cancel out. On the other hand, designed “polar” arrays with all subunits oriented in the same direction—such as p6\_9—have advantages for functionalization, as the two sides are distinct and can be addressed separately.

It is notable that, for all three designs, extensive crystalline arrays form unsupported in *E. coli* and from purified protein in vitro. The coherent arrays can extend up to 1  $\mu\text{m}$  in length but are only 3 to 8 nm thick by design (fig. S2). We anticipate that even larger and perhaps more highly ordered crystals would form on a solid support, which will be useful for future nanotechnology applications. The ability to precisely design 2D arrays at the near atomic level should enable new approaches in structural biology [fusing proteins of unknown structure to array components for electron crystallography or using these to nucleate 3D crystal growth for x-ray and MicroED (26) applications], new sensing modalities with the coupling of analyte binding domains to the arrays, and the organization of enzyme networks and light-harvesting chromophores in two dimensions.

#### REFERENCES AND NOTES

1. N. C. Seeman, *Annu. Rev. Biochem.* **79**, 65–87 (2010).
2. P. W. Rothmund, *Nature* **440**, 297–302 (2006).
3. D. Pum, J. L. Toca-Herrera, U. B. Sleytr, *Int. J. Mol. Sci.* **14**, 2484–2501 (2013).
4. D. L. Caspar, D. A. Goodenough, L. Makowski, W. C. Phillips, *J. Cell Biol.* **74**, 605–628 (1977).
5. J. E. Rash, K. G. Davidson, T. Yasumura, C. S. Furman, *Neuroscience* **129**, 915–934 (2004).
6. C. J. Lanci et al., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7304–7309 (2012).

7. J. C. Sinclair, K. M. Davies, C. Vénien-Bryan, M. E. Noble, *Nat. Nanotechnol.* **6**, 558–562 (2011).
8. J. Zhang, F. Zheng, G. Grigoryan, *Curr. Opin. Struct. Biol.* **27**, 79–86 (2014).
9. J. D. Brodin et al., *Nat. Chem.* **4**, 375–382 (2012).
10. F. Baneyx, J. F. Matthaei, *Curr. Opin. Biotechnol.* **28**, 39–45 (2014).
11. E. D. Levy, E. B. Erba, C. V. Robinson, S. A. Teichmann, *Nature* **453**, 1262–1265 (2008).
12. N. P. King et al., *Science* **336**, 1171–1174 (2012).
13. N. P. King et al., *Nature* **510**, 103–108 (2014).
14. I. André, P. Bradley, C. Wang, D. Baker, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 17656–17661 (2007).
15. I. André, C. E. Strauss, D. B. Kaplan, P. Bradley, D. Baker, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 16148–16152 (2008).
16. B. L. Nannenga, M. G. Iadanza, B. S. Vollmar, T. Gonen, in *Current Protocols in Protein Science*, chap. 17, unit 17.15 (Wiley, New York, 2013).
17. S. W. Wukovitz, T. O. Yeates, *Nat. Struct. Biol.* **2**, 1062–1067 (1995).
18. F. DiMaio, A. Leaver-Fay, P. Bradley, D. Baker, I. André, *PLOS ONE* **6**, e20450 (2011).
19. S. J. Fleishman et al., *PLOS ONE* **6**, e20161 (2011).
20. M. C. Lawrence, P. M. Colman, *J. Mol. Biol.* **234**, 946–950 (1993).
21. B. Kuhlman et al., *Science* **302**, 1364–1368 (2003).
22. Z. S. Hendsch, T. Jonsson, R. T. Sauer, B. Tidor, *Biochemistry* **35**, 7621–7625 (1996).
23. E. H. Kellogg, O. F. Lange, D. Baker, *J. Phys. Chem. B* **116**, 11405–11413 (2012).
24. L. G. Nivón, R. Moretti, D. Baker, *PLOS ONE* **8**, e59004 (2013).
25. B. Gipson, X. Zeng, Z. Y. Zhang, H. Stahlberg, *J. Struct. Biol.* **157**, 64–72 (2007).
26. D. Shi, B. L. Nannenga, M. G. Iadanza, T. Gonen, *eLife* **2**, e01345 [abstr.] (2013).

#### ACKNOWLEDGMENTS

We thank D. Shi and J. de la Cruz for help with EM, S. Sanchez-Martinez for help with protein expression, J. Bale and N. King for support and helpful discussions, and W. Sheffer for Rosetta code. We would also like to thank members of both the Baker and Gonen labs for scripts and useful discussions. We thank HHMI's Janellia Research Campus visitor program, the University of Washington's Biochemistry Department and Biological Physics, Structure, and Design program, U.S. Defense Threat Reduction Agency, U.S. Air Force Office of Scientific Research, and the Howard Hughes Medical Institute for funding and support. S.G., F.D., T.G., and D.B. are inventors on a provisional patent application that covers the method for the design of self-assembling 2D protein material. Supporting materials, methods and the design models and sample design files are available in the supplementary materials. Author contributions: S.G. and F.D. worked on the docking and design. S.G. worked on design optimization. F.D. wrote Rosetta code. S.G., F.D., and D.B. computationally analyzed the designs. S.G. worked on the biochemistry, electron microscopy, and data analysis. S.G. and T.G. analyzed the EM data. All authors designed the research and wrote, edited, and contributed to the manuscript.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6241/1365/suppl/DC1  
Materials and Methods  
Supplementary Text  
Figs. S1 to S5  
Tables S1 and S2  
References (27–39)

25 February 2015; accepted 11 May 2015  
10.1126/science.aaa9897

---

*This copy is for your personal, non-commercial use only.*

---

**If you wish to distribute this article to others**, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

**Permission to republish or repurpose articles or portions of articles** can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of June 20, 2015 ):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/348/6241/1365.full.html>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/content/suppl/2015/06/17/348.6241.1365.DC1.html>

This article **cites 37 articles**, 7 of which can be accessed free:

<http://www.sciencemag.org/content/348/6241/1365.full.html#ref-list-1>

This article appears in the following **subject collections**:

Biochemistry

<http://www.sciencemag.org/cgi/collection/biochem>