

SCIENTIFIC REPORTS



OPEN

Modular assembly of a protein nanotriangle using orthogonally interacting coiled coils

Won Min Park¹, Mostafa Bedewy^{2,5}, Karl K. Berggren^{2,3} & Amy E. Keating^{1,4}

Synthetic protein assemblies that adopt programmed shapes would support many applications in nanotechnology. We used a rational design approach that exploits the modularity of orthogonally interacting coiled coils to create a self-assembled protein nanotriangle. Coiled coils have frequently been used to construct nanoassemblies and materials, but rarely with successful prior specification of the resulting structure. We designed a heterotrimer from three pairs of heterodimeric coiled coils that mediate specific interactions while avoiding undesired crosstalk. Non-associating pairs of coiled-coil units were strategically fused to generate three chains that were predicted to preferentially form the heterotrimer, and a rational annealing process led to the desired oligomer. Extensive biophysical characterization and modeling support the formation of a molecular triangle, which is a shape distinct from naturally occurring supramolecular nanostructures. Our approach can be extended to design more complex nanostructures using additional coiled-coil modules, other protein parts, or templated surfaces.

Protein assembly is a versatile bottom-up approach for creating precise structures at the nanometer scale¹. Such structures are critical to the development of novel materials and biologically functional devices. Short peptides provide minimal units for non-covalent association to mediate self-assembly, and substantial advances have been achieved using naturally derived or synthetic peptides to make fibers, nanotubes and mesh-like scaffolds². Folded protein assembly is guided by shape- and chemical-complementarity that can establish highly specific non-covalent associations. Protein-protein interfaces can be harnessed or newly designed to build geometrically well-defined nanostructures, although the complex and cooperative interactions at interfaces are difficult to control rationally. Recently, computational methods have enabled stunning *de novo* design of protein nanostructures including two-dimensional arrays³ and cages with diverse structural features⁴⁻⁷. A rational approach to building with native components is to combine known folds into fused assembly units. Symmetric cages and cubes have been constructed in this way, with distinct domain interfaces aligned into controlled geometries by rigid helical linkers⁸⁻¹⁰. Similarly, protein assembly has been demonstrated using oligomeric protein domains connected by short linkers^{11,12} biotin-streptavidin binding¹³, or metal-directed coordination¹⁴.

As an alternative to designing specialized components for each assembly task, a set of modular building blocks can potentially provide enhanced flexibility for constructing biomolecular nanostructures. For example, coiled coils, which are rod-like complexes of two or more supercoiled α -helices, have designable specificity, oligomeric state, length, and helix orientation. The interactions between helices in coiled coils is controlled by hydrophobic-polar and charge patterning of amino acids in a seven-residue repeat that is conventionally labeled *abcdefg*, with *a* and *d* positions typically hydrophobic^{15,16}. Based upon a rich understanding of coiled-coil sequence-structure relationships that has been obtained over the past ~25 years, *de novo* designed coiled-coil toolkits, i.e. sets of coiled-coil units with known interaction properties, have been generated by using systematic search and/or rational^{17,18} or computational approaches¹⁹⁻²³. Because of their utility as short self-assembling protein modules, coiled coils have been used to construct a variety of assemblies, including fibers²⁴, cages^{25,26}, and nanotubes²⁷ with extended and geometrically irregular structures²⁸.

¹Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139, USA. ²Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139, USA. ³Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139, USA. ⁴Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts, 02139, USA. ⁵Present address: Department of Industrial Engineering, University of Pittsburgh, 3700 O'Hara Street, Pittsburgh, Pennsylvania, 15261, USA. Correspondence and requests for materials should be addressed to A.E.K. (email: keating@mit.edu)

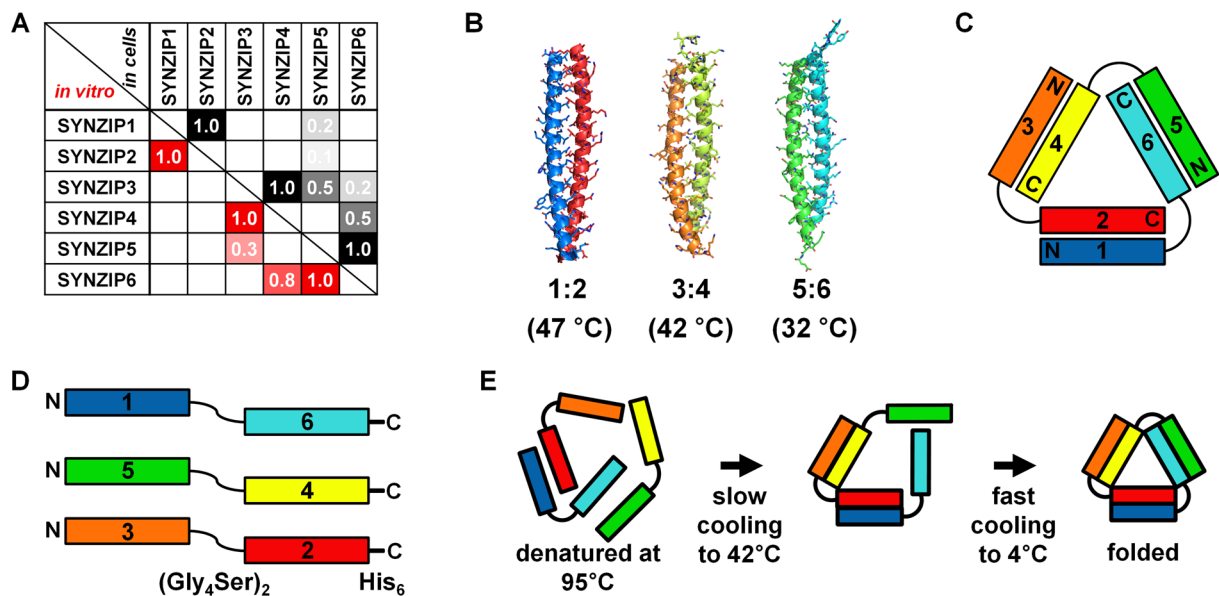


Figure 1. Orthogonally interacting SYNZIPs underlie the design of a protein nanotriangle. **(A)** Pairwise interaction profiles of SYNZIPs (1 to 6) determined by protein microarrays (*in vitro*, lower left)²⁰ and the yeast two-hybrid assays (*in cells*, upper right)¹⁹. Assay scores were normalized to a scale where 1.0 indicates tight binding. The corresponding K_D values are <10, <30, and <15 nM for 1:2, 3:4, and 5:6, respectively, and are approximated to be >400 nM for 3:5, and 4:6¹⁹. **(B)** Crystal structures of 1:2 (PDB ID: 3HE5) and 5:6 (PDB ID: 3HE4)²⁰, and a model of 3:4; melting temperatures are given in parentheses. **(C)** Topology of the designed heterotrimeric protein nanotriangle composed of the orthogonal SYNZIPs. **(D)** The composition of individual linked-SYNZIP fusion proteins that self-assemble into the nanotriangle shown in panel C. **(E)** A two-step thermal annealing process for assembly of the linked-SYNZIP fusion proteins, and illustration of desired assemblies at each step.

Modular design of coiled-coil assemblies with specific atomically definable three-dimensional structures²⁸ has been challenging, and has only been demonstrated in few cases. In one striking example, multiple distinct coiled coils were arranged into a single chain that folds into a tetrahedron, with the final structure dictated by specific coiled-coil associations²⁹. The single-chain design relied on a careful ordering of elements within the chain to control the topology, which may limit the broader utility of this approach. So far, no other single-chain folded shapes based on coiled-coil modules have been reported. A different method using coiled coils to create polygonal nanoscale objects was reported, in which linked coiled coils formed self-assembling structures³⁰. Because only a single type of coiled-coil heterodimer was used in this work, control over the number of subunits was introduced using linker lengths, limiting modularity and potential generalizability.

Here, we report a simple and rational design of a protein nanotriangle that exploits the modularity of multiple orthogonally interacting coiled coils. The construction and characterization process involved: (1) design of appropriately linked coiled-coil modules; (2) recombinant production and purification of designed proteins; (3) mixing and annealing according to an optimized schedule; and (4) biophysical analysis of the assembled nanotriangle. Three previously characterized heterodimeric coiled-coil modules that specifically associate while avoiding undesired crosstalk were arranged into self-assembling building blocks (Fig. 1). The assembly geometry was encoded by the strategic fusion of pairs of non-associating sequences in combinations that disfavored competing assemblies. Specific dimerization of the designed modules favored formation of coiled-coil edges connected into a triangular shape via flexible linkers.

Results

Design of a self-assembled protein nanotriangle. Three orthogonally interacting pairs of coiled-coil modules were selected from 23 previously reported synthetic heterodimeric coiled coils (SYNZIPs) (Fig. 1A)^{19–21}. The three SYNZIP pairs 1:2, 3:4, and 5:6 (where a colon indicates non-covalent interaction) interact with high affinity (assay scores ~1.0) and are favored over weak undesired crosstalk with other SYNZIPs (assay scores < 0.8). The dissociation constants for the three binding pairs (K_D < 30 nM) are at least an order of magnitude lower than those for any other pairwise interactions within this set of proteins (K_D > 400 nM)¹⁹. Crystal structures indicate that complexes 1:2 (PDB ID: 3HE5) and 5:6 (PDB ID: 3HE4) are parallel, heterodimeric coiled coils (Fig. 1B)²⁰. Complex 3:4 is also a parallel heterodimer, and the axial sequence alignment has been determined experimentally^{19, 20}. Based on these characterizations, we built a structural model for 3:4 by comparative modeling³¹ and refined it by molecular dynamics simulation for 10 ns³². The model shows interactions between polar and charged residues at *e* and *g* positions, and hydrogen bonding of asparagine residues at the *a*-*a'* positions (Figs 1B and S1) that are characteristic of many parallel coiled-coil dimers¹⁵.

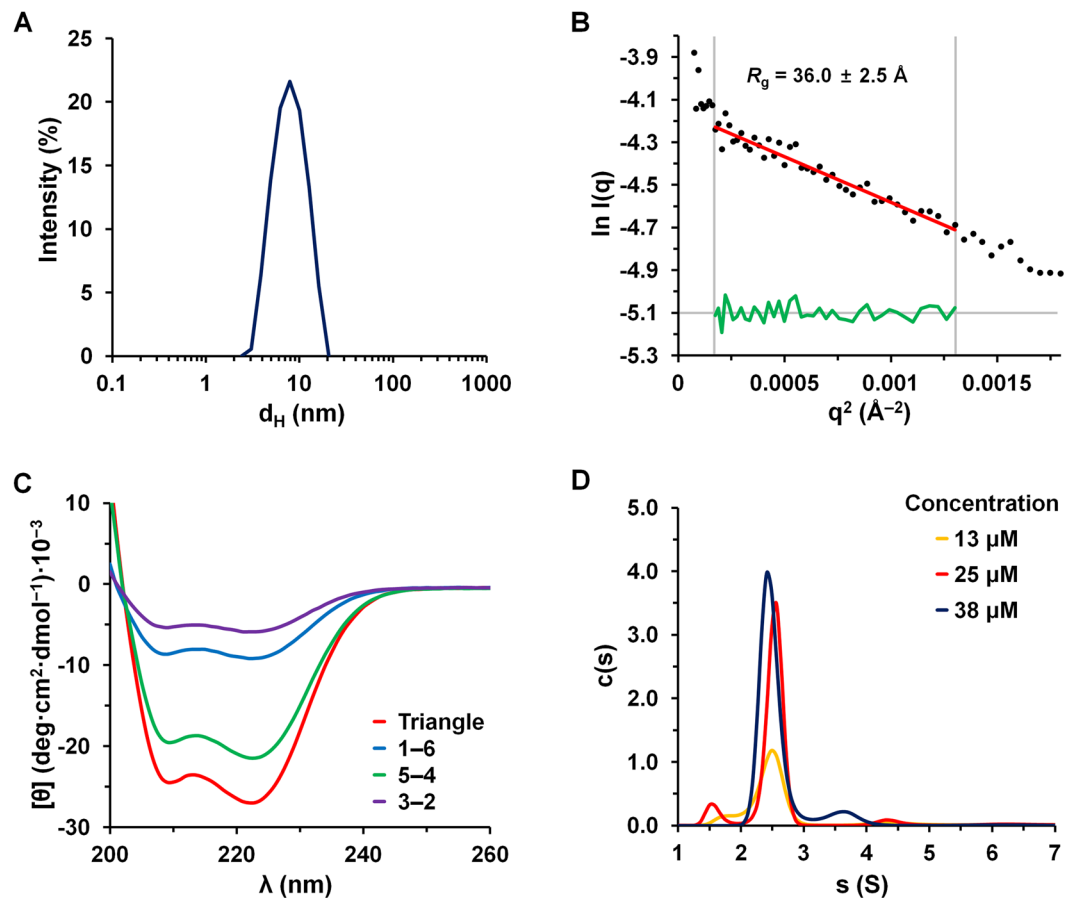


Figure 3. Characterization of size, folding, and assembly of the protein nanotriangle. (A) The hydrodynamic diameter (d_H) of the protein nanotriangle as determined by DLS. (B) Guinier plot to determine R_g from SAXS measurements at $44\ \mu\text{M}$ (black dots; Fig. S8). The fit (red line with residuals in green) is for data in the low- q range ($q \times R_g < 1.3$). (C) CD spectra for 1–6, 5–4, and 3–2 and the purified assembly (triangle) ($5\ \mu\text{M}$ of each protein and $15\ \mu\text{M}$ total for mixtures, $20\ ^\circ\text{C}$). (D) Distributions of sedimentation coefficient (s), in Svedbergs (S), for the protein nanotriangle at concentrations of 13, 25, and $38\ \mu\text{M}$, estimated from interference boundary fits (Fig. S11).

gyration (R_g) of $36.0 \pm 2.5\ \text{\AA}$ (Fig. 3B). These measurements are consistent with the expected size of the nanotriangle, given that the SYNZIP coiled-coil edges have lengths of 6.7, 5.7, and 5.5 nm (Fig. 1B). The data also confirm that the assembly does not form aggregates.

The helicity of the folded protein triangle was quantified using circular dichroism (CD) spectroscopy. Peaks at 208 and 222 nm in the CD spectrum indicated a helical structure consistent with the designed coiled-coil edges (Fig. 3C). The helix content estimated from the mean residue ellipticity ($\theta_{222} = -27.0 \times 10^3\ \text{deg cm}^2\ \text{dmol}^{-1}$) was 79%, which is consistent with the value of 81% calculated based on the design. To estimate the expected helicity, we used SYNZIP crystal structures²⁰ and homology models (see Methods), and assumed that the linker regions connecting SYNZIPs do not adopt a helical conformation. Unfolding with heat showed a cooperative structural transition with a melting temperature (T_m) of $\sim 50\ ^\circ\text{C}$ (Fig. S9), which is slightly higher than the T_m of $47\ ^\circ\text{C}$ for 1:2 ($42\ ^\circ\text{C}$ for 3:4 and $32\ ^\circ\text{C}$ for 5:6)²⁰. In contrast, each of the unmixed, annealed fusion proteins showed less ellipticity at 222 nm than the nanotriangle (Fig. 3C); the low solubility of 3–2 may contribute to the lower CD intensity for this protein. Interestingly, 5–4 showed evidence of a partially helical structure, $\theta_{222} = -21.4 \times 10^3\ \text{deg cm}^2\ \text{dmol}^{-1}$, despite the fact that interactions were not previously reported between SYNZIPs 5 and 4. A helical structure is apparently populated by joining these sequences in the same chain. Although we did not perform detailed characterization of folded 5–4, SEC experiments indicate that this protein elutes as expected for a dimer that is mixed with, or exchanging with, a smaller population of monomer (Fig. S10).

The molecular weight determined by analytical ultracentrifugation (AUC) was consistent with the value expected for the designed heterotrimer. Sedimentation velocity analysis was performed on samples at concentrations of 13, 25, and $38\ \mu\text{M}$, and we calculated sedimentation coefficient distributions $c(s)$ ³³. The average s value for peaks measured at different concentrations was $2.5 \pm 0.1\ \text{S}$ (Fig. 3D), corresponding to a molecular weight of $34.9 \pm 1.0\ \text{kDa}$. The deviation from the value predicted based on protein sequences (35.8 kDa) was less than 3% (Table S3). We also determined the molecular weight from SAXS data using an analysis that makes no

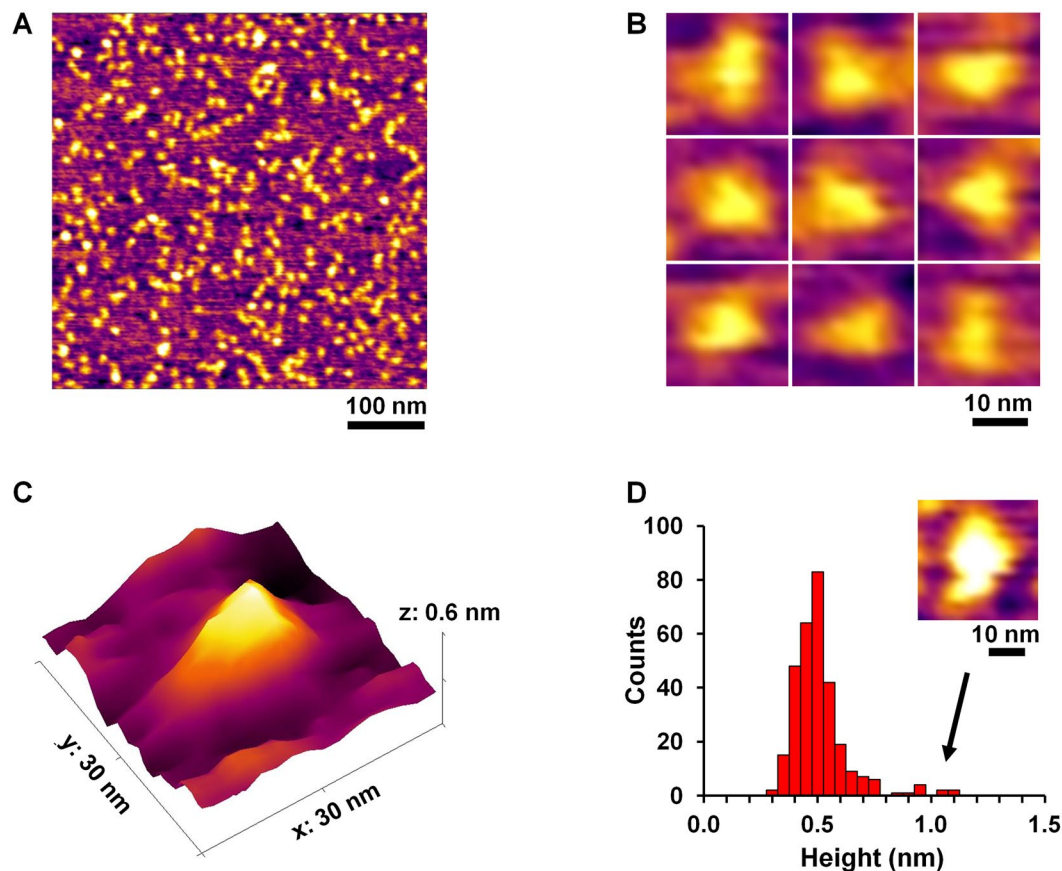


Figure 4. Characterization of the shape of the protein nanotriangle by AFM imaging. Height images of the triangles at low (A) and high magnifications (B), and a three-dimensional AFM image of a single assembly (C). (D) A histogram of the heights of the particles shown in panel A, and an image of a “tall” particle with height of ~ 1 nm (inset).

assumptions about shape (Fig. 3B). The method we employed defines and uses the volume of correlation, V_c , a SAXS invariant derived from the scattered intensities³⁴. The molecular weight determined using this approach was 32.6 kDa, which is within the reported error range of this method ($\sim 10\%$)³⁴.

We used atomic force microscopy (AFM) to image the designed protein assembly. Height images showed monodisperse nano-objects (Fig. 4A), consistent with the d_H value determined by DLS (Fig. 3A). Imaged objects appeared triangular in close-up images (Fig. 4B), with an average height of 0.51 ± 0.12 nm (Fig. 4C and D). The unimodal distribution indicated that the nanotriangles were uniform in height and thus likely to be discrete single particles (Fig. 4D). We observed a few tall particles with heights of ~ 1 nm, which we assumed to be aggregates or overlapped nanotriangles (see Fig. 4D inset). The average height for imaged particles was close to the previously measured height of a dimeric coiled coil on a mica substrate (~ 0.6 nm)³⁵.

Structural modeling with SAXS data. To further investigate the conformation of the protein nanotriangle in solution, we built all-atom models representing many different conformations and tested which models were consistent with the SAXS data. Using the crystal structures of 1:2 and 5:6²⁰ and the homology model of 3:4, we built models of the heterotrimer as described in the methods and, for each model, compared the predicted SAXS profiles for that structure to the experimental data³⁶. We assessed agreement using the χ value, which has a value of 1.0 for models that fit the data within the accuracy of the experimental noise and larger values for less-good fits. Among 1996 diverse comparative models, which included highly collapsed and maximally expanded arrangements of coiled coils, the distribution of R_g values ranged from 26 to 47 Å, and χ values ranged from 1.01 to 3.77 (see Figs 5A and S12). Guinier analysis of the SAXS data gave $R_g = 36.0 \pm 2.5$ Å (Fig. 3B), and it is apparent in Fig. S12 that most of the models with $\chi \leq 1.05$ had R_g values consistent with that range of values. Fig. 5C shows two models (*i* and *ii*) that are an excellent match to the data, with χ values of 1.01 and 1.03 and R_g values of 36.7 and 35.2 Å, respectively (Fig. 5B). In contrast, Fig. 5D shows models (*iii* and *iv*) with extreme values of R_g that had high χ values (2.46 and 3.77) and were clearly not consistent with the experimental SAXS profile shown in Fig. 5B. Given the flexible linkers used in the design, the triangle structure is best described as an ensemble of many different conformations. A two-structure ensemble consisting of 64% of model *i* and 36% of model *ii* agrees with the experimental observations within the noise level. Many other ensembles would also be consistent with the SAXS observations. However, our analysis rules out high

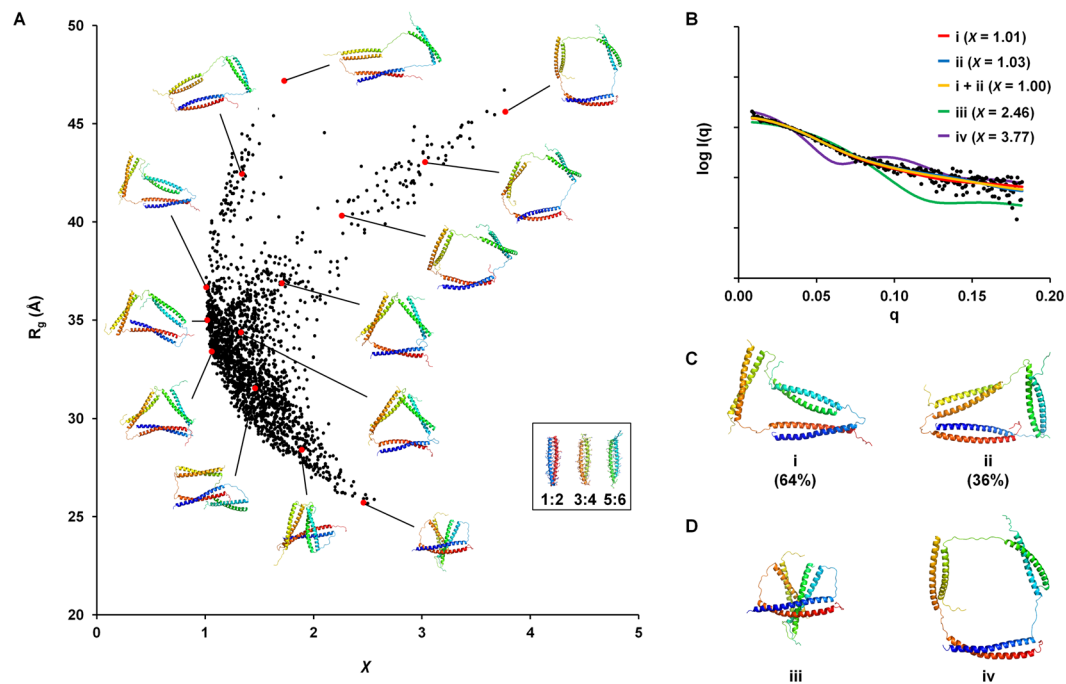


Figure 5. Structural models of the trimeric assembly and their agreement with experimental SAXS data. (A) R_g values for 1996 models plotted against χ values. Structures for representative models are shown next to the corresponding red dots. (B) Experimental data (black dots) and single-conformation or ensemble fits for structure models shown in panels C and D (colored lines). (C) Models with low χ values (*i* and *ii*) and the weights that define the ensemble in panel C. (D) Models with high χ values (*iii* and *iv*).

occupancy of extremely compact structures in which the coiled-coil units themselves are strongly associating (as in Fig. 5D, *iii*), and also shows that the ensemble is not composed exclusively of highly expanded structures (Fig. 5A).

Discussion

We designed and characterized a monodisperse, two-dimensional protein nanotriangle with a characteristic dimension of ~ 10 nm. To our knowledge, this is the first confirmed design of a protein structure with this topology. In prior work³⁰, supramolecular assemblies were designed using a single type of heterodimeric coiled coil. Linkers of variable lengths were used to control the mode of assembly, from fibers to discrete nanoscale objects. One of the designs was predicted to adopt a triangular shape, although its shape was not directly characterized³⁰.

In our simple strategy, self-assembly of a protein nanostructure was directed by the modularity of protein association, and a triangle structure was formed efficiently using a mixing and annealing procedure. Although we encountered some obstacles, such as unexpected folding of 5–4, and limited solubility of unassembled proteins, these were overcome by modulating the annealing process and increasing the relative concentration of the poorly soluble component 3–2. Flexibility in mixing and annealing procedures are advantages of our approach that here allowed us to realize a successful design. Our strategy has other advantages compared to alternatives for supramolecular assembly. For example, folding of a single protein chain into a tetrahedron, in prior work, was directed by six different coiled-coil modules, with the structure encoded in the ordered arrangement of sequences in the protein chain²⁹. The fact that intra- over intermolecular interaction is favored at high dilution allowed for the formation of a complex structure. Micromolar concentrations lead to aggregation, as expected due to the propensity for self-association that is part of the design, and the strategy is not easily generalized. The design of cages or arrays using globular oligomeric domains requires symmetry matching, along with geometrical alignment of the domains using rigid^{8–10} or flexible linkers with optimal spacing^{11,13}. This is non-trivial to accomplish, and small differences can lead to the wrong stoichiometry⁹ or an infinite rather than finite assembly³⁰. Finally, for some applications, it may be advantageous to trigger assembly at a specific time; in our scheme, this can be accomplished by mixing the components at a controlled time.

In this work, the pairwise interactions of three parallel, heterodimeric SYNZIPs were sufficient to direct supramolecular assembly. 23 SYNZIPs have been described^{19,20}, and binding pairs such as 10:22 and 17:18 could be useful as additional modules because their interactions are reported as stronger than potentially competing associations with any of SYNZIPs 1–6. Furthermore, synthetic coiled coils beyond SYNZIPs, as well as large sets of coiled-coil homo- and heterodimers from animal bZIP transcription factors that have been comprehensively tested for associations provide candidates for extended sets of orthogonal coiled coils^{17,18,23,37,38}. With larger sets of coiled-coil modules that display strong interaction preferences, our approach may be extended to design more complex structures that can potentially be linked together in 2- or 3-dimensions, functionalized, attached to surfaces or incorporated with other designed elements into increasingly complex nanoassemblies.

Methods

Design of a protein nanotriangle. We designed a protein nanotriangle assembled from three chains of genetically linked SYNZIPs^{19–21}. There are 16 candidate designs of this type that differ in which SYNZIPs are connected (Fig. S2A). We excluded designs that linked hetero-associating SYNZIPs, because such constructs could form complexes such as homo-oligomers and/or fibers that would compete with triangle assembly³⁰. SYNZIP pairs 4:6, 3:5, 3:6, and 1:5 are also known to interact weakly (Fig. 1A). Although the affinities of these pairs are orders of magnitude lower than the binding pairs (1:2, 3:4, and 5:6), the interactions may become significant when protein chains are linked together, increasing effective concentrations and potentially leading to kinetic traps. For example, linked-SYNZIP fusion proteins 4–6 or 3–5 can potentially form homodimers, with remaining coiled-coil segments forming heterodimers such as 5–2/1–3 (or 4–2/1–6); Fig. S2B illustrates some of these possibilities. We therefore excluded designs containing 3–5, 5–3, 4–6, 6–4, 3–6, 6–3, 1–5, or 5–1. Among the remaining four designs (in the top row of Fig. S2A), we chose to make the three chains 1–6, 5–4, and 3–2. A flexible linker of tetraglycine-serine repeats (Gly₄Ser)₂ was placed between SYNZIPs, and an affinity purification tag (His₆) was placed at the C-terminus of each polypeptide (Fig. 1D). The final sequences are given in Table S2.

Construction of plasmids. We used standard molecular biology techniques to construct three plasmids that express the linked-SYNZIP fusion proteins. The genes encoding SYNZIPs were PCR-amplified using forward and reverse primers listed in Table S1. The amplified fragments were assembled into the designed DNA constructs, which were further amplified by PCR. After digestion by the restriction enzymes *NdeI* and *XhoI* (New England Biolabs), the inserts were ligated into an expression vector, pET-43.1a (Novagen). The resulting plasmids were sequenced to confirm insertion. Plasmids encoding individual SYNZIPs are available via AddGene (www.addgene.org).

Protein expression and purification. Plasmids were separately transformed into *E. coli* strain BL21(DE3) cells (Agilent). Cell cultures for each fusion protein (1 L) were grown at 37 °C in Luria-Bertani liquid medium containing ampicillin (100 mg/L). At an optical density at 600 nm (OD₆₀₀) of 0.6, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce protein expression (final concentration 1.0 mM). After 5 hours at 37 °C, cells were harvested by centrifugation. The harvested cells were resuspended in lysis buffer (8 M urea, 10 mM Tris-HCl, and 100 mM Na₂HPO₄ pH 8.0) and lysed by a cycle of freezing-thawing and sonication. The cell lysate was cleared by centrifugation, and incubated with nickel-nitrilotriacetic acid resin (Qiagen). In buffers containing 8 M urea, 10 mM Tris-HCl, and 100 mM Na₂HPO₄, the fusion proteins were washed at pH 6.3 and collected by elution at pH 4.5. The purified proteins were dialyzed into deionized water.

Size-exclusion chromatography (SEC). The annealed mixture of linked-SYNZIP fusion proteins was prepared at a total protein concentration of 100 μM (molar ratio of 1–6, 5–4, and 3–2 was 1:1:3) in 50 mM Tris-HCl pH 7.4, 150 mM NaCl. Samples (10 mL) were run on a fast protein liquid chromatography system over a Superdex 75 26/60 column (GE Healthcare) at a flow rate of 0.5 mL/min. Purity was confirmed by native-polyacrylamide gel electrophoresis (Native-PAGE).

Circular dichroism (CD) spectroscopy. CD spectra were recorded on an AVIV 420 spectropolarimeter (Aviv Biomedical, Inc.). Protein solution samples were prepared in 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA). Measurement was performed at 5 μM of each protein (15 μM total protein) in a 0.1-cm-length cuvette. The spectra were obtained at 20 °C in 1 nm increments in a wavelength range from 200 to 260 nm, averaging for 30 s at each wavelength. The α-helical content was estimated from the mean residual ellipticity at 222 nm, using the following equation³⁹:

$$\text{Helical content (\%)} = ([\theta]_{\text{obs}} \times 100) / \{[\theta]_{\text{helix}} \times (1 - 2.57/n)\}$$

where $[\theta]_{\text{obs}}$ and $[\theta]_{\text{helix}}$ are the ellipticities of a helix of n and infinite residues, respectively. We used a $[\theta]_{\text{helix}}$ value of $-34,546 \text{ deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$, which we computed based on the CD signal for SYNZIP1:SYNZIP2 and the number of helical residues observed in crystal structure (PDB ID: 3HE5)²⁰. The thermal unfolding curve was determined by measuring ellipticity at 222 nm from 10 to 80 °C in 2 °C steps, with an averaging time of 30 s and an equilibration time of 1.5 min.

Dynamic light scattering (DLS). The hydrodynamic diameter of the protein nanotriangle was measured on a DynaPro NanoStar (Wyatt Technology), with a laser operating at a wavelength of 658 nm and at a detection angle of 90°. Protein solution (100 μL) was prepared in a cuvette, and measurements were performed at 25 °C. The raw correlation data were processed to generate a size distribution using DYNAMICS Software (Wyatt Technology).

Analytical ultracentrifugation (AUC). Sedimentation velocity AUC was performed using a Beckman XL-I centrifuge with interference optics. Protein samples were prepared at concentrations of 13, 25, and 38 μM in 50 mM Tris-HCl pH 7.4, 150 mM NaCl, and 1 mM EDTA. The samples were dialyzed against the same buffer overnight. All measurements were made at 20 °C with a rotor speed of 42,000 rpm. Data were analyzed using SEDFIT (<http://www.analyticalultracentrifugation.com>)³³, and sedimentation coefficient distributions $c(s)$ were obtained. Parameters required for data analysis (protein partial specific volume, buffer viscosity, and buffer density) were calculated using the SEDNTERP web server (Biomolecular Interaction Technologies Center, <http://sednterp.unh.edu>).

Atomic force microscopy (AFM). Freshly cleaved mica was incubated with poly-L-lysine (Sigma, MW 1,000–5,000) dissolved in deionized water (0.015 w/v %) for 10 min. The protein nanotriangle in 50 mM Tris-HCl pH 7.4, 150 mM NaCl, and 1 mM EDTA, and at a concentration of 500 nM, was deposited onto the mica treated with poly-L-lysine for 1 min. The mica was washed with 1 mL of MilliQ-filtered deionized water and blown dry in a stream of nitrogen gas.

Imaging was carried out using the Asylum Research Cypher microscope at the Center for Nanoscale Systems (CNS) at Harvard University. Imaging was done in air for topography measurement, using tapping mode in the repulsive interaction regime. Tapping parameters were tuned such that the tip started tracking the surface in the attractive interactions regime, and then the strength of interactions was increased gradually, into the repulsive regime, at which point high-resolution AFM images were obtained. We used AFM tip AC240BSA-R3 from Asylum Research ($f = 75$ kHz, $k = 2$ N/m).

Biological small-angle X-ray scattering (BioSAXS). The protein nanotriangle solution samples were prepared in 50 mM Tris-HCl pH 7.4, 150 mM NaCl, and 1 mM EDTA, and at concentrations of 11, 22, and 44 μ M. Synchrotron X-ray scattering data was collected at the G-line of the Cornell High Energy Synchrotron Source (G1)^{40,41}. Area Detector System CCD was used to collect the scattering patterns, and data analysis was done using BioXTAS RAW (version 1.0.0) to determine R_g and molecular weight.

Modeling. A comparative protein structure model for 3:4 was built with MODELLER³¹, using structures of coiled coils (PDB ID: 1KD8, 1KD9⁴², and 3HE4²⁰) as templates. Using this model as a starting structure, molecular dynamics (MD) simulations were performed using the NAMD 2.11 package⁴³ and the CHARMM22 all-atom force field⁴⁴. The protein nanotriangle was solvated within a $60.7 \times 95.2 \times 85.2$ Å³ water box, using periodic boundary conditions. The system was simulated in the constant temperature and pressure ensemble at 298 K for over 10 ns.

Comparative structure models for the protein nanotriangle were built using MODELLER³¹ and fit to experimental SAXS data using the FoXS/MES web server software (<https://modbase.compbio.ucsf.edu>)³⁶. As templates for modeling, we used x-ray structures for 1:2 and 5:6 (PDB ID: 3HE5 and 3HE4)²⁰. For 3:4, we used the two coiled-coil structures 1KD8 and 1KD9⁴². We confirmed that models of the coiled-coil part of 3:4, in the context of the triangle models, agreed well with structures from the MD simulation (at 10 ns) shown in Fig. S1 (backbone atom RMSD < 3.5 Å for 1984 models). No spatial restraints were applied to the flexible linkers between SYNZIPs. Numerous models were generated by choosing different, random initial positions for the coiled coils in the starting templates, and then refining the models using molecular dynamics with simulated annealing. The end-to-end distance (d_{linker}) of the 1996 resulting structures ranged from 3 to 33 Å. The 1996 resulting structures were fit to the SAXS data, and the quality of the fit was measured by the χ function³⁶:

$$\chi = \sqrt{\frac{1}{M} \sum_{i=1}^M \left(\frac{I_{\text{exp}}(q_i) - cI(q_i)}{\sigma(q_i)} \right)^2} \quad (2)$$

where M is the number of points in the SAXS profile, $I_{\text{exp}}(q_i)$ and $I(q_i)$ are the experimental and computed profiles, c is the scale factor, and $\sigma(q_i)$ is the experimental error that represents standard deviation of intensity values. Various combinations of models with $\chi \leq 1.05$ were selected and used to fit ensembles to the SAXS data using ensemble search software (MES)⁴⁵. Scattering intensities of the multiple conformations of a minimal ensemble were computed by averaging individual scattering patterns.

References

- Luo, Q., Hou, C., Bai, Y., Wang, R. & Liu, J. Protein assembly: Versatile approaches to construct highly ordered nanostructures. *Chem. Rev.* **116**, 13571–13632, doi:10.1021/acs.chemrev.6b00228 (2016).
- Zhang, S. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* **21**, 1171–1178, doi:10.1038/nbt874 (2003).
- Gonen, S., DiMaio, F., Gonen, T. & Baker, D. Design of ordered two-dimensional arrays mediated by noncovalent protein-protein interfaces. *Science* **348**, 1365–1368, doi:10.1126/science.aaa9897 (2015).
- King, N. P. *et al.* Computational design of self-assembling protein nanomaterials with atomic level accuracy. *Science* **336**, 1171–1174, doi:10.1126/science.1219364 (2012).
- King, N. P. *et al.* Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103–108, doi:10.1038/nature13404 (2014).
- Bale, J. B. *et al.* Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science* **353**, 389–394, doi:10.1126/science.aaf8818 (2016).
- Hsia, Y. *et al.* Design of a hyperstable 60-subunit protein icosahedron. *Nature* **535**, 136–139, doi:10.1038/nature18010 (2016).
- Padilla, J. E., Colovos, C. & Yeates, T. O. Nanohedra: Using symmetry to design self-assembling protein cages, layers, crystals, and filaments. *Proc. Natl. Acad. Sci. USA* **98**, 2217–2221, doi:10.1073/pnas.041614998 (2001).
- Lai, Y.-T. *et al.* Structure of a designed protein cage that self-assembles into a highly porous cube. *Nat. Chem.* **6**, 1065–1071, doi:10.1038/nchem.2107 (2014).
- Lai, Y.-T., Cascio, D. & Yeates, T. O. Structure of a 16-nm cage designed by using protein oligomers. *Science* **336**, 1129–1129, doi:10.1126/science.1219351 (2012).
- Sciore, A. *et al.* Flexible, symmetry-directed approach to assembling protein cages. *Proc. Natl. Acad. Sci. USA* **113**, 8681–8686, doi:10.1073/pnas.1606013113 (2016).
- Kobayashi, N. *et al.* Self-assembling nano-architectures created from a protein nano-building block using an intermolecularly folded dimeric de novo protein. *J. Am. Chem. Soc.* **137**, 11285–11293, doi:10.1021/jacs.5b03593 (2015).
- Sinclair, J. C., Davies, K. M., Venien-Bryan, C. & Noble, M. E. M. Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nat. Nanotechnol.* **6**, 558–562, doi:10.1038/nnano.2011.122 (2011).

14. Laganowsky, A. *et al.* An approach to crystallizing proteins by metal-mediated synthetic symmetrization. *Protein Sci.* **20**, 1876–1890, doi:10.1002/pro.727 (2011).
15. Mason, J. M. & Arndt, K. M. Coiled coil domains: Stability, specificity, and biological implications. *ChemBioChem* **5**, 170–176, doi:10.1002/cbic.200300781 (2004).
16. Grigoryan, G. & Keating, A. E. Structural specificity in coiled-coil interactions. *Curr. Opin. Struct. Biol.* **18**, 477–483, doi:10.1016/j.sbi.2008.04.008 (2008).
17. Gradišar, H. & Jerala, R. De novo design of orthogonal peptide pairs forming parallel coiled-coil heterodimers. *J. Pept. Sci.* **17**, 100–106, doi:10.1002/psc.1331 (2011).
18. Fletcher, J. M. *et al.* A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology. *ACS Synth. Biol.* **1**, 240–250, doi:10.1021/sb300028q (2012).
19. Thompson, K. E., Bashor, C. J., Lim, W. A. & Keating, A. E. Synzip protein interaction toolbox: *In vitro* and *in vivo* specifications of heterospecific coiled-coil interaction domains. *ACS Synth. Biol.* **1**, 118–129, doi:10.1021/sb200015u (2012).
20. Reinke, A. W., Grant, R. A. & Keating, A. E. A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering. *J. Am. Chem. Soc.* **132**, 6025–6031, doi:10.1021/ja907617a (2010).
21. Grigoryan, G., Reinke, A. W. & Keating, A. E. Design of protein-interaction specificity gives selective bzip-binding peptides. *Nature* **458**, 859–864, doi:10.1038/nature07885 (2009).
22. Negron, C. & Keating, A. E. A set of computationally designed orthogonal antiparallel homodimers that expands the synthetic coiled-coil toolkit. *J. Am. Chem. Soc.* **136**, 16544–16556, doi:10.1021/ja507847t (2014).
23. Crooks, R. O., Baxter, D., Panek, A. S., Lubben, A. T. & Mason, J. M. Deriving heterospecific self-assembling protein–protein interactions using a computational interactome screen. *J. Mol. Biol.* **428**, 385–398, doi:10.1016/j.jmb.2015.11.022 (2016).
24. Papapostolou, D. *et al.* Engineering nanoscale order into a designed protein fiber. *Proc. Natl. Acad. Sci. USA* **104**, 10853–10858, doi:10.1073/pnas.0700801104 (2007).
25. Fletcher, J. M. *et al.* Self-assembling cages from coiled-coil peptide modules. *Science* **340**, 595–599, doi:10.1126/science.1233936 (2013).
26. Park, W. M. & Champion, J. A. Thermally triggered self-assembly of folded proteins into vesicles. *J. Am. Chem. Soc.* **136**, 17906–17909, doi:10.1021/ja5090157 (2014).
27. Burgess, N. C. *et al.* Modular design of self-assembling peptide-based nanotubes. *J. Am. Chem. Soc.* **137**, 10554–10562, doi:10.1021/jacs.5b03973 (2015).
28. Der, B. S. & Kuhlman, B. Cages from coils. *Nat. Biotechnol.* **31**, 809–810, doi:10.1038/nbt.2670 (2013).
29. Gradišar, H. *et al.* Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nat. Chem. Biol.* **9**, 362–366, doi:10.1038/nchembio.1248 (2013).
30. Boyle, A. L. *et al.* Squaring the circle in peptide assembly: From fibers to discrete nanostructures by de novo design. *J. Am. Chem. Soc.* **134**, 15457–15467, doi:10.1021/ja3053943 (2012).
31. Webb, B. & Sali, A. In *Current Protocols in Bioinformatics* **47**, 5.6.1–5.6.32 (John Wiley & Sons, Inc., 2014).
32. Phillips, J. C. *et al.* Scalable molecular dynamics with namd. *J. Comput. Chem.* **26**, 1781–1802, doi:10.1002/jcc.20289 (2005).
33. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619, doi:10.1016/S0006-3495(00)76713-0 (2000).
34. Rambo, R. P. & Tainer, J. A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481, doi:10.1038/nature12070 (2013).
35. van Noort, J. *et al.* The coiled-coil of the human rad50 DNA repair protein contains specific segments of increased flexibility. *Proc. Natl. Acad. Sci. USA* **100**, 7581–7586, doi:10.1073/pnas.1330706100 (2003).
36. Schneidman-Duhovny, D., Hammel, M., Tainer, John, A. & Sali, A. Accurate saxs profile computation and its assessment by contrast variation experiments. *Biophys. J.* **105**, 962–974, doi:10.1016/j.bpj.2013.07.020 (2013).
37. Reinke, A. W., Baek, J., Ashenberg, O. & Keating, A. E. Networks of bzip protein–protein interactions diversified over a billion years of evolution. *Science* **340**, 730–734, doi:10.1126/science.1233465 (2013).
38. Chen, R. *et al.* A barcoding strategy enabling higher-throughput library screening by microscopy. *ACS Synth. Biol.* **4**, 1205–1216, doi:10.1021/acssynbio.5b00060 (2015).
39. Chen, Y.-H., Yang, J. T. & Chau, K. H. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry* **13**, 3350–3359, doi:10.1021/bi00713a027 (1974).
40. Acerbo, A. S., Cook, M. J. & Gillilan, R. E. Upgrade of macchess facility for x-ray scattering of biological macromolecules in solution. *J. Synchrotron Radiat.* **22**, 180–186, doi:10.1107/S1600577514020360 (2015).
41. Skou, S., Gillilan, R. E. & Ando, N. Synchrotron-based small-angle x-ray scattering of proteins in solution. *Nat. Protocols* **9**, 1727–1739, doi:10.1038/nprot.2014.116 (2014).
42. Keating, A. E., Malashkevich, V. N., Tidor, B. & Kim, P. S. Side-chain repacking calculations for predicting structures and stabilities of heterodimeric coiled coils. *Proc. Natl. Acad. Sci. USA* **98**, 14825–14830, doi:10.1073/pnas.261563398 (2001).
43. Kalé, L. *et al.* Namd2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151**, 283–312, doi:10.1006/jcph.1999.6201 (1999).
44. Brooks, B. R. *et al.* Charmm: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217, doi:10.1002/jcc.540040211 (1983).
45. Pelikan, M., Hura, G. L. & Hammel, M. Structure and flexibility within proteins as identified through small angle x-ray scattering. *Gen. Physiol. Biophys.* **28**, 174–189, doi:10.4149/gpb_2009_02_174 (2009).

Acknowledgements

We thank D. Pheasant and R. Rezaei Araghi for help with experiments, V. Xue, and V. Frappier for help with modeling, and members of the Keating lab for comments on the work and the manuscript. We thank J. Hopkins and R. Gillilan from the Cornell High Energy Synchrotron Source (CHESS) at Cornell University for help with collecting the BioSAXS data. CHESS is supported by the NSF & NIH/NIGMS via NSF award DMR-1332208, and the MacCHESS resource is supported by NIH/NIGMS award GM-103485. We thank J. Tresback at the Center for Nanoscale Systems at Harvard University for help with collecting AFM images. The Biophysical Instrumentation Facility for the Study of Complex Macromolecular Systems (NSF-0070319) is gratefully acknowledged. This research was financially supported by DuPont and by a Professor Amar G. Bose Research Grant to A.K. and K.B.

Author Contributions

W.M.P. and A.E.K. designed the research. W.M.P. and M.B. performed the experiments, and W.M.P. performed the modeling. W.M.P., M.B., K.K.B., and A.E.K. analyzed the data and wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-10918-6](https://doi.org/10.1038/s41598-017-10918-6)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017