

PROTEIN DESIGN

Computational design of mechanically coupled axle-rotor protein assemblies

A. Courbet^{1,2,3†}, J. Hansen^{1†}, Y. Hsia^{1,2}, N. Bethel^{1,2,3}, Y.-J. Park¹, C. Xu^{1,2,3}, A. Moyer^{1,2}, S. E. Boyken^{1,2†}, G. Ueda^{1,2}, U. Nattermann^{1,2}, D. Nagarajan^{1,2}, D.-A. Silva^{1,2,4,5}, W. Sheffler^{1,2}, J. Quispe¹, A. Nord⁶, N. King^{1,2}, P. Bradley⁷, D. Veessler^{1,3}, J. Kollman¹, D. Baker^{1,2,3*}

Natural molecular machines contain protein components that undergo motion relative to each other. Designing such mechanically constrained nanoscale protein architectures with internal degrees of freedom is an outstanding challenge for computational protein design. Here we explore the de novo construction of protein machinery from designed axle and rotor components with internal cyclic or dihedral symmetry. We find that the axle-rotor systems assemble in vitro and in vivo as designed. Using cryo-electron microscopy, we find that these systems populate conformationally variable relative orientations reflecting the symmetry of the coupled components and the computationally designed interface energy landscape. These mechanical systems with internal degrees of freedom are a step toward the design of genetically encodable nanomachines.

Intericate protein nanomachines in nature have evolved to process energy and information by coupling biochemical free energy to mechanical work. Among the best studied and most sophisticated are protein rotary machines such as the F_1 motor of adenosine triphosphatase or the bacterial flagellum, which contain axle-like and ring-like symmetric protein components capable of constrained dynamic motion relative to each other (1–3). Feynman's 1959 lecture on nanotechnology as a means to leverage the properties of materials at the molecular scale (4) inspired interest in synthetic nanomachines (5, 6). Synthetic chemists were the first to design molecules with mechanically coupled components (7–9). Nucleic acid nanotechnologies have more recently been used to construct rotary systems (10). Designed dynamic protein mechanical systems are of great interest given the richer functionality of proteins, but with this functionality comes more complex folding and a greater diversity of noncovalent interactions, which, despite recent advances in design of static protein nanostructures (11–19), has made the design of protein machines an outstanding challenge (20).

We explored the design of protein mechanical systems through a first-principle, bottom-up approach that focuses on operational concepts independent from the complex evolutionary trajectory of natural nanomachines. Previous two-component protein assembly design studies

have focused on nanomaterials such as icosahedral nanocages (21) and two-dimensional (2D) arrays (19) in which the components have fixed orientations relative to one another. In this work, we sought to design a nanoscale simple machine or kinematic pair (22, 23) in which two mechanically coupled protein components undergo Brownian diffusion along internal degrees of freedom (DOFs). We used a hierarchical design approach with three steps: (i) de novo design of stable and rigid protein components suitable for assembly into constrained mechanical systems, (ii) directed self-assembly of these components into hetero-oligomeric complexes, and (iii) shaping of the multistate energetic landscape along the mechanical DOFs. A major challenge is to design the interface between the two designed rigid bodies to have sufficiently low energy to drive self-assembly, while still allowing relative motion of the components. We started from a machine blueprint that consists of two coupled structural components resembling an axle and rotor (Fig. 1A), in which, similar to natural protein rotary systems, the features of the energy landscape are determined by the symmetry of the interacting components, their shape complementarity, and specific interactions across the interface.

Computational design of protein mechanical components

We first sought to design ringlike protein topologies, or rotors, with a range of inner diameter sizes capable of accommodating an axle-like binding partner (Fig. 1B). In a first approach, we started from de novo α -helical tandem repeat proteins (24) and redesigned them to form C1 single-chain structures or symmetric C3 or C4 homo-oligomers. In a second approach, we started from de novo helical repeat proteins (DHRs) and helical bundle heterodimers and used a hierarchical design procedure based on architecture-guided

rigid helical fusion (14) to build C3 and C5 cyclic symmetric rotor structures. To facilitate subsequent microscopy characterization and modularity, we fused another set of DHRs at the outer side of the rotors, generating armlike extensions (Fig. 1, A and B). Synthetic genes encoding these rotor designs (12xC3s, 12xC4s, 2xC5s) were synthesized and the proteins expressed in *Escherichia coli*. All designed proteins were soluble after purification on nickel-nitrilotriacetic acid (Ni-NTA) columns, and ~23% (6/26) had size exclusion chromatography (SEC) profiles that matched the expected theoretical elution profile for the oligomerization state (figs. S1 and S2 and table S1). These designs were further examined using small-angle x-ray scattering (SAXS) (25, 26), negative stain electron microscopy, or cryo-electron microscopy (cryo-EM) (fig. S1). For the C3_R1 rotor, SAXS data analysis was consistent with the computational model [volatility ratio (V_r) = 4.684] (table S2 and fig. S2), and we were able to determine using cryo-EM a 6.0-Å 3D reconstruction that was close to the design model [backbone root mean square deviation (RMSD) = 3.451 Å] (Fig. 1B; figs. S1, S4, and S5; and table S3). Similar results were obtained for another design of the same topology (C3_R2) (fig. S1). For the C4 design C4_1, we obtained a 7.9-Å cryo-EM density map closely consistent with the design model (backbone RMSD = 1.8 Å) (Fig. 1B; figs. S1, S5, and S6; and table S3). C3 and C5 rotors with larger inner diameter and different topology (C3_R3 and C5_2) were characterized using negative stain EM, yielding low-resolution 3D reconstructions consistent with the design model (Fig. 1B and fig. S1).

We next sought to design high-aspect ratio protein components, or axles, onto which the designed rotor protein could be threaded, using three different design approaches. In a first approach, single-helix backbones were parametrically generated, and the sequence was optimized in D2, D3, or D4 dihedral symmetry using buried hydrogen bond networks and hydrophobic contacts to produce self-assembling homo-oligomer interfaces with the high level of specificity needed for dihedral assembly (Fig. 2A). To increase the total mass and diversify the shape for subsequent EM analysis, the termini of these rod-shaped structures were rigidly fused to cyclic homo-oligomers of matching symmetry (i.e., D_n dihedral assemblies were fused with C_n cyclic assemblies) to create dumbbell-shaped structures. In a second approach, two copies of designed cyclic homo-oligomers were assembled into dihedral structures by connecting them with rigid helical bundle connectors built using fragment sampling (Fig. 2B). In a third approach, parametrically generated homotrimer backbones consisting of helical hairpin monomer topologies (27) were circularly permuted and

¹Department of Biochemistry, University of Washington, Seattle, WA, USA. ²Institute for Protein Design, University of Washington, Seattle, WA, USA. ³Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. ⁴Division of Life Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. ⁵Monod Bio, Inc., Seattle, WA, USA. ⁶Centre de Biologie Structurale (CBS), INSERM, CNRS, Université Montpellier, Montpellier, France. ⁷Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

*Corresponding author. Email: dabaker@uw.edu

†These authors contributed equally to this work.

‡Present address: Outpace Bio, Inc., Seattle, WA, USA.

elongated to generate extended C3 homo-oligomers (Fig. 2C). Details of the methods, as well as scripts for carrying out the design calculations, are provided in the supplementary materials. Synthetic genes encoding axle designs generated from the three approaches (12xC3s, 12xC5s, 12xC8s, 6xD2s, 12xD3s, 6xD4s, 6xD5s, and 12xD8s) were obtained, and the proteins were expressed in *E. coli*. The designed proteins that were well expressed, soluble, and readily purified by Ni-NTA affinity chromatography were further purified on SEC. Success rates for the first, second, and third approach were 37.5% (6/16), 43% (14/32), and 33% (4/12), respectively, as assessed by the match between estimated molecular weight (MW) from SEC with the MW of the design model (Fig. 2D; figs. S1 to S3; and table S1). Designs with matching SEC traces were further examined using SAXS, negative stain EM, and cryo-EM (figs. S1 to S3).

The first approach generated D2, D3, and D4 axle-like structures with folds featuring interdigitated helices with extended hydrogen bond networks. We obtained a 4.2-Å 3D reconstruction of a D3 axle (D3_3) with backbone nearly identical to the design model (backbone RMSD = 1.9 Å) (Fig. 2A and figs. S3, S4, and S7); SAXS data were also consistent with the design model ($V_r = 6.0$) (table S2 and figs. S1 and S2). The central homohexameric

50-residue helices (D3_2) could also be solubly expressed and formed an oligomeric self-assembly that eluted at the expected volume (fig. S3 and table S1). D3 design D3_1, consisting of 36-residue-long single helices, was produced by chemical peptide synthesis and assembled into a homohexamer (figs. S3 and S8), and fusion to wheel-like C3s generated a larger D3 oligomer as designed (D3_4) (fig. S3). A D4 peptide homo-oligomer designed using the same approach (D4_1) had a SEC profile consistent with the expected oligomeric state (figs. S2 and S3 and table S1). Negative stain EM of a D2 design (D2_2) yielded a low-resolution 3D reconstruction with the overall features of the design model (Fig. 2D and fig. S3); the corresponding central 50-residue D2 peptide (D2_1) could also be expressed, and the SEC elution volume corresponded to the expected oligomeric state (fig. S3 and table S1).

The second approach generated D3, D4, D5, and D8 axle-like structures, with interdigitated helices with internal cavities in the D5 and D8 cases where each central helix only contacts the two neighboring ones (Fig. 2B). We obtained a 7.4-Å electron density map of a D8 design (D8_1) revealing a backbone structure nearly identical to the design model (backbone RMSD = 2.9 Å) (Fig. 2B and figs. S3, S5, and S6). This cylinder-shaped homodecahexamer has a large central cavity, an 84-

residue helix, and opposing N and C termini close to its center (Fig. 2B and fig. S3). Negative stain EM 3D reconstructions of D8_2, D8_3, D5_2, and D4_2 were consistent with the design models (Fig. 2D and fig. S3). We converted several of these designs from dihedral to cyclic symmetry by connecting N and C termini, and two such designs, one C5 (C5_1) and one C8 (C8_1), yielded EM reconstructions with good agreement with the design model (Fig. 2D and figs. S1 and S3). SAXS profiles of additional designs (4xD3s, 2xD4s, and 1xD5) were consistent with the design models with $V_r < 10$ in most cases and measured MW within 15% of the design model for D3_1 and D3_8, and within 1% for D5_1 (table S2 and figs. S2 and S3).

The third approach yielded four C3 axles with smaller aspect ratios and overall sizes, containing a large wheel-like feature at one end, a narrow central three-helix section, and a six-helix section at the other end. SAXS profiles together with SEC traces suggested that the designed oligomerization state is realized in solution ($V_r \sim 12$) (tables S1 and S2 and figs. S1 and S2). For design C3_A1, we obtained a low-resolution cryo-EM map that recapitulates the general features of the design model, with prominent C3 symmetric DHR extremities and opposing prism-like extensions (Fig. 2C and figs. S1 and S4).

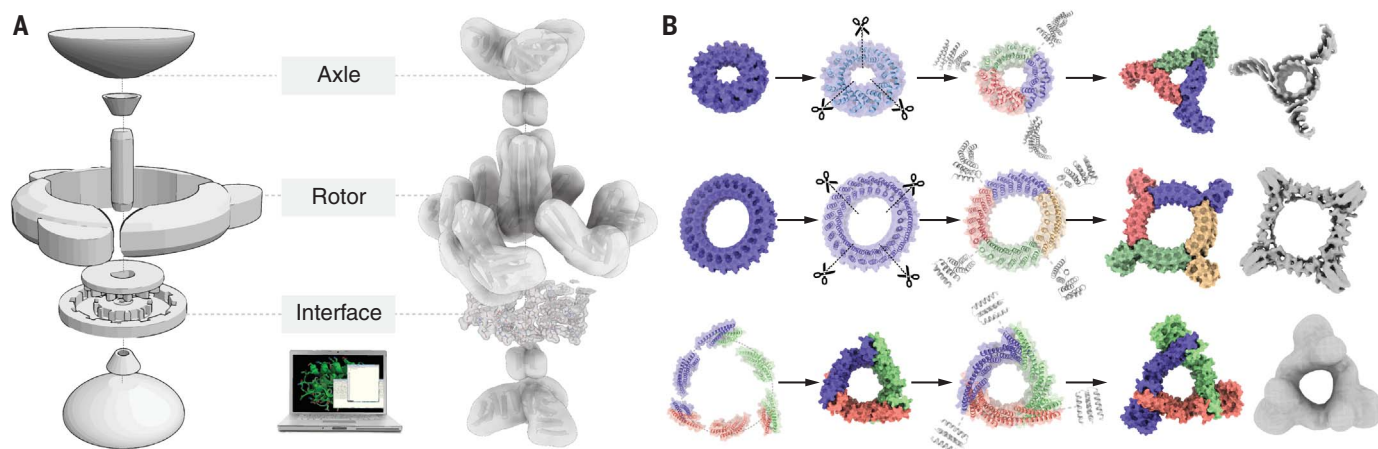


Fig. 1. Overview of protein machine assembly and rotor component design approaches. (A) (Left) A blueprint of a simple two-component machine consisting of an assembly of an axle and a rotor mechanically constrained by the shape of the interface between the two. (Middle) Systematic generation by computational design of a structurally diverse library of machine components and design of interfaces between axle and rotor that mechanically couple the components and direct assembly. (Right) Example of hierarchical design and assembly of a protein machine from axle and rotor components, here a D3 axle and C3 rotor, and interacting interface residues. Wheel-like cyclic DHRs are fused to the end of the axle and rotor components to increase mass and provide a modular handle and a structural signature to monitor conformational variability. (B) Hierarchical design strategies for rotor components. (Top) A single-chain C1 symmetric and internally C12 symmetric α -helical tandem repeat protein is

split into three subunits, and each is fused to DHRs by means of helical fusion (HelixFuse) to generate a C3 rotor (C3_R1) with an internal diameter of 28 Å. The 6.0-Å cryo-EM electron density (shown in gray) shows agreement with the design model. (Middle) A single-chain C1 symmetric and internally C24 symmetric α -helical tandem repeat protein is split into four subunits, and each is fused to DHRs to generate a C4 rotor (C4_1) with an internal diameter of 57 Å. The 7.9-Å cryo-EM electron density (shown in gray) shows agreement with the design model. (Bottom) Hetero-oligomeric helical bundles and DHRs are fused and assembled into a higher-ordered closed C3 structure through helical fusion, after which another round of helical fusion protocol is used to fuse DHRs to each subunit, to generate a C3 rotor (C3_R3) with an internal diameter of 41 Å. The negative stain electron density (shown in gray) shows agreement with the design model. Monomer subunits are colored by chain. Scale bar, 10 nm.

Design of mechanically coupled axle-rotor assemblies

We next investigated the construction of mechanically constrained axle-rotor assemblies from the designed axles and rotors. As noted above, an inherent challenge for the *de novo* design of dynamic protein complexes is to incorporate sufficient energetically favorable interactions to enable directed self-assembly without

creating deep energy minima that lock the assembly into a single state and prevent Brownian diffusion along the mechanical DOFs. We explored three approaches for constructing axle-rotor assemblies, which result in interfaces with widely varying energetics, shape complementarity, and symmetry.

First, we sought to construct two-component assemblies in which the rigid body orientation

of the axle and rotor was minimally constrained. We designed symmetry-mismatched axle-rotor interfaces with low orientational specificity and loose interface packing, allowing only small numbers of close contacts across the interface and using primarily electrostatic interactions between rotor and axle, which are longer range and less dependent on shape matching than the hydrophobic interactions

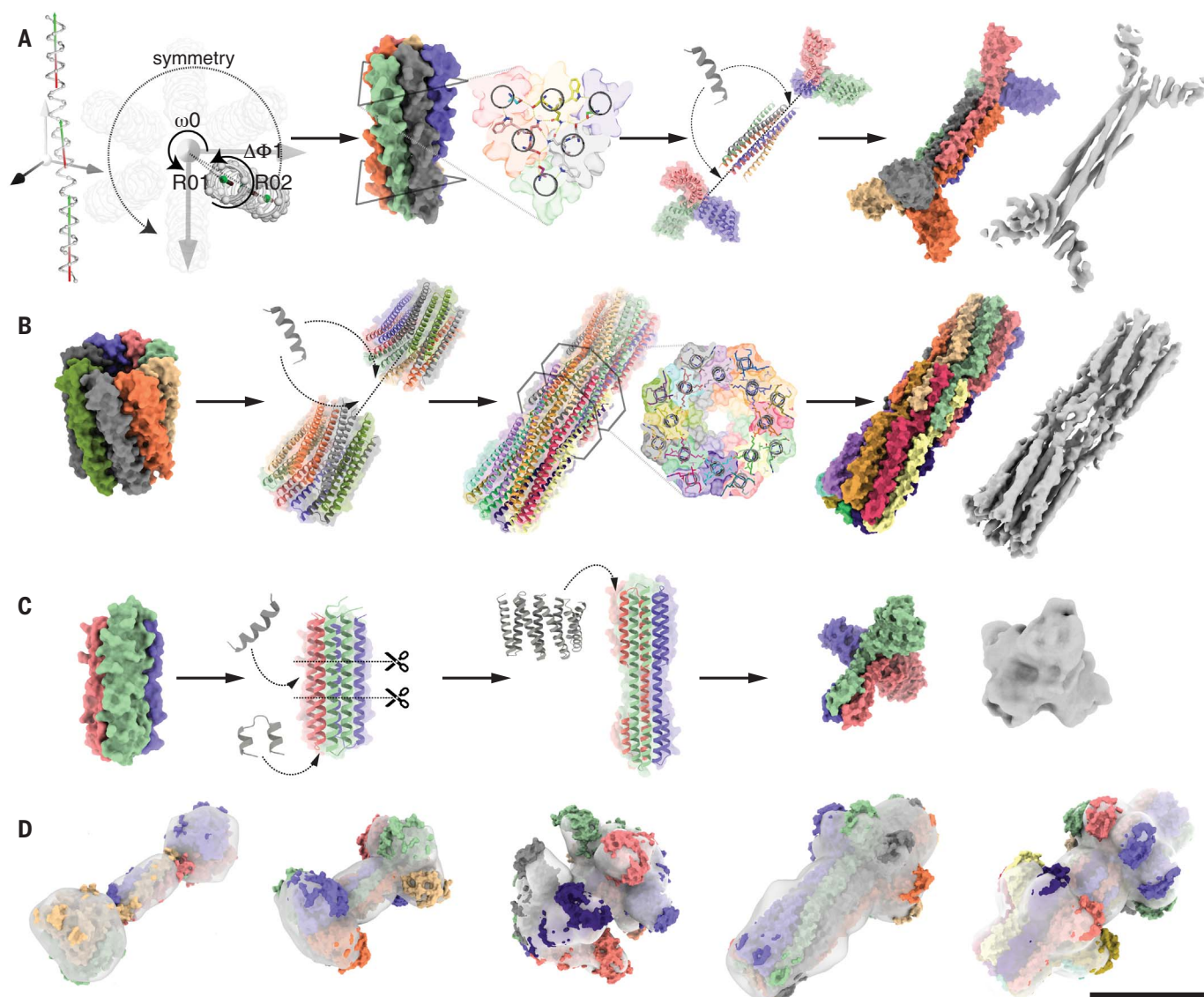


Fig. 2. Design of axle machine components. (A) Hierarchical design of a D3 symmetric homohexamer axle (D3_3). Parametric design of interdigitated helices in D3 symmetry was achieved by sampling supercoil radius (R01, R02), helical phase ($\Delta\phi_{1-1}$, $\Delta\phi_{1-2}$), supercoil phase ($\Delta\phi_{0-1}$, $\Delta\phi_{0-2}$) of two helical fragments, and the z-offset and supercoil twist (ω_0). The interface was designed using the HBNet protocol to identify hydrogen bond networks spanning the six helices mediating high-order specificity. The design was then fused to C3 wheel-like homotrimers using RosettaRemodel. The 4.2-Å cryo-EM electron density is consistent with the design model. (B) Hierarchical design of a D8 axle (D8_1). Interdigitated helical extensions at the termini of a parametrically designed C8 homohexamer were sampled using Rosetta BlueprintBuilder, and hydrogen bond

networks were identified using HBNet, while sampling rotation and translation in D8 symmetry using Rosetta SymDofMover. The 7.4-Å cryo-EM electron density is in close agreement with the design model. (C) Hierarchical design of a C3 homotrimer axle (C3_A1). A parametrically designed C3 homotrimer was circularly permuted and an extra heptad repeat added to increase the aspect ratio, after DHRs were fused to each subunit using HelixFuse. The negative stain electron density is consistent with the design model. (D) Additional axle components overlaid with experimental negative stain electron density, corresponding to (from left to right) D2 (D2_2), D4 (D4_2), D5 (D5_2), C8 (C8_1), and D8 (D8_3) designs. Model monomer subunits are colored by chain, and electron densities are shown as gray surfaces. Scale bar, 10 nm.

generally used in protein design. To prevent potential disassembly at low concentrations owing to weak axle-rotor interactions, we sought to kinetically trap the rotor around the axle by installing disulfide bonds at the rotor subunit-subunit interfaces. To gain stepwise control on the in vitro assembly process, we introduced buried histidine-mediated hydrogen bond networks at the asymmetric interfaces between rotor subunits to enable pH-controlled rotor assembly (fig. S9, and see methods in the supplementary materials). To test this approach, we selected three of the machine components described above—a D3 axle, a C3 rotor, and a C5 rotor—and constructed axle-rotor assemblies with D3-C3 and D3-C5 symmetries (designs D3-C5 and D3-C3, respectively) (Fig. 3A and fig. S10). To thread axles and rotor

together, we computationally sampled rotational and translational DOFs and designed complementary electrostatic interacting surfaces excluding positively charged residues on the axle (lysine and arginine) and negatively charged residues on the rotor (aspartate and glutamate). Given the shape complementarity between the internal diameter of the rotors and the axle thickness, the interface is tight for D3-C3, constraining the rotor on the axle, and loose for D3-C5. By design, the D3-C3 can rotate and translate along the main symmetry axis (z), whereas the D3-C5 rotor has rotation and translation components along x , y , and z (Fig. 3, A and B, and fig. S11). Synthetic genes encoding the one axle and two rotor designs were obtained, and the proteins were separately expressed in *E. coli* and purified

by Ni-NTA affinity chromatography and SEC, which indicated that the surface redesign did not affect solubility or oligomerization state (figs. S1 and S3). After stoichiometric mixing of the designed D3 axle and C3 rotor, EM analysis showed a collection of assembled and isolated axle and rotor molecules (fig. S9A, top panel). After dropping the pH and reducing the disulfide, the particles appeared as a mixture of opened, linear, and hard-to-distinguish particles (fig. S9A, middle panel). After restoring the pH under oxidizing conditions, the particles appeared fully assembled by EM (fig. S9A, bottom panel). Biolayer interferometry assays showed that the rotor and axle associated rapidly with an approximate association rate of $10^3 \text{ M}^{-1}\text{s}^{-1}$ and a dissociation constant (K_d) in the micromolar range (fig. S12). Similar results

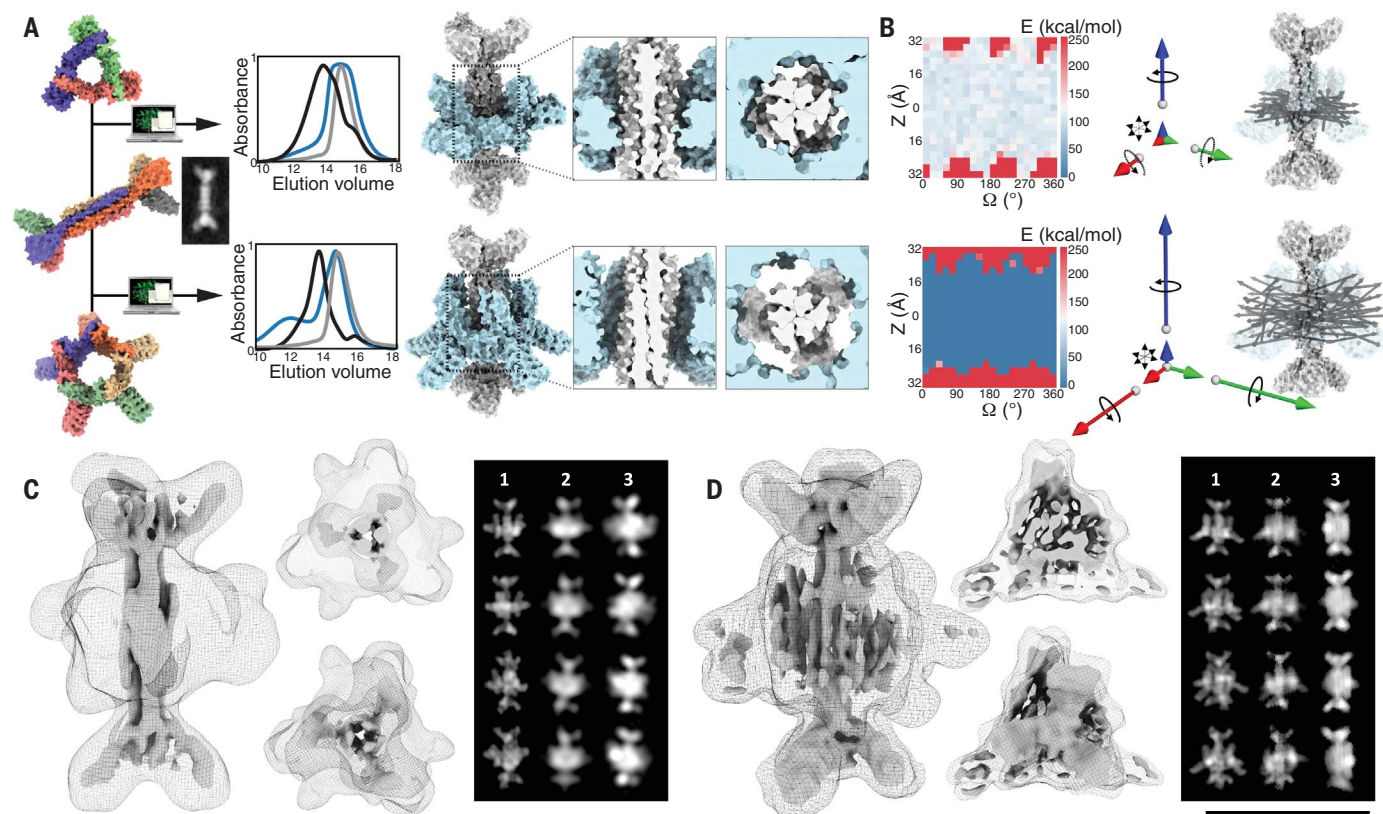


Fig. 3. Design of symmetry-mismatched D3-C3 and D3-C5 axle-rotor assemblies. (A) (Left to right) Models of a D3 axle (D3_3) and C3 (C3_R3) and C5 (C5_2) rotors and cryo-EM 2D average of axle alone before assembly. Overlaid SEC chromatograms (absorbance at 215 nm) of axle (gray), rotor (blue), and full assembly (black). Models of D3-C3 and D3-C5 assemblies with top-view and side-view close-up on interfaces; shape and symmetry results in different DOFs. (B) (Left) 2D rotation-translation energy landscapes showing a large area of low energy where the rotor can be positioned on the axle. (Right) MD simulation results are shown as vectors whose magnitude corresponds to the computed mean square displacement of the rotor relative to the axle along the six DOFs. The D3-C3 system is largely constrained to rotation along the z axis (blue), whereas the D3-C5 assembly allows rotation along x (green), y (red), and z and translation in z , x , and y . N- and C-terminal unit vectors of an ensemble of MD trajectories are superimposed on an axle-rotor

model structure. (C) (Left) 3D cryo-EM reconstruction of D3-C3, processed in D3 at 7.8-Å resolution suggests that the rotor sits midway across the D3 axle, consistent with the designed mechanical DOF. The maps are shown in side views, end-on views, and transverse slices, as surface for the axle and as mesh for the rotor, at two different thresholds. (Right) Simulated 2D class averages without (1) and with (2) conformational variability, and experimental averages (3). (D) (Left) 3D cryo-EM reconstruction of D3-C5, processed in C1 at 8.6-Å resolution, has the overall features of the designed structure, shown as surface and mesh at different thresholds. The 2D averages capture secondary structure corresponding to the C5 rotor but could not be fully resolved, which is consistent with the rotor populating conformationally variable states. (Right) Simulated 2D class averages without (1) and with (2) conformational variability, and experimental averages (3). Scale bar for cryo-EM density, 10 nm.

were obtained with D3-C5 rotary assemblies, and SEC and SAXS profiles were in agreement with the design model in both cases ($V_r < 15$) (tables S1 and S2 and figs. S2 and S10).

Second, we experimented with more direct steric coupling to limit conformational variability primarily to rotation of the rotor around

the axle. We used shape-complementary axle and rotor components to enable the incorporation of steric constraints restricting translation, leveraging Rosetta's ability to design tightly packed interfaces and hydrogen bond network-mediated specificity (27). We designed seven axle-rotor assemblies using this approach: three

with C3 symmetric axes with C1 rotors (C3-C1_1 to C3-C1_3) (fig. S10) and four larger designs with C3 axes and rotors (C3-C3_1 to C3-C3_4) (Fig. 4A and fig. S10) with DHR arm extensions. The C3 symmetry matching of the rotor and axle differs from the mismatching in the other designed assemblies, and the extent of

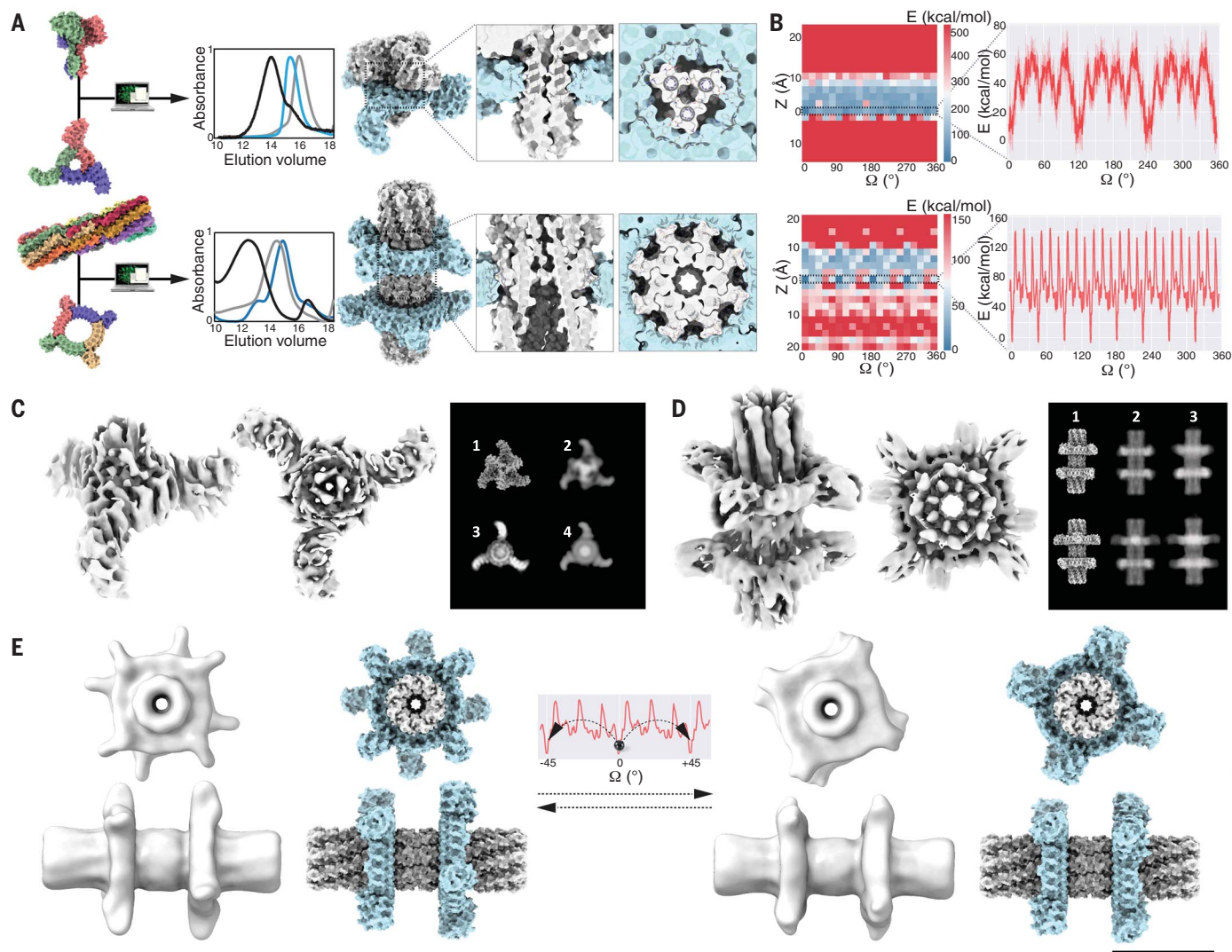


Fig. 4. Computational sculpting of the energy landscape by design of interface side-chain interactions. (A) (Left to right) Models of C3 axle (C3_A1), C3 rotor (C3_R1), D8 axle (D8_1), and C4 rotor (C4_1) used to design C3-C3 and D8-C4 axle-rotor assemblies. Overlaid SEC chromatograms (absorbance at 215 nm) of axle (gray), rotor (blue), and full assembly (black). Models of symmetry-matched C3-C3_1 and quasi-symmetric D8-C4 assemblies and close-ups on the interface reveal the shape-complementary cogwheel topology. (B) Energy landscapes corresponding to the C3-C3 (top) and D8-C4 (bottom) axle-rotor assemblies. (Left) 2D rotation-translation energy landscapes showing a narrow band of low energy where the rotor sits on the axle. (Right) 1D rotational energy landscape has three main minima corresponding to the C3 symmetry of the interface with nine additional lesser energy minima for C3-C3 and eight main energy minima corresponding to the C8 symmetry of the interface and additional 18 lesser minima for D8-C4. The energy landscapes were computed by scoring 10 independent Rosetta backbone and side-chains relax

and minimization trajectories (solid red line with error bars depicting the standard deviation; kilocalories per mole, as calculated by Rosetta). (C) Single-particle cryo-EM analysis of the C3-C3 assembly. The rotor is evident in the 6.5-Å resolution electron density in the side and top views; only a portion of the axle is resolved. In the panel to the right, the experimental 2D class averages (3) match the projection of the design model (1) more closely with conformational variability (4) than without (2). (D) Single-particle cryo-EM analysis of the designed D8-C4 rotor. The electron density (in gray) at 5.9-Å resolution shows the main features of the designed structure and two distinct rotational states (1), also visible in the simulated projections (2), which closely resemble the experimental 2D class average (3). (E) 3D variability analysis and computed rotational landscape of the D8-C4 axle-rotor assembly. The two resolved structures (shown in gray on the left and right) are separated by a 45° rotational step. Corresponding computational models are shown in spacefill (blue and gray). Top row: top view; bottom row: side view. Scale bar, 10 nm.

alignment of axle and rotor DHR arms relative to each other provides a measure of conformational variability. Design was carried out by systematically sampling rotational and translational DOFs, removing arrangements with backbone-to-backbone clashes (see methods), and then using the Rosetta HBNNet protocol and FastDesign (28) to identify interacting residues and optimize the interface energy. Each interface design trajectory generates widely different periodic energy landscapes according to interface metrics and design specifications (fig. S13) and results in shape-complementary axle-rotor interfaces with an overall cogwheel topology. C3-C1 designs were experimentally screened for assembly by expressing rotor and axle pairs bicistronically and carrying out Ni-NTA purification relying on a single His tag on the rotor component (fig. S14A). Six out of 12 designs expressed solubly and copurified, suggesting that the two components assembled in cells (fig. S14B), and three designs (C3-C1_1 to C3-C1_3) (fig. S10) were selected for further characterization. The SEC profiles in combination with native mass spectrometry indicated an oligomeric state consistent with the designed assembly, and SAXS data were also consistent with the design model ($V_r < 12$ and MW within ~10% of expected values for C3-C1_3, and ~15% for C3-C1_1 and C3-C1_2) (tables S1 and S2 and figs. S2, S10, and S14, C and D). The C3-C3 designs (C3-C3_1 to C3-C3_4) (fig. S10) were screened for in vitro assembly by stoichiometric mixing of axle and rotor, followed by SEC and SAXS analysis, which were consistent with assemblies of the expected oligomeric state ($V_r < 10$) (tables S1 and S2 and figs. S2 and S10). Biolayer interferometry showed that the designed C3 axle and C3 rotor rapidly assemble with an approximate association rate of $10^3 \text{ M}^{-1}\text{s}^{-1}$ and a K_d in the micromolar range (fig. S12).

Third, we sought to design further constrained axle-rotor assemblies by increasing the surface area of the interfaces between axle and rotor to enable more extensive sculpting of the energy landscape. We designed a symmetry-mismatched assembly consisting of a D8 axle around which two C4 rotors are assembled (D8-C4), a symmetry-mismatched assembly consisting of a C5 axle and C3 rotor (C5-C3_1 and C5-C3_2), and a C8-C4 assembly corresponding to a circular permutation version of D8-C4 (C8-C4) (Fig. 4A and fig. S10). The D8-C4 assembly with one axle for two rotors tests the incorporation of multiple coupled rotational DOFs in a multicomponent system and also provides a simple way to monitor the position of rotors relative to each other by experimental structural characterization. For the D8-C4, C5-C3, and C8-C4 designs, because the symmetry of the rotor is internally mismatched to the axle, we used a quasi-symmetric design

protocol (see methods). The C4 rotor has internal C24 symmetry, and hence is symmetry-matched to both D8 and C8 axles. In contrast, the C5-C3 arrangement has broken symmetry with a resulting energy landscape with 15 energy minima, with periodicities reflecting the constituent C5 and C3 symmetries (fig. S13). Twelve D8-C4, twelve C5-C3, and six C8-C4 designs were screened for in vitro assembly by isolating axle and rotors individually by Ni-NTA purification and mixed stoichiometrically. We selected four of these designs for further experimental investigation and obtained SEC data indicative of assembly of axle-rotor complexes, while SAXS analysis of a C5-C3 design suggested assembly of the axle-rotor complex ($V_r = 6.9$ and predicted MW within 6% of expectation) (tables S1 and S2 and figs. S2 and S9). Biolayer interferometry binding kinetics and negative stain EM data were also consistent with quantitative assembly into the designed hetero-oligomeric complex (figs. S10 and S13).

Correspondence between designed energy landscape and observed mechanical DOFs

We subjected one construct from each design approach and symmetry class to single-particle cryo-EM examination and related these data to energy landscape calculations based on the design model (Figs. 3 and 4). Comparison of the electron density data on the axle-rotor assemblies to data on the isolated rotors and axles suggests considerable variation in their rigid body orientations, as summarized in figs. S17 to S19 and S21.

For the D3-C3 and D3-C5 assemblies produced by the first approach, we obtained 2D class averages that clearly resembled projection maps computed from the design models and 3D reconstructions in close agreement with the overall design model topology and designed hetero-oligomeric state (Fig. 3, C and D; figs. S15 and S16; and table S3). For both designs, the D3 axle was clearly visible, and we obtained a high-resolution structure of the axle nearly identical to the design model. 3D reconstructions in C1, C3, and D3 of the D3-C3 axle-rotor assembly at 7.8-Å resolution showed visible density corresponding to the rotor in the middle of the axle with the C3 rotor arms clearly evident (Fig. 3C and fig. S15). 3D reconstructions of the D3-C5 design also showed clear density for the rotor, which could be isolated by masking the axle, but its resolution could not be further improved, as the secondary structure placement relative to the axle appeared variable (Fig. 3D and fig. S16). The particle alignment algorithm is likely dominated by features of the axle that is mostly in side-view in the data (figs. S17 and S18), and thus the lack of resolution of the electron density corresponding to the rotor (see supplementary materials) is probably due to variability in

the axle-rotor rigid body transform. CryoSPARC 3D variability analysis (29) suggests that the rotor can populate multiple translational and rotational conformational states around the axle (movies S1 to S4). Inspection of the cryo-EM 3D reconstruction also suggests that the rotor arms populate multiple positions along the rotational axis (Fig. 3, C and D, and figs. S17 and S18). Rosetta energy landscapes generated by rotating and translating the rotor relative to the axle suggest that a broad range of orientations are energetically accessible (Fig. 3B), and the rotor-axle rigid body orientation fluctuated in molecular dynamics (MD) simulations, with the D3-C5 assembly showing increased displacement compared with the D3-C3 assembly (Fig. 3B and figs. S11, S17, and S18). Explicit modeling of conformational variability along the designed DOFs was necessary to produce computed projections closely resembling the experimental 2D class averages (Fig. 3, C and D, and fig. S18). Taken together, the cryo-EM data, Rosetta models, and molecular dynamics simulations are consistent with the design goal of constrained mechanical coupling of axle and rotor components (see figs. S17 and S18 for summary of data indicating conformational sampling of rotor-axle rigid body DOFs).

Among the assemblies generated with the second approach, single-particle cryo-EM analysis of a C3-C3 assembly yielded 2D class averages with the axle and rotor clearly visible. Resolution was limited by the orientation bias of the particle in ice, resulting in few side views, but we were able to obtain a 6.5-Å 3D reconstruction that resembled the design model (Fig. 4A; figs. S4, S10, and S19; and table S3). 2D averages and the 3D reconstruction clearly capture the rotor component, but the axle was only partially resolved; the rotor has a mass greater than the C3 axle and clear armlike features, which likely bias the alignment algorithm in its favor. Aligning on the rotor yielded a density map with diffuse density for the axle near the rotor (fig. S19). The contrast between the diffuse density for the axle and the well-resolved density of the rotor likely reflects conformational variability (Fig. 4, C and D, and figs. S4, S18, and S19). The Rosetta energy landscape suggests that the axle-rotor assembly can primarily sample rotational rather than translational DOFs (Fig. 4B), and rotational averaging increased the similarity between projections of the design model and the experimental data (Fig. 4, C and D, and figs. S18 and S19). Taken together, the data (summarized in fig. S19) are consistent with variability along the rotational DOF, in accordance with the designed energy landscape, which has three energy minima at a 60° rotation distance and nine other 30°-spaced degenerate alternative wells separated by low energy barriers (Fig. 4B and figs. S13 and S20).

The D8-C4 design generated by the third approach has a rugged energy landscape, with a dynamic range of 151 kcal/mol (as estimated by Rosetta), and eight steep wells spaced 45° stepwise along the rotational axis corresponding to the high symmetry of the interface (Fig. 4B). Consistent with the deep minima in this landscape, we obtained a cryo-EM map of ~5.9-Å resolution that is close to the design model (Fig. 4, C and D; fig. S6; and table S3). 3D variability analysis calculations using cryoSPARC (30) suggested two nearly equiprobable states in which the rotor arms are either aligned or offset, as in the eclipsed and staggered arrangements of ethane (Fig. 4, D and E; figs. S6, S18, and S21; and movie S5). The two rotational states of one rotor relative to the other suggest energy minima spaced by 45° along the rotational axis, consistent with an eightfold step-like feature in the frequency spectrum analysis of the computed energy landscape (fig. S21). Although cryo-EM provides a frozen snapshot of molecules and not a real-time measurement of diffusion, these data (summarized in fig. S20) suggest that the system populates multiple rotational states consistent with the designed energy landscape. Taken together, these results suggest that the explicit design of side-chain interactions and deep energy minima reduces the degeneracy of conformational states observed with purely electrostatic interactions and support a correspondence between the energy landscape and the observed conformational variability.

Conclusions

Our proof-of-concept axle-rotor assemblies demonstrate that protein nanostructures with internal mechanical constraints can now be systematically designed. Key to this advance is the ability to computationally design protein components with complex complementary shapes, symmetries, and topologies, such as the high-aspect ratio dihedral axle structures (D2 homotetramers to D8 homo-16-mers (Figs. 1 and 2) with oligomerization states and sizes considerably larger than previously designed dihedral structures. Our studies of assembly of these shape-complementary homo-oligomeric components into higher-order hetero-oligomeric structures with internal DOFs provide insights toward the design of complex protein nanomachines. First, computational sculpting of the interface between the components can be used to promote self-assembly of constrained systems with chosen internal DOFs. Second, the shape and periodicity of the resulting energy landscape is determined by the symmetry of components, the shape complementarity of the interface, and the balance between hydrophobic packing and conformationally promiscuous electrostatic interactions (Figs. 3, A and B, and 4, A and B). Symmetry mismatch generates assemblies with larger numbers of en-

ergy minima than symmetry-matched ones evident in the frequency domain (figs. S13 and S20), and explicit design of close side-chain packing across the interface results in deeper minima and higher barriers than nonspecific interactions (Figs. 3 and 4 and fig. S13). In general, the surface area of the interface between axle and rotor scales with the number of subunits in the symmetry, with larger surface areas providing a larger energetic dynamic range accessible for design (Figs. 3 and 4 and fig. S13). The combination of the conformational variability apparent in the cryo-EM data of D3-C3, D3-C5, and C3-C3 designs (Figs. 3, C and D, and 4, C and D, and figs. S4 and S15 to S19), the Rosetta and MD simulations (Figs. 3B and 4B and fig. S11), and the discrete states observed for the D8-C4 design (Fig. 4, D and E, and figs. S6 and S21) suggests that these assemblies populate multiple rotational states (the axle-rotor assemblies also have multiple symmetrically identical yet physically distinct rotational states—for example, rotation of the C3 rotor around the C3 axle by 120°—which cannot be distinguished by cryo-EM). Our cryo-EM analysis cannot distinguish whether the conformational variability reflects rotational motion or states captured during axle-rotor assembly and does not report on energy barrier heights; time-resolved microscopy at the single-molecule level will be required to reveal the dynamics of transitions between the different states and to relate the computational sculpting of the rotational energy landscapes to Brownian dynamics.

The internal periodic but asymmetric rotational energy landscapes of our mechanically coupled axle-rotor systems provide one of two needed elements for a directional motor. Coupling to an energy input to break detailed balance and drive directional motion remains to be designed: for example, the interface between machine components could be designed for binding and catalysis of a small-molecule fuel (22). Symmetry mismatch, which plays a crucial role in torque generation in natural motors (31, 32), can be incorporated in synthetic protein motors, as illustrated here for our axle-rotor assemblies. Modular assembly could lead to compound machines for advanced operation or integration within nanomaterials, and the components can be further functionalized using reversible heterodimer extensions (33) (fig. S22). Our protein systems can be genetically encoded for multicomponent self-assembly within cells (fig. S14) or in vitro (figs. S9 and S12). Taken together, these approaches could enable the engineering of a range of nanodevices for medicine, material sciences, or industrial bioprocesses. More fundamentally, de novo design provides a bottom-up platform to explore the fundamental principles and mechanisms underlying nanomachine

function that complements long-standing studies of the elaborate molecular machines produced by natural evolution.

REFERENCES AND NOTES

- W. Junge, N. Nelson, *Annu. Rev. Biochem.* **84**, 631–657 (2015).
- D. Okuno, R. Iino, H. Noji, *J. Biochem.* **149**, 655–664 (2011).
- J. C. Deme *et al.*, *Nat. Microbiol.* **5**, 1553–1564 (2020).
- R. P. Feynman, *Eng. Sci.* **23**, 22–36 (1960).
- L. Zhang, V. Marcos, D. A. Leigh, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9397–9404 (2018).
- K. Drexler, *Trends Biotechnol.* **17**, 5–7 (1999).
- B. L. Feringa, *J. Org. Chem.* **72**, 6635–6652 (2007).
- J.-P. Sauvage, *Angew. Chem. Int. Ed.* **56**, 11080–11093 (2017).
- C. Cheng, J. F. Stoddart, *ChemPhysChem* **17**, 1780–1793 (2016).
- H. Ramezani, H. Dietz, *Nat. Rev. Genet.* **21**, 5–26 (2020).
- D. Baker, *Protein Sci.* **28**, 678–683 (2019).
- G. L. Butterfield *et al.*, *Nature* **552**, 415–420 (2017).
- Z. Chen *et al.*, *Science* **368**, 78–84 (2020).
- Y. Hsia *et al.*, *Nat. Commun.* **12**, 2294 (2021).
- R. A. Langan *et al.*, *Nature* **572**, 205–210 (2019).
- G. Ueda *et al.*, *eLife* **9**, e57659 (2020).
- C. Xu *et al.*, *Nature* **585**, 129–134 (2020).
- R. Divine *et al.*, *Science* **372**, eabd9994 (2021).
- A. J. Ben-Sasson *et al.*, *Nature* **589**, 468–473 (2021).
- B. Kuhlman, P. Bradley, *Nat. Rev. Mol. Cell Biol.* **20**, 681–697 (2019).
- J. B. Bale *et al.*, *Science* **353**, 389–394 (2016).
- H. Flechsig, A. S. Mikhailov, *J. R. Soc. Interface* **16**, 20190244 (2019).
- T. A. Hearson, *Philos. Trans. R. Soc. London Ser. A* **187**, 15–40 (1896).
- L. Doyle *et al.*, *Nature* **528**, 585–588 (2015).
- K. N. Dyer *et al.*, *Methods Mol. Biol.* **1091**, 245–258 (2014).
- G. L. Hura *et al.*, *Nat. Methods* **10**, 453–454 (2013).
- S. E. Boyken *et al.*, *Science* **352**, 680–687 (2016).
- J. B. Maguire *et al.*, *Proteins* **89**, 436–449 (2021).
- A. Punjani, D. J. Fleet, *J. Struct. Biol.* **213**, 107702 (2021).
- A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, *Nat. Methods* **14**, 290–296 (2017).
- M. Sobti *et al.*, *Nat. Commun.* **11**, 2615 (2020).
- D. D. Majewski *et al.*, *Nat. Commun.* **10**, 626 (2019).
- D. D. Sahtoe *et al.*, *Science* **375**, eabj7662 (2022).
- A. Courbet, alexiscourbet/Computational-design-of-protein-nanomachines: v1.0, Zenodo (2022); <https://doi.org/10.5281/zenodo.6370776>.

ACKNOWLEDGMENTS

We thank B. Carragher, C. Potter, E. Eng, L. Yen, and M. Kopylov of the New York Structural Biology Center for assistance and helpful discussions. We especially thank S. Scheres of MRC-LMB for helpful discussions and guidance regarding cryo-EM data processing. We thank the Rosetta@Home user base for donating their computational hours to run forward folding simulations. We thank F. Busch and V. Wysocki at Ohio State University for providing expert support with native mass spectrometry experiments and V. Wysocki at Ohio State University for providing expert support with native mass spectrometry experiments. We especially thank D. Sahtoe for all the scientific support and many insightful discussions. Thanks to F. Praetorius for the brainstorming sessions dedicated to designing de novo protein motors. An additional thanks to A. Bera, M. Bick, and J. Decarreau for support in crystallography and optical microscopy respectively. Thanks to T. Daniel and L. Ceze for all the great interactions and fascinating ideas and discussions regarding the design of protein nanomachines, either computational or mechanical. Thanks to B. Coventry for very helpful advice and computational help and L. Stewart for expert help, advice, perspectives, and discussions. **Funding:** This work was supported by National Science Foundation (NSF) award 1629214 (D.B.); a generous gift from the Audacious Project (D.B.); the Open Philanthropy Project Improving Protein Design Fund (D.B.); University of Washington Arnold and Mabel Beckman cryo-EM center (D.B., D.V., J.K., and J.Q.); National Institute of Allergy and Infectious Diseases grants DP1AI158186 and HHSN272201700059C (D.V.); a Pew Biomedical Scholars Award (D.V.); an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund (D.V.); a Human

Frontiers Science Program Long-Term Fellowship (A.C.); a Washington Research Foundation Senior Fellowship (A.C.); Howard Hughes Medical Institute research (A.C.); Howard Hughes Medical Institute Hanna Gray fellowship grant GT11817 (N.B.). An S10 award (S10OD032290 to D.B., D.V., J.K., and J.Q.) funded the purchase of a Glacios microscope. SAXS data were collected at the Advanced Light Source (ALS) SIBYLS beamline on behalf of US DOE-BER, through the Integrated Diffraction Analysis Technologies (IDAT) program. Additional support comes from the NIGMS project ALS-ENABLE (P30 GM124169) and a High-End Instrumentation Grant (S10OD018483). The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health (NIH), National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. This research used resources of the ALS, a US DOE Office of Science User Facility under contract DE-AC02-05CH11231. Some of this work was performed at the Pacific Northwest Center for Cryo-EM (PNCC), which was supported by NIH grant U24GM129547 and performed at the PNCC at Oregon Health & Science University and accessed through EMSL (grid.436923.9), a DOE Office of Science User

Facility sponsored by the Office of Biological and Environmental Research. Molecular graphics and analyses were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. **Author contributions:** Conceptualization: A.C. and D.B. Methodology: A.C., D.B., J.H., and J.K. Software: A.C., Y.H., C.X., S.E.B., G.U., U.N., P.B., D.B., D.-A.S., A.M., N.K., W.S., and N.B. Validation: A.C., D.B., J.H., J.K., and Y.H. Formal analysis: A.C., J.H., and N.B. Investigation: A.C., J.H., N.B., Y.-J.P., A.N., D.N., and J.Q. Resources: A.C., D.B., J.H., J.K., and J.Q. Data curation: A.C., J.H., and Y.H. Writing – original draft: A.C. Writing – review & editing: A.C., D.B., J.H., Y.H., and J.K. Visualization: A.C., J.H., and N.B. Supervision: D.B. and J.K. Project administration: A.C. and D.B. Funding acquisition: D.B., J.K., D.V., A.C., and Y.H. **Competing interests:** A.C., D.B., J.H., J.K., N.B., and Y.H. are inventors on a provisional patent application submitted by the University of Washington for the design, composition, and function of the proteins created in this study. **Data and materials availability:** All data are available in the main text or

the supplementary materials. All the EM maps have been deposited in the Electron Microscopy Data Bank (accession codes EMD-25575, EMD-25576, EMD-25577, EMD-25578, EMD-25579, and EMD-25580). All code is available from GitHub (<https://github.com/alexiscourbet/>) and is archived on Zenodo (34).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm1183

Materials and Methods

Figs. S1 to S22

Tables S1 to S3

References (35–59)

Movies S1 to S5

Data S1 to S5

MDAR Reproducibility Checklist

[View/request a protocol for this paper from Bio-protocol.](#)

26 August 2021; accepted 21 March 2022

10.1126/science.abm1183



Computational design of mechanically coupled axle-rotor protein assemblies

A. Courbet, J. Hansen, Y. Hsia, N. Bethel, Y.-J. Park, C. Xu, A. Moyer, S. E. Boyken, G. Ueda, U. Nattermann, D. Nagarajan, D.-A. Silva, W. Sheffler, J. Quispe, A. Nord, N. King, P. Bradley, D. Veessler, J. Kollman, and D. Baker

Science **376** (6591), . DOI: 10.1126/science.abm1183

Steps toward a nanomachine

Protein rotary machines such as ATP synthase contain axle-like and ring-like components and couple biochemical energy to the mechanical work of rotating the components relative to each other. Courbet *et al.* have taken a step toward designing such axle-rotor nanomachines. A structural requirement is that interactions between the components must be strong enough to allow assembly but still allow different rotational states to be populated. The authors met this design challenge and computationally designed ring-like protein topologies (rotors) with a range of inner diameters that accommodate designed axle-like binding partners. The systems assemble and populate the different rotational states anticipated by the designs. These rotational energy landscapes provide one of two needed elements for a directional motor. —VV

View the article online

<https://www.science.org/doi/10.1126/science.abm1183>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works