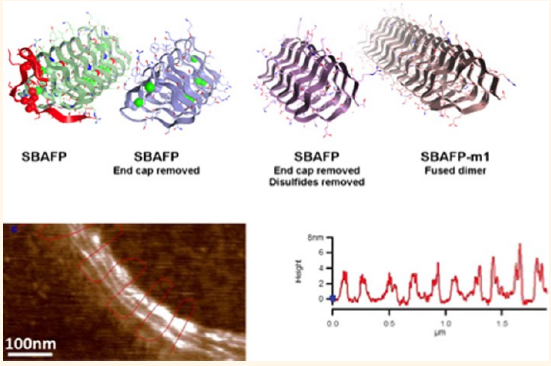
Engineering Amyloid Fibrils from β-Solenoid Proteins for Biomaterials Applications

Maria D.R. Peralta,[†] Arpad Karsai,[†] Alice Ngo,[†] Catherine Sierra,[†] Kai T. Fong,[†] Natha Robert Hayre,[‡,§] Nima Mirzaee,[‡,§] Krishnakumar Mayuram Ravikumar,[‡,§] Alexander J. Kluber,[§,∥] Xi Chen,[†] Gang-yu Liu,[†] Michael D. Toney,*[,†] Rajiv. R. Singh,[‡,§] and Daniel Lee Cox*[,‡,§]

[†]Department of Chemistry, [‡]Department of Physics, and [§]Institute for Complex Adaptive Matter, University of California, 1 Shields Avenue, Davis, California 95616, United States and [∥]Department of Chemistry, Rice University, Space Science 201, Houston, Texas 77251, United States

ABSTRACT Nature provides numerous examples of self-assembly that can potentially be implemented for materials applications. Considerable attention has been given to one-dimensional cross-β or amyloid structures that can serve as templates for wire growth or strengthen materials such as glue or cement. Here, we demonstrate controlled amyloid self-assembly based on modifications of β-solenoid proteins. They occur naturally in several contexts (e.g., antifreeze proteins, drug resistance proteins) but do not aggregate *in vivo* due to capping structures or distortions at their ends. Removal of these capping structures and regularization of the ends of the spruce budworm and rye grass antifreeze proteins yield micron length amyloid fibrils with predictable heights, which can be a platform for biomaterial-based self-assembly. The design process, including all-atom molecular dynamics simulations, purification, and self-assembly procedures are described. Fibril formation with the predicted characteristics is supported by evidence from thioflavin-T fluorescence, circular dichroism, dynamic light scattering, and atomic force microscopy. Additionally, we find evidence for lateral assembly of the modified spruce budworm antifreeze fibrils with sufficient incubation time. The kinetics of polymerization are consistent with those for other amyloid formation reactions and are relatively fast due to the preformed nature of the polymerization nucleus.



KEYWORDS: amyloid · antifreeze protein · β-solenoid · atomic force microscopy · thioflavin-T fluorescence

An important goal of nanotechnology is bottom-up manufacturing of useful devices and materials via self-assembly at room temperature in environmentally benign solvents. Living systems provide numerous examples of such self-assembly in the guise, for example, of protein structures such as microtubules,[1] viral capsids,[2] bacterial s-layers,[3] and amyloid fibrils[4] in which proteins grow in one-dimensional filaments with β-strands perpendicular to the growth axis.

The programmable design of DNA-based nanostructured scaffolds is extraordinary,[5] allowing for the templating of ordered heterogeneous arrays of, e.g., metallic nanoparticles,[6] proteins,[7] and semiconducting wires.[8] However, it is plagued with technical barriers to advancement, including: (a) difficulties in scaling it to industrial applications, (b) high error rates of DNA replication, (c) denaturation of DNA scaffolds/bundles at moderate temperatures (~60 °C), and (d) loss of integrity under exposure to ultraviolet light and enzymes.[9]

Belcher and collaborators have been highly successful at using the M13 virus as a scaffold for self-assembly of a wide variety of inorganic materials. Their strategy relies on modifying coat proteins with peptides that are selected through phage display for templating a specific material.[10] In one example, the M13 major coat protein was coated by a peptide with $FePO_4$-nanoparticle templating activity, while the attachment proteins at the end of the virus were fused to a peptide known to adhere to carbon nanotubes.[11] Incubation of the virus with iron and phosphate ions together with single-walled carbon nanotubes generated

* Address correspondence to dlcox@ucdavis.edu, mdtoney@ucdavis.edu.

Received for review October 2, 2014 and accepted January 6, 2015.

Published online January 06, 2015
10.1021/nn5056089

a self-assembled working cathode. However, the M13 approach is limited by several factors: (a) viruses are large (M13 is nearly a micron in length); (b) templating sites are limited to the coat proteins, and the geometry is restricted to that provided by the virus; and (c) while the viruses can order as liquid crystals, the ordering is on the micron scale. Hence, precise, programmable nanometer-scale ordered heterogeneity, as achieved with DNA, is not feasible with viruses.

Amyloid fibrils are self-assembled one-dimensional protein arrays with β-strands perpendicular to the linear axis.[4] They arise both in unregulated self-assembly in numerous diseases including Alzheimer's disease and type II diabetes, as well as in regulated contexts in biofilm extracellular matrices,[12] synapse formation,[13] and hormone reservoir manufacture.[14] These fibrils have bending and twisting persistence lengths on the micron scale,[15,16] which contribute to the remarkable tensile strength of spider silk[17] and the structural stability of barnacle cement.[18] They have previously been used to template metallic nanowire growth[19−21] and have been used to produce mechanically strong oriented films.[22]

Amyloid structures are remarkably robust. Generally, they can survive heating to the boiling point of water[23−25] although there is monomer size and sequence dependence to this result. They are resistant to protease degradation[26,27] and UV light exposure. To date, amyloids have not been assembled to produce a significant level of transverse order, nor have they been used to template material growth other than the examples given above. There is also little systematic understanding of amyloid structure because the lack of transverse order makes it difficult for X-ray diffraction to reveal more than the generic cross β-stacking,[28] although in some instances additional scattering rings in fiber diffraction have provided information about transverse dimensions of fibrils and longer periodicity repeats along the fiber axis.[29]

Here, we present a new approach to amyloid design that may eventually allow programmable nanoscale structural precision for self-assembly of materials under mild conditions. Our focus is on naturally occurring β-solenoid proteins (BSPs) such as those shown in Figure 1. These proteins have backbones that twist helically in either a left- or right-handed sense from the N-terminus to form β-sheets and have regular geometric structures (triangles, rectangles, *etc.*) with 1.5−2 nm sides. The wild type (WT) proteins are inhibited from amyloid aggregation (end-to-end polymerization to give cross β-fibrils) by natural capping features and/or structural distortions on one or both ends. We describe the modifications necessary to make linear polymers (amyloids) from these proteins, explain molecular simulations used to assess structural stability and geometric properties for comparison to measurements, and describe the protocol for expressing and folding of the engineered proteins. We show that the correct monomeric structures are obtained after purification and folding, that amyloid fibrils are produced by incubation at elevated temperatures, and that the kinetics of fibril formation are consistent with, though slightly faster than, other amyloid polymerization reactions. These conclusions are supported by measurements of circular dichroism (CD), thioflavin-T (ThT) fluorescence, dynamic light scattering (DLS), turbidity, and atomic force microscopy (AFM).

We note that these modified BSP proteins may offer excellent platforms for functionalization in nanotechnology without interfering with the native β-sheet structure. For example, the large area faces together with their designable length can, in principle, support templating peptides of more than one kind of nanoparticle to grow ordered heterogeneous nanoparticle arrays. Additionally, staggered placement of identical templating peptides control the nanoparticle aspect ratio. Even as one face is being used for nanoparticle templating, another can be used for binding to the surface or assuring designed lateral assembly of the fibrils. In contrast, strategies based upon small amyloidogenic peptides do not immediately offer this level of functionalization diversity.

RESULTS AND DISCUSSION

Most amyloid fibrils used in the literature for engineering purposes, e.g., templating nanowire growth,[19,20] are derived from naturally occurring proteins or peptides known to form amyloid fibrils under specific conditions. For example, nanowires have been previously grown on a prion variant from *Saccharomyces cerevisiae* known to self-assemble[19] as well as on an Alzheimer's β-amyloid diphenylalanine peptide.[20] These examples and various others in the literature[22] exploit naturally occurring amyloid fibrils. Approaches from other groups use harsh conditions (e.g., treatment with concentrated hydrochloric acid at elevated temperatures for various days[23]) to produce self-assembled amyloid fibrils out of intrinsically nonamyloidogenic proteins like lysozyme.[30] Instead of using naturally occurring peptides already known to self-assemble or exposing a protein to harsh conditions, we propose here a rational design concept to render intrinsically nonamyloidogenic proteins with natural cross-β-structure into proteins that readily self-assemble into amyloid fibrils under benign conditions.

Protein Design. *Spruce Budworm Antifreeze Protein.* Here, isozyme 501 of the β-solenoid antifreeze protein from the spruce budworm (SBAFP; PDB entry 1M8N)[31] was used to engineer one-dimensional fibrils. The polypeptide backbone is triangular about the long axis of the helix (Figure 1). The structurally homologous 2PIG PDB entry (Figure 1) shows that the left-handed

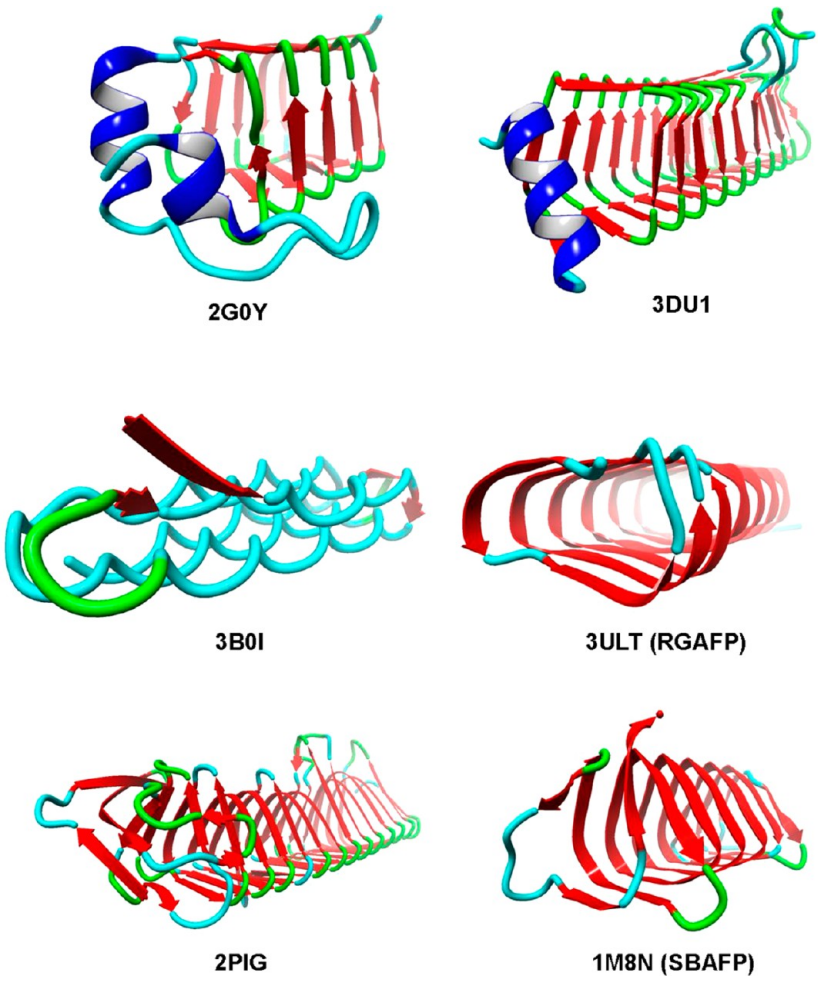


Figure 1. Examples of wild type β-solenoid proteins. PDB IDs are given below each. 2G0Y[68] and 3DU1[69] are four-sided pentapeptide repeat proteins. 3B0I[70] is a two-sided snow flea antifreeze protein. 3ULT[71] is the two-sided rye grass antifreeze protein abbreviated RGAFP herein. 2PIG[72] is the three-sided "type I" left-handed β-helical enzyme ydcK from *Salmonellae cholera*. 1M8N[73] is the three-sided "type II" left-handed β-helical solenoid antifreeze protein from the spruce budworm abbreviated SBAFP herein.

β-solenoid scaffold is tolerant to substitutions at the apexes of the triangular scaffold and therefore likely to be robust for materials engineering.

There were two major considerations in the rational design of the first engineered protein, termed SBAFP-m1. The first was seamless and stable end-to-end interactions, and the second was ease of biochemical handling. The first was addressed as follows and as illustrated in Figure 2. The β-hairpin-like capping motif at the C-terminus of WT SBAFP was removed to give a clean C-terminus in the β-strand conformation (Figure 2A). This required eliminating the last 21 residues. The N-terminus was modified by removing the first six amino acids present in the crystal structure of SBAFP. To avoid a heterogeneous N-terminus, Met-Ala were used for the first two amino acids of the SBAFP-m1 sequence. The *E. coli* methionine aminopeptidase has a strong preference for small amino acids such as alanine at the second position.[32] Thus, the Met-Ala sequence increases the likelihood that the N-terminal amino acid is homogeneously processed to Ala instead of a possible mixture of Met and a different second amino acid. This design gives a seamless interface between the N- and C-termini of successive monomers, as illustrated in Figure 2B.

Two additional salt bridges were placed at the interface of the termini to increase the stability of the interaction between monomers. These are illustrated in Figure 2B. As modeled, the end-to-end interface includes three salt bridge interactions: one between the N- and C-terminal ammonium and carboxylate groups, and two between the Arg-Glu side chain pairs introduced. These, along with the intermonomer β-sheet hydrogen bonding, proved sufficient to keep the interface structure stable in MD simulations at 100 °C for 2 ns, and for 20 ns runs at 25 °C as described further below.

WT SBAFP has a total of five disulfide bonds. These undoubtedly stabilize the folded protein, yet disulfides often present difficulties to high-level expression of recombinant proteins in *E. coli* and their subsequent handling. Nevertheless, recombinant expression of WT

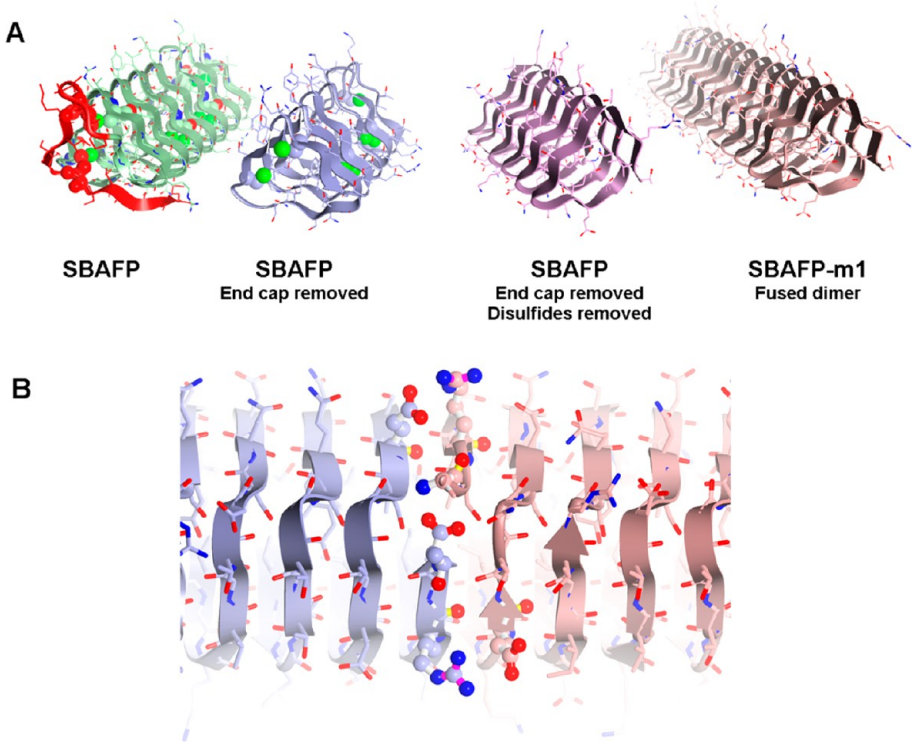


Figure 2. Stages in the design of SBAFP-m1. (A) The SBAFP wild type protein has a C-terminal capping motif shown in red. This was removed so that the N- and C-termini form a gapless interface when brought together. All cysteine residues (shown in space-filling on the second from the left) were changed to serines, eliminating the disulfide bonds. Two monomers were fused to form a larger protein that is more manageable genetically and biochemically. Finally, the monomer−monomer interface was optimized, including the addition of two Arg/Glu salt bridges. (B) Close-up of the optimized interface between two SBAFP-m1 proteins. The N- and C-termini form a salt bridge, as do the two Arg/Glu pairs placed at the interface. These are shown in ball-and-stick representation.

SBAFP in *E. coli* was previously achieved and reproduced here (see Methods),[33] but the expression levels are modest. Modeling the replacement of all the cysteines by serines showed a good steric fit and enabled hydrogen-bonding interactions to partially replace the disulfide bonds stabilizing interactions. An alignment of sequences summarizing deletions and mutations between WT SBAFP and SBAFP-m1 are presented in Figure 3A. Molecular dynamics (MD) simulations also showed the Cys-to-Ser mutant stable at 25 °C for 20 ns. Therefore, the Cys-to-Ser mutations were kept in the design. The cysteine-less design also has the advantage that future engineering efforts can employ the introduction of a unique cysteine to enable chemically specific modification of the protein.

The alterations described above left a relatively small protein of 90 amino acids. For ease of recombinant expression, purification, and functionalization of the protein, two of the 90 amino acid monomers were genetically fused together to give a single contiguous protein of 180 amino acids in length.

Ryegrass Antifreeze Protein. As illustrated in Figure 4, the same basic design concepts used for SBAFP-m1 were applied to a ryegrass (*Lolium perenne*) antifreeze left-handed β-solenoid protein (RGAFP, PDB entry 3ULT)[34] with the addition of using FoldIt[35] to optimize the interface. In contrast to SBAFP, which has a triangular cross section with three β-sheet faces, RGAFP has two β-sheet faces in a rectangular arrangement along the long axis of the β-helix. The structure has eight β-sheet rungs, each containing 14−15 amino acids per turn, and an exceptionally flat β-sheet on the ice-binding face. There is an additional amino acid present in the first three rungs, causing a bulge at the N-terminus of the protein structure (Figure 4A). These amino acids were removed to regularize the β-helix. Residues 1−4, which are missing in the crystal structure, were excluded from the designed protein, termed RGAFP-m1. Additionally, residue 5 (Pro), which might interfere with polymerization, was mutated to Ala, which additionally facilitates a homogeneous N-terminus as with the SBAFP-m1 design.

To engineer an idealized interface, a dimer model (Figure 4B) was used. Residue Lys110 appeared to hinder interface formation and hence was mutated to Asn. As with the design of SBAFP-m1, an Arg/Glu salt bridge was added at the interface. To further strengthen the interaction between monomers at the binding interface, FoldIt[35] optimization was used as follows. The last 16 residues of the first monomer and the first 16 residues of the second monomer were allowed to mutate to get the best FoldIt score. This

A

```
SBAFP:      1 DGTCVNTNSQITANSQCVKSTATNCYIDNSQLVDTSICTRSQYSDANVKKSVTTDCNIDK
SBAFP-m1:   1 .......ASRIT.NSQIVKSEATNSDINNSQLVD.SISTRSQYSDANVKKSVTTDSNIDK

SBAFP:     61 SQVYLTTCTGSQYNGIYIRSSTTTGTSISGPGCSISTCTI
SBAFP-m1:  52 SQVYLTTSTGSQYNGIYIRSSDTTGSEISGSSISTSRITI
```

B

```
RGAFP:      1 PNTISGSNNTVRSGSKNVLAGNDNTVISGDNNSVSGSNNTVVSGNDNTVTGSNHVVSGTN
RGAFP-m1:   1 ANDIDGTNNEVD-GSENVLAGNDNT-VSGDNNSVSGSNNT-VSGNDNTVTGSNHVVSGTN

RGAFP:     61 HIVTDNNNNVSGNDNNVSGSFHTVSGGHNTVSGSNNTVSGSNHVVSGSNKVVTD
RGAFP-m1:  58 HIVTDNNNNVSGNDNNVSGSFHTVSGGHNTVSGSNNTVSGKRHRVQGTNNRVTD
```

Figure 3. (A) Sequence alignment between WT SBAFP (PDB entry 1M8N) and the SBAFP-m1 protein derived from it. Only the first half of SBAFP-m1 (itself composed of two fused monomers) was used in the alignment. The last 21 residues of 1M8N were deleted in the design of SBAFP-m1, and the first Met is not shown in the alignment. (B) Alignment between WT RGAFP and RGAFP-m1, which has a total of three deletions and 13 mutations, compared to WT RGAFP.

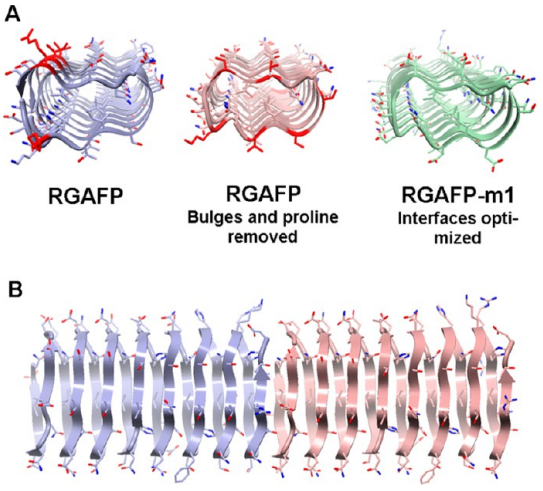


Figure 4. Stages in the design of the RGAFP-m1. (A) The N-terminal proline and amino acids causing the bulge (marked in red in the left structure) were deleted. Next, amino acids at the monomer interface (marked in red in the center structure) were mutated to optimize binding interactions. (B) Close up of end-to-end dimer interface of RGAFP-m1.

resulted in 10 mutations that largely increased the overall charge of the proteins at the interface. Thus, a total of 13 mutations, summarized in Figure 3B, were incorporated in the design. Further testing using MD simulations showed that the modeled interface was stable for 20 ns at 25 C.

Molecular Dynamics Simulations. The designed proteins were tested using MD simulations for 20 ns at constant pressure and temperature (25 °C) using AMBER 12[36] suite (see Methods). Both individual monomers and longer 11-unit multimers were analyzed for stability in MD simulations. Monomer simulations determined the inherent stability of each design, particularly with respect to the modifications of the sequence at the termini. A concern was that new steric and electrostatic interactions might lead to local instability of the β-helical motif, which could disrupt polymerization. In the case of SBAFP-m1, the monomer simulation also probed the viability of the novel intermonomer interface, with fused identical 90-residue segments used as proxies, where deregistry or layer separation of the ideal motif might have occurred.

Models of 11-unit multimers were also constructed to observe the behavior of ideal fibrils. Here, a sample of ten dimer interfaces could be observed simultaneously for defects or instability, along with the macroscale superhelical tendencies of a long polymeric fibril. Molecular dynamics simulations demonstrated that both designs were stable out to 20 ns. A measure of this is seen in Figure 5 where we plot the β-sheet content of simulated fibrils over time, averaged over five different runs for the SBAFP-m1 fibril and the RGAFP-m2 fibril. The latter is similar in design to the experimentally studied RGAFP-m1. The β-sheet content is determined by counting all residues within the β-sheet region of the Ramachandran plot (for further details see the Methods section) within the Visual Molecular Dynamics program (VMD).[37] For the idealized three-sided structure, this yields 80% β-sheet content (12 β-strand residues per 15 residue rung), and for the idealized two-sided structure this yields 70% (ten β-strand residues per 14 residue rung). The monomeric form retained all native β-helical contacts, including those corresponding to terminal sequence segments that one might expect to separate or fray. Simulations of polymer fibrils showed that all dimer interfaces remained intact, and imperfections in monomer docking during the model construction process became annealed into seamless registry.

Another role for the simulations is to provide analyses of possible height profiles of fibrils to

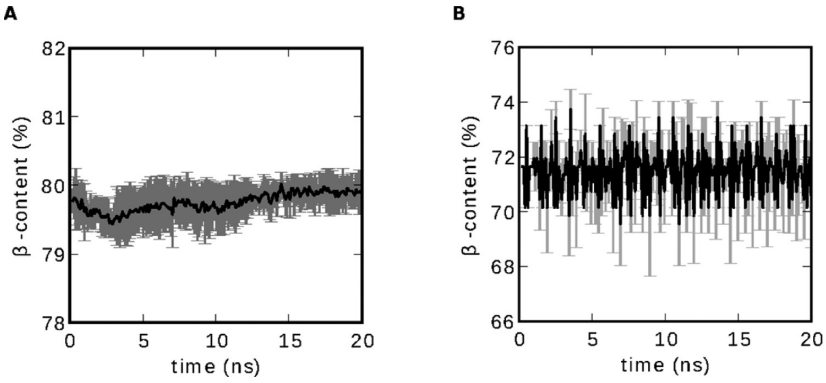


Figure 5. β-Sheet content vs time during fibril simulations. (A) SBAFP-m1. (B) RGAFP-m2. The data are averaged over five different runs each with random number seeds of the Langevin thermostat in the AMBER12 simulation suite.[36] The β-content was determined by counting residues in the β-sheet region of the Ramachandran plots using VMD.[37]

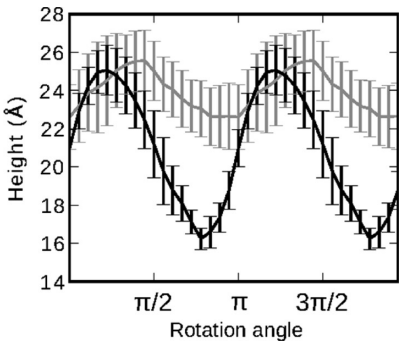


Figure 6. Height profile of SBAFP-m1 (gray) and RGAFP-m2 (black) monomers. For each case, the height minimum above a constraining surface corresponds to having a face parallel to and in contact with the surface. The maximum corresponds to having a line contact with an edge such that a face is perpendicular to the surface.

compare with the AFM topographic information. Note that this variation is mainly due to the triangular and rectangular cross sections of the proteins. Figure 6 shows height curves for monomers of SBAFP-m1 and RGAFP-m2 found using the untwisted monomers. The SBAFP-m1 monomer has a mean height ranging from 2.25 to 2.55 nm, while the RGAFP-m2 monomer has a height variation between 1.6 and 2.6 nm.

Taken together, the simulations, carried out in advance of the experimental fibril synthesis, portended the experimental finding that the designed β-solenoid proteins form stable amyloid fibrils and allowed us to exclude several intermediate designs in the process. Also, simulations provide important height profiles for comparison to AFM data to provide corroborating evidence that the observed fibrils have the desired structure.

Protein Expression, Purification, and Folding. The *E. coli* codon-optimized genes for both SBAFP-m1 and RGAFP-m1 were procured from Life Technologies. The SBAFP-m1 variant expressed well in *E. coli* BL21-(DE3) from the pET28a vector, although it was found almost completely in inclusion bodies (IBs). The protein was purified using a protocol of repeated IB washing, denaturation in 8.5 M urea, and purification by anion exchange chromatography (see Methods).[38] Purified, unfolded protein was folded to its native state by stepwise urea removal via dialysis. The progress of this procedure is shown in the SDS-PAGE gel presented in Figure S1A, Supporting Information. Generally, a yield of 30−40 mg of pure protein per liter of growth medium was obtained. Once the protein was purified and refolded, it was placed in an incubator at 37 °C to allow fibril formation. Expression of the naturally occurring SBAFP followed a literature procedure, with minor changes detailed in the Methods section.[33]

The RGAFP-m1 gene in pET28a and WT RGAFP gene in pET24a were expressed in *E. coli* BL21 (DE3) cells. Whole cells were lysed by boiling for 10 min, releasing the heat-stable RGAFP-m1 and WT RGAFP into the soluble fraction. This was followed by a 2 h cooling period to room temperature to refold the protein and storage at 4 °C. The heat stable properties of WT RGAFP are retained in RGAFP-m1, as shown in Figure S1B, Supporting Information. A yield of ∼10 mg of pure RGAFP-m1 per liter of growth medium was obtained. WT RGAFP gave a yield of ∼50 mg of pure protein per liter of growth medium. As with SBAFP, both RGAFP-m1 and WT RGAFP were incubated at 37 °C to promote fibril growth.

Spectroscopic Characterization. *Mass Spectroscopy.* SBAFP-m1 has a calculated molecular mass of 19 397 Da with the N-terminal Met and 19 265 Da without it. Electrospray ionization mass spectrometry (ESI-MS) analysis (Figure S2A, Supporting Information) gives a molecular weight of 19 267 ± 4.8 Da, corresponding to the protein without the N-terminal Met, as desired. SBAFP-m1 runs on SDS-PAGE with an apparent molecular weight of 20 kDa.

The RGAFP-m1 protein has a calculated molecular mass of 11 410 Da with the N-terminal Met and 11 279 Da without it. ESI-MS analysis gives a molecular weight of 11 280 ± 2.8 Da (Figure S2B, Supporting Information), again corresponding to the mutant without the N-terminal Met. RGAFP-m1 runs on

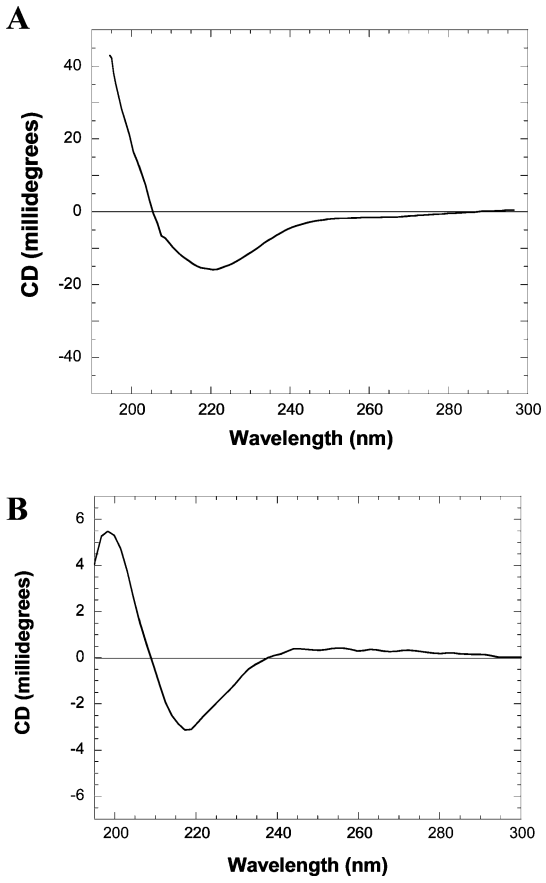


Figure 7. Circular dichroism spectra of (A) SBAFP-m1 in 10 mM sodium phosphate, pH 7.8, and (B) RGAFP-m1 in 10 mM sodium phosphate, pH 7.8. The spectrum for SBAFP-m1 was recorded in a 1 cm path length cell, while that for RGAFP-m1 was recorded in a 0.1 cm path length cell. Both spectra show a single minimum at ~220 nm indicative of predominantly β-sheet secondary structure for both proteins.

SDS-PAGE with an apparent molecular weight of 25 kDa.

This MS data provides important confirmation that the sequences of the proteins produced correspond exactly to those designed and tested for stability. No modifications to the proteins were detected.

Circular Dichroism. The CD spectrum for SBAFP-m1 is presented in Figure 7A. It shows a single peak at ~220 nm, suggesting a largely β-sheet structure. Deconvolution of the spectrum using the Contin[39] Set 4[40] program on the online DichroWeb[41] server gave the following estimates for secondary structure content: 4% α-helix, 64% β-sheet/turn and 32% random coil. Simulation gives 80% β-sheet structure, per Figure 5. This may indicate somewhat less stability for the experimental fibrils than expected by simulations. The SBAFP model has no α-helix, 89% β-sheet, 3% turns, and 8% coil as calculated by YASARA.

The CD spectrum for RGAFP-m1 is presented in Figure 7B. As with SBAFP-m1, it shows a single peak at ~220 nm, indicative of a largely β-sheet structure. Deconvolution of the spectrum using CONTIN-Set 4 gave a secondary structure content of 2% α-helix, 63% β-sheet/turn, and 35% random coil. Simulations give 72% β-sheet as shown in Figure 5. The RGAFP model has no α-helix, 88% β-sheet, and 12% coil as calculated by YASARA.

The secondary structure assignments agree reasonably well between structural models and the CD deconvolutions. The average error in secondary structure assignment by the CONTIN software is ~5% for α-structure and ~10% for β-structure,[42] although this depends on the protein and the lowest wavelength used in analysis.[43] Additional evidence of extended β-sheet formation is presented below.

Thioflavin-T Fluorescence. Thioflavin-T is a small molecule commonly used for the detection of amyloid cross-β-structure in peptides and proteins.[44,45] The binding mechanism is not well understood but is thought to involve either binding into "channels" between outward facing side chains of β-sheets and/or ThT micelle formation.[46] Empirically, ThT fluorescence is significantly altered when it binds to cross-β-structures: compared to ThT free in solution, the excitation maximum shifts from 385 to 450 nm in the presence of β-sheet fibrils and the emission maximum changes from 445 to 482 nm.[47] This fluorescence shift has been used extensively to probe peptides for β-sheet secondary structure, primarily with amyloid fibrils such as those from Aβ(1−42),[47] insulin,[48] and immunoglobulin light chain variable domain SMA.[49]

During the refolding of SBAFP-m1 by stepwise dialysis against solutions of decreasing urea concentrations, protein samples at each urea concentration were collected and analyzed for ThT fluorescence. Figure 8A shows the ThT fluorescence of SBAFP-m1 at urea concentrations from 8 M urea to 0 M urea at 4 °C and SBAFP-m1 in the absence of urea after incubation at 37 °C for 24 h. There is a gradual increase in fluorescence at 482 nm for SBAFP-m1 as the urea concentration decreases; given the long time of incubation at 4 °C (8 days total to change from 8 to 0 M urea), the increase in ThT signal may arise from partial polymerization.

SBAFP-m1 at 4 °C gives low ThT fluorescence, which is significantly above background and greater than that for WT SBAFP as shown in Figure 8B. The larger signal from SBAFP-m1 compared to WT SBAFP implies a small degree of polymerization during the extended period (8 days) required for stepwise dialysis at 4 °C. This is further evidence that SBAFP-m1 folds into the predicted β-helical structure after removal of urea by dialysis. Incubation of SBAFP-m1 at 37 °C for 24 h significantly increases ThT fluorescence at 482 nm, indicating formation of an extended β-sheet species that has higher ThT binding capacity or greater effects on the spectral changes of ThT. In general, incubation of amyloid-prone proteins and peptides at elevated temperatures is known to facilitate amyloid fibril

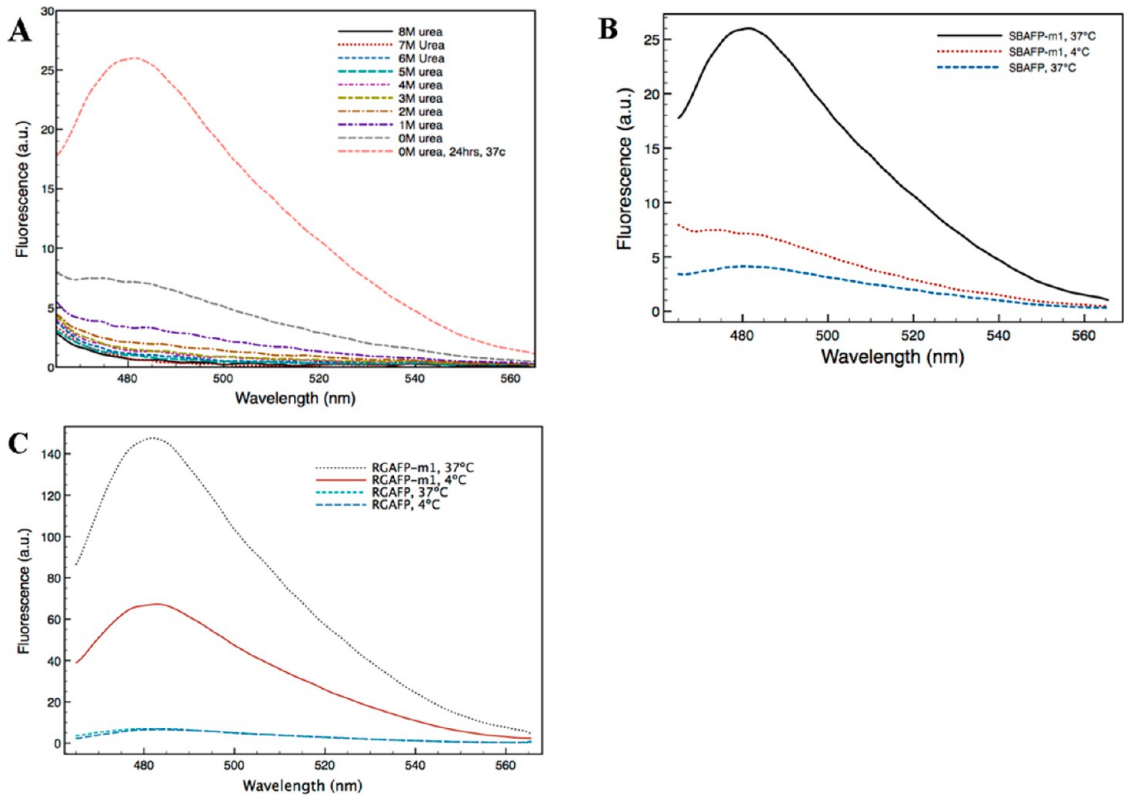


Figure 8. ThT fluorescence analysis of SBAFP-m1 and RGAFP-m1. (A) ThT fluorescence assay of SBAFP-m1 as a function of urea concentration in 0.1 M Tris-HCl, pH 7.8. The samples were dialyzed against decreasing concentrations of urea in buffer over the course of 8 days. Incubation at 37 °C leads to an increase in β-sheet structure. (B) ThT fluorescence of SBAFP-m1 at 4 and 37 °C compared to WT SBAFP at 37 °C. WT SBAFP has low fluorescence at 482 nm even after incubation at 37 °C. (C) ThT assay comparing WT RGAFP to RGAFP-m1. WT RGAFP has low fluorescence both at 4 and 37 °C, while RGAFP-m1 has significant fluorescence at 4 °C, which increases with incubation at 37 °C as expected for the growth of amyloid fibrils. For the assay, final protein concentrations, except RGAFP-m1, were 5 μM, and ThT was 10 μM in PBS, pH 7.4. RGAFP-m1 was added to a final concentration of 1.8 μM due to loss of protein on incubation at 37 °C. ThT data for RGAFP-m1 were normalized for the decrease in concentration.

formation.[50] The increase in ThT fluorescence with SBAFP-m1 after incubation at 37 °C implies that the elevated temperature promotes fibril formation since an identical sample held at 4 °C does not show the same large increase over 24 h (Figure 8B). The kinetic studies discussed below confirm this.

The RGAFP-m1 protein has an increased ThT fluorescence emission peak at 482 nm compared to WT RGAFP as shown in Figure 8C. Incubation of the WT RGAFP sample at 37 °C did not change the ThT emission intensity, suggesting that WT RGAFP does not form amyloid fibrils. RGAFP-m1 kept at 4 °C exhibits significant ThT fluorescence, albeit at a weaker intensity than the sample after incubation at 37 °C; this suggests some fibril formation at the lower temperature. This is also evidenced by the DLS data discussed below, which shows evidence for aggregates in the 4 °C sample. After incubation, fluorescence at 482 nm increases considerably, undergoing the same pattern seen with SBAFP-m1. This indicates the formation of longer amyloid fibrils with β-sheet conformation. Only weak fluorescence is seen for WT RGAFP both at 4 °C and after incubation at 37 °C, indicating the changes made to the protein template are causing the controlled aggregation. Fibril formation for both SBAFP-m1 and RGAFP-m1 is further evidenced by DLS data, discussed below.

Dynamic Light Scattering. DLS measures the hydrodynamic size of species present in solutions and is a nondestructive method for investigation of self-assembly. Measuring the size of the species in solution by DLS complements AFM imaging in characterizing the size distribution of fibrils.

For SBAFP-m1, DLS measurements performed before incubation show a species having an apparent hydrodynamic diameter of 6.6 $\pm$ 1.4 nm that constitutes ~99.8% of the sample, with the remainder consisting of minor species between 59 and 5560 nm in diameter. After incubation at 37 °C for 24 h, the species at 6.6 nm (presumably the monomer since the calculated diameter of gyration is ~4 nm for the unhydrated model) is absent. Larger species at 32 $\pm$ 8 and 230 $\pm$ 43 nm in apparent diameter are present. Thus, SBAFP-m1 may self-assemble to a minor extent at 4 °C, while the monomer is absent and larger species are present after 24 h incubation at 37 °C.

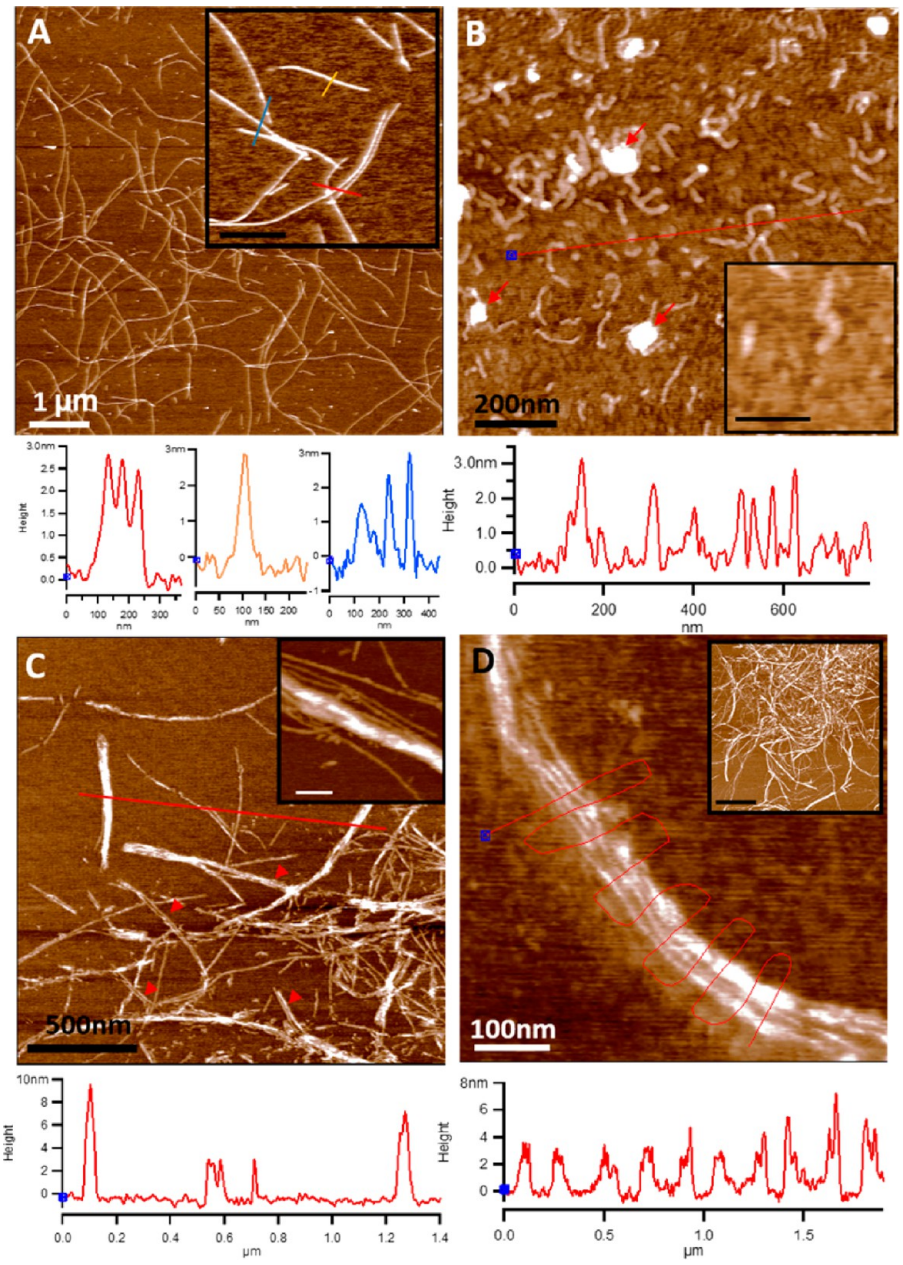


Figure 9. AFM topography images of SBAFP-m1 and WT SBAFP. (A) SBAFP-m1 fibrils after 48 h incubation. The inset shows fibrils at higher magnification (bar = 200 nm). Height profiles show that the height varies between 1.5 and 3.0 nm. (B) AFM image of WT SBAFP. Elongated structures and protein aggregates (indicated by red arrows) are present. The inset shows a higher magnification image (bar = 50 nm). Height profiles show that the heights of fibril-like structures are 2.0−3.0 nm. (C) Mature, three-week-old SBAFP-m1 fibrils. Red arrowheads indicate lateral assembly of SBAFP-m1 fibrils. Height profiles show variable (3, 8, and 10 nm) heights. The inset shows SBAFP-m1 fibrils in a parallel or antiparallel arrangement (bar = 100 nm). (D) Internal structure of mature SBAFP-m1 fibril. Image indicates that mature fibrils contain at least approximately four 3 nm tall individual fibrils bundled together. The red line corresponds to the height profile plotted under the image. The inset shows randomly attached fibrils and fibril bundles of SBAFP-m1 (bar = 2 μm).

Dynamic light scattering results with RGAFP-m1 also show an increase in fibrils after incubation at 37 °C. Before incubation, the sample consisted of species with apparent hydrodynamic diameters of 5.0 ± 0.3 nm (~98.7%), 28 ± 4 nm (~1.1%), and 143 ± 40 nm (~0.2%). After incubation at 37 °C, the monomer at 5 nm is absent, and a species with an apparent diameter of 268 ± 48 nm constitutes the entire sample. The presence of small amounts of large species in the sample before incubation at 37 °C could either be due to polymerization at 4 °C, as with SBAFP-m1, or the procedure by which RGAFP-m1 was purified and folded. Unlike SBAFP-m1, RGAFP-m1 was purified and folded by boiling followed by slow cooling to room temperature. This process gives the newly folded protein substantial time at both 4 °C and higher temperatures to undergo polymerization.

AFM Imaging. The formation of fibers by SBAFP-m1 was analyzed after 48 h and 3 weeks of incubation at 37 °C. After 48 h at 37 °C, the proteins assembled into fibers, as revealed by AFM imaging (Figure 9A). These fibers exhibit no branching or bundling, and their lengths measure ~1−7 μm. The expansion (Figure 9A inset) allows an accurate height determination of 1.5−3 nm, but predominantly between 2.5 and 3 nm. In contrast to SBAFP-m1, WT SBAFP does not yield well-defined fibrils under similar conditions. Instead, it forms arch-shaped assemblies (40−150 nm in length) and amorphous aggregates (see red arrows in Figure 9B). The height of the WT SBAFP monomers is 1 nm (see Figure 9B inset).

Extended (e.g., 3 week) incubation at 37 °C leads to bundles of fibrils in the case of SBAFP-m1, as shown in Figure 9C. The expanded view shows that the long axes of the fibrils are parallel to each other within each bundle (Figure 9C inset and Figure 9D) although the fibrils themselves, in the sense of N-to-C terminal, may be parallel or antiparallel. The number of fibrils in each bundle varies from ~2−4 as shown on Figure 9D. The overall height of the bundle varies from 6 to 10 nm, depending on the number of component fibers. The inset of Figure 9C shows the number of individual SBAFP-m1 fibrils and overall diameter of a typical bundle. After 3 weeks of incubation at 37 °C, SBAFP-m1 also forms random, haystack-like aggregates along with fibrils and bundles as shown in Figure 9D.

AFM images of RGAFP-m1 samples were taken after 3 days incubation at 37 °C. Figure 10A shows high coverage of RGAFP-m1 fibrils. These fibrils are unbranched and very uniform in size, with height of 1.5−2.0 nm and length of 150−400 nm. In contrast, no fibrillar structures were observed under the same condition using WT RGAFP, as shown in Figure 10B. The bright features in AFM topographs measure 1−10 nm tall with various lateral dimensions, which is consistent with monomeric and aggregates of RGAFP monomers. The time-dependence of fibril and bundle formation is qualitatively consistent with known amyloid fibril formation kinetics.[51]

Kinetic Analysis. The kinetics of fibril formation were followed for SBAFP-m1. DLS measurements on samples at 4 °C before the initiation of polymerization at 37 °C show the monomer as the only species present. Within 10 min of incubation at 37 °C the monomer is completely gone and replaced by larger species. The formation of large polymers was monitored by turbidity at 300 nm. A time course is presented in Figure 11. A lag phase is indicated by the data, and a simple exponential model gives a poor fit. The data were analyzed within the framework presented by Ferrone.[52] Considering the 12 rung cross-β-structure of SBAFP-m1, we assumed that the critical nucleus size is 1 so that the initial concentration of polymerization nuclei is equal to the initial concentration of protein. This leads to a simplified expression for the polymerization kinetics, given in eq 1

$$y = 1 - \mathrm{sech}(k_+ M_0 t) \quad (1)$$

where k_+ is the effective forward rate constant for polymerization, M_0 is the initial concentration of monomers, and t is time. This equation results from a new analytic solution to the Ferrone nucleated linear polymerization equations,[52] the full details of which will be presented elsewhere. By effective k_+ value here, we mean that we are not explicitly accounting for nucleation effects such as the need for a different, rare monomer conformation to nucleate fibrils as has been argued for PolyQ aggregation.[53]

Fitting the turbidity data to eq 1 (Figure 11) gives a value of 14 ± 1 M^{-1} s^{-1} for the rate constant, while fitting to ThT time course data (not shown) gives a value of 42 ± 7 M^{-1} s^{-1}. These effective values for k_+ are rapid compared to those measured for other cross-β-fibril formation reactions. In particular, effective k_+ values from the literature are 0.3 M^{-1} s^{-1} for insulin,[54] 3.3 M^{-1} s^{-1} for WT PrP^{Sc} (prion) elongation in mice (assuming a concentration of PrP^{C} in mice brain tissues),[55] 0.034, 0.3, and 1.4 M^{-1} s^{-1} for PolyQ polymers with 28, 36, and 47 glutamines,[56] respectively, and 0.14, 0.58, 1.5, and 6.9 M^{-1} s^{-1} for the spider silk amyloid peptide constructs eADF4(C2), eADF4(C4), eADF4(C8), and eADF4(C16), respectively.[57] While a critical nucleus for aggregation of 1 was determined for the PolyQ case,[56] it is unknown in the others and assumed to be 2 for insulin.[54] Nonetheless, it is apparent that despite the large size of the SBAFP monomers, the fact that they have preformed β-sheet structure affords much more rapid aggregation than other amyloids, particularly in the short time limit where aggregation goes generically as k_+.[2,52]

The fit assumes the turbidity is proportional to the total mass of polymer, which is a reasonable assumption provided that the fibril lengths exceed the wavelength of light (300 nm here).[58] This may break down in detail at short times in our experiments. We also note that our effective k_+ values fall many orders of magnitude below the Smoluchowski diffusion limit k_D of about 10^9 M^{-1} s^{-1},[59] a point noted before for polyQ aggregation.[53] In our case it is unlikely that this arises from a small probability nucleation complex as argued for PolyQ since the monomer form appears highly stable on its own in simulations and experimentally in solution. It is possible for anisotropic geometric constraints to significantly reduce k_+, a point we shall explore in future work.

In preliminary experiments, we find that polymerization under identical conditions, except when NaCl is added to 1 M final concentration, occurs approximately 3-fold slower. When fibrils are harvested from polymerization reactions by centrifugation and resuspended in fresh buffer at 60 °C, approximately 30% of the initial

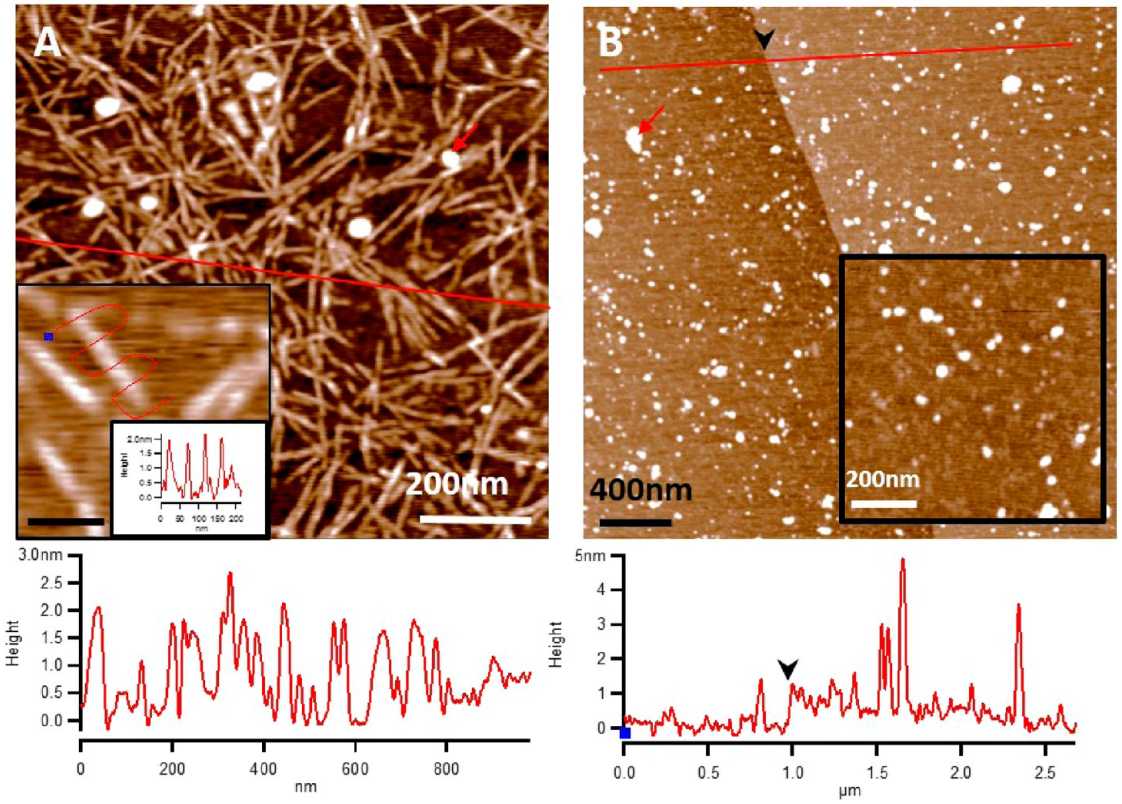


Figure 10. AFM topography images of RGAFP-m1 and RGAFP on poly-L-lysine coated mica (0001) surfaces. (A) A 1 μm × 1 μm AFM image of RGAFP-m1. The red line corresponds to the height profile under the image. The height of RGAFP fibrils is around 2.0 nm. Tall features (indicated by red arrows) are likely aggregates of RGAFP-m1 monomers. The inset shows the height profile for a single fibril (bar = 40 nm). (B) 3 μm × 3 μm AFM image of WT RGAFP. The red line corresponds to the height profile under the image. The inset shows a higher magnification, with individual bright features clearly shown. The red line indicated by the black arrow is a single layer step of mica (0001) with a known height of 1 nm.

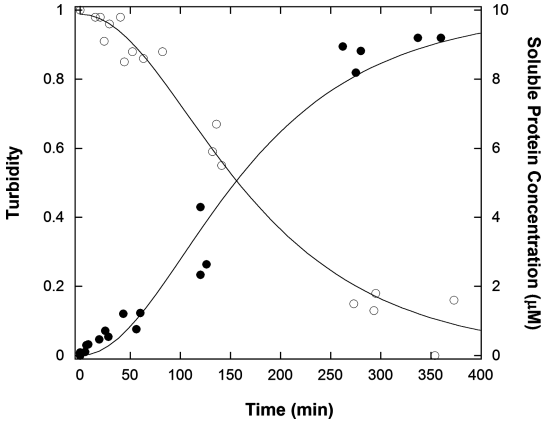


Figure 11. Kinetics of SBAFP-m1 fibril formation monitored by turbidity at 300 nm and total soluble protein in solution. A reaction mixture containing 10 μM SBAFP-m1 initially on ice was placed in a 37 °C incubator and shaken at 250 rpm to keep the sample mixed. At various times the reaction mixture was homogenized on a vortex mixer, a sample removed, and turbidity measured. The endpoint turbidity is very close to 1 in these experiments. The sample was then centrifuged to remove insoluble protein, and the soluble protein concentration measured. The data are a combination of three independent experiments and were fitted to eq 1. The forward rate constant for polymerization calculated from both types of data is 14 ± 1 M^{-1} s^{-1}.

turbidity is lost in 1.5 h, after which it is stable over time. Similar experiments show that approximately 5% of the initial turbidity is lost when fibrils are resuspended in buffer containing 80% ethanol in 1.5 h, after which it is stable over time.

CONCLUSIONS

In conclusion, we have shown that two monomeric wild type antifreeze BSPs (from Spruce Budworm and Ryegrass) can be engineered to polymerize into amyloid fibrils as evidenced by CD, DLS, ThT fluorescence, and AFM imaging, with the expected height profiles observed in AFM. To our knowledge, this is the first confirmation of amyloid formation from large BSP monomers, albeit in a synthetic context, despite the extant speculation that BSP structures should arise for known wild type amyloidogenic proteins[60−63] (we refer here to large BSPs with tightly packed interiors, although the Het-S fungal prion is known to form a more open β-solenoid amyloid structure[64]).

This demonstration is the first step toward applying these BSPs with extraordinary geometries for precision nanoscale applications such as templating nanoparticles for functional devices, enzyme arrays, peptide arrays for regenerative medicine, *etc*. Recently,

enzymatic[65] and charge transfer[66] activity for peptide-based amyloids has been demonstrated. The larger and more readily manipulated BSPs employed here hold greater promise for tailor-made nanoscale templates with precisely spatially defined functionalization on the 2−10 nm scale.

METHODS

Molecular Dynamics Simulations. All molecular dynamics simulations of both designed peptides were performed using the AMBER 12 package[36] at our custom built STRIDER GPU cluster at UC Davis. The ff12SB parameter set was employed with a time step of 0.002 ps and fully constrained bonds to hydrogen atoms. The aqueous peptide environment was simulated explicitly with TIP3P water at constant pressure, in a long, rectangular box with periodic boundary conditions (PBC). In simulating long, fibrillar multimers, a novel adaptive box algorithm was employed to economize the computation (avoiding an exceedingly large solvation geometry) all while maintaining a consistent solvent environment: (i) The minimum pairwise distance of the macromolecule (solute) with all of its periodic images was regularly recomputed, specifically anticipating rotational drift. (ii) When this distance decreased below a cutoff of 15 Å, waters beyond this distance from any solute atom were stripped; the solute and close waters were reoriented in a new rectangular box wherein the box boundaries were at least 20 Å from any solute atom; and this box was resolvated with TIP3P waters of an appropriate density for the fixed pressure. (iii) The simulation was recommenced, accepting a picosecond-scale duration to accommodate the re-equilibration of outer-shell waters in the new periodic box. The simulations for the SBAFP-m1 and RGAFPm2 fibrils were carried out for 20 ns with five different random number seeds for the Langevin thermostat.

β-Sheet content was measured from simulation time series by using VMD[37] to count the number of residues within the typical β-sheet secondary structure region of the Ramachandran plot of $\varphi-\psi$ torsion angles (i.e., $-180° < \varphi < 0°$, $-180° < \psi < -150°$ and $-180° < \varphi < 0°$, $60° < \psi < 180°$).

The height profiles of the monomers for comparison with AFM experiments were obtained as follows. First, to remove any inherent twist in the monomers, we constrain Cα atoms on one side of the monomer to lie in a plane and use energy minimization within the AMBER suite to relax this constrained structure. Then, each monomer was rotated such that its helical axis is aligned along the *x*-axis. The monomers were then rotated about the *x*-axis at 10° angle intervals. At each rotation the maximum *z*-coordinate height difference was measured for 5 Å thick slabs along the length of the helical axis. The average and standard deviations of this height along the length of the monomer was then obtained. Only heavy atoms were used for the height measurements.

Protein Expression, Purification, and Folding. The SBAFP-m1 and RGAFP-m1 genes in pET28a were procured from Life Technologies (Grand Island, NY). Proteins were expressed in *E. coli* BL21 (DE3) cells. For SBAFP-m1, 1 L cultures were inoculated with overnight cultures and grown at 37 °C, until OD_{600} reached 0.9−1.0. Cultures were cooled on ice for 30 min after which isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Protein expression proceeded at 30 °C for 3 h, and cells were collected via centrifugation and resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM EDTA, and 0.5% Triton X-100). Cells were lysed by sonication and soluble and insoluble fractions separated by centrifugation. Insoluble inclusion bodies were purified by repeated sonication in lysis buffer without Triton X-100 and centrifugation, a total of four times. The purified inclusion bodies were resuspended in folding buffer (100 mM Tris and 50 mM glycine, pH 8.0) and were added dropwise into denaturing buffer (100 mM Tris-HCl, 50 mM glycine, and 8.5 M urea, pH 8.0) at 4 °C with stirring overnight. Denatured SBAFP-m1 was purified on Fast Q Sepharose anion exchange column (GE Healthcare Lifesciences, U.K.). The loading buffer was 50 mM Tris-HCl, 10 mM NaCl, and 8 M urea, pH 8.0, while the elution buffer was identical except for being supplemented with 500 mM NaCl. Elution used a linear gradient. Purified SBAFP-m1 was concentrated using Amicon centrifugal devices (EMD Millipore, Germany) with a molecular weight cutoff of 3500 Da. Purified, concentrated SBAFP-m1 was refolded by stepwise dialysis out of urea using dialysis membrane with a molecular weight cutoff of 3000 Da and into 0.1 M Tris-HCl, pH 8.0 at 4 °C. Each day the concentration of urea was decreased by 1 M until it reached zero.

The WT SBAFP gene in pET20b was acquired from the Davies lab and expressed according to the published protocol with minor changes.[33] Briefly, protein was expressed in BL21(DE3) *E. coli* cells in LB with IPTG induction when cell density reached 0.9 at 600 nm. Cells were collected via centrifugation and resuspended in a lysis buffer of 10 mM Tris-HCl, pH 9.0, 1 mM EDTA, and 10 mM 2-mercaptoethanol. Cells were disrupted by sonication and IBs denatured in lysis buffer containing 8 M urea and allowed to stir overnight at 4 °C. Crude, denatured SBAFP was purified like SBAFP-m1 with buffers supplemented with 10 mM 2-mercaptoethanol to keep cysteines reduced. Dialysis buffer for refolding was supplemented with 2% w/v glycerol.

The RGAFP-m1 and WT RGAFP proteins were expressed by inoculating a 1L culture with overnight cultures and grown at 37 °C until OD_{600} reached 0.9−1.0. Cultures were chilled in an ice−water bath for 20 min followed by IPTG addition to 0.5 mM. Protein expression continued at 18 °C for 20 h. The cells were pelleted by centrifugation and resuspended in PBS, pH 7.4 (10 mM sodium phosphate, 138 mM NaCl, and 2.7 mM KCl) followed by freezing at −80 °C. The frozen cell pellet was thawed at 37 °C and boiled for 10 min to lyse the cells. After a 2 h cooling period, RGAFP-m1 was sonicated extensively. DNaseI (Worthington Biochemical Corp., Lakewood, NJ) was added to a final concentration of 1.6 μg/mL to the sonicated sample and incubated for 40 min followed by heat inactivation of the DNaseI by boiling at 100 °C for 10 min. The sample was allowed to cool to 4 °C and dialyzed into PBS with 6−8 kDa dialysis tubing (Spectrum Laboratories, Irving, Texas). As a final purification step, the dialyzed protein was loaded onto a Fast Q Sepharose anion exchange column (GE Healthcare Lifesciences, U.K.) and washed with PBS pH 7.4. The elution buffer was supplemented with 0.5 M NaCl. Fractions containing proteins were pooled and dialyzed into PBS buffer before incubating samples at 37 °C for fibril formation.

The boiled cell lysate of WT RGAFP was purified using a nickel−NTA resin pre-equilibrated with binding buffer (50 mM Tris-HCl, 0.5 M NaCl and , 5 mM imidazole, pH 7.5) in a 10 cm × 1.5 cm column. The bound WT RGAFP was washed with 10 column volumes of binding buffer and the protein eluted with elute buffer (50 mM Tris-HCl, 0.5 M NaCl, and 200 mM imidazole, pH 7.5). Fractions containing protein were pooled and dialyzed into PBS using a 6−8 kDa dialysis membrane tubing. Protein concentrations for both RGAFP-m1 and WT RGAFP were determined with the bicinchonimic acid assay (Thermo Scientific, Rockford, IL).

Amyloid Fibril Formation. Purified SBAFP-m1 at a concentration of 70 μM in 0.1 M Tris-HCl, pH 8.0, was transferred to an Eppendorf tube and incubated at 37 °C to promote fibril formation. Purified RGAFP-m1at at a concentration of 98 μM in PBS pH 7.4 was incubated at 37 °C in an Eppendorf tube until further analysis.

Thioflavin-T Fluorescence Assay. ThT fluorescence was measured as described.[45,67] ThT stock solutions were prepared by dissolving ∼2 mg of ThT (Sigma-Aldrich) in 2 mL of PBS, pH 7.4, and filtered through a 0.22 μm filter. Stock concentrations were determined using an extinction coefficient of 26 620 M^{-1} cm^{-1} at 416 nm. A working 500 μM ThT solution was prepared from the stock solution. For the assay, SBAFP-m1 was added to a final concentration of 5 μM and ThT to a final concentration of 10 μM

in PBS, pH 7.4. The emission spectrum was recorded from 465 to 565 nm with excitation at 450 nm. For the assay, all protein concentrations, except for RGAFP-m1, were added to a final concentration of 5 μM and ThT to a final concentration of 10 μM in PBS, pH 7.4. RGAFP-m1 was added to a final concentration of 1.8 μM due to a loss of protein after incubation. ThT data for RGAFP-m1 were normalized for the decrease in concentration.

Circular Dichroism. Protein secondary structure was analyzed using CD. For SBAFP-m1, the spectrum is a concentration-normalized combination of those for 0.02 mg/mL sample from 190 to 200 nm and 0.2 mg/mL sample from 200 to 300 nm in 10 mM sodium phosphate, pH 7.4, in a 1 cm cell at 25 °C. Spectra were collected at a scan rate proportional to high voltage using an OLIS DSM 20 instrument. Reported spectra are an average of 5 scans. For RGAFP-m1, spectra were taken with 0.2 mg/mL sample in a 1 mm path length cell at 25 °C in 10 mM sodium phosphate, pH 7.4.

Dynamic Light Scattering. DLS measurements were performed using a Zetasizer NanoS (Malvern Instruments, Worcestershire, U.K.). Sample preparation for DLS measurements consisted of clarification by centrifugation at 13 000 × *g* for 5 min. Protein concentrations were ~1 mg/mL in PBS, pH 7.4. Measurements were made at either 4 or 37 °C depending on prior sample treatment. A protein refractive index of 1.450 and a water refractive index of 1.330 were used. Each reported value is an average of 10 acquisitions, each lasting 300 s. The averages and standard deviations of these 10 runs are reported.

Atomic Force Microscopy. *Sample Preparation.* Pieces of 8.0 mm × 8.0 mm × 0.5 mm muscovite mica were cut and mounted on standard microscope slides using composite epoxy glue (5 min Epoxy, ITW Performance Polymers and Fluids, FL, USA). Before protein deposition, the top layer(s) of mica was peeled mechanically to expose fresh (0001) surfaces. For SBAFP-m1 or WT SBAFP, 20 μL of sample in Tris buffer (100 mM, pH 8.0) was deposited on the freshly exposed mica (0001) surfaces. After 5 min incubation, the surfaces were washed three times with 200 μL of Tris buffer (100 mM, pH 8.0) to remove weakly bound proteins and fibrils. The samples were immediately imaged in the Tris buffer. Sample preparation for RGAFP-m1 or WT RGAFP followed similar protocols except for the surface coating and imaging medium. The freshly cleaved mica (0001) surfaces were coated with poly-L-lysine, by dropping 80 μL of 0.1% (w/v) poly-L-lysine (Sigma P8920, MW 150−300 kDa) onto the surface, incubating for 5 min, then washing with Milli-Q water. The surfaces were dried with clean air before protein deposition. The samples were imaged under ambient conditions.

Imaging. AFM was performed using an MFP-3D AFM (Asylum Research, Santa Barbara, California, USA). Most images were acquired using AC or tapping mode to minimize perturbation and damage to surface bound proteins and fibrils. The typical set point was at 70%−80% of original amplitude, and scan rate was 0.8−1 Hz. For imaging in aqueous media, two types of probes were used. The first was a Biolever A cantilever (BL-RC150, Olympus, Japan). Its resonance frequency (f) was determined by the built-in software of MFP3D AFM. The spring constant (k) was provided by the manufacturer. Typically, $f \approx$ 10 kHz and $k \approx$ 30 pN/nm. The second type was an MSNL E cantilever (Bruker, USA) with $f \approx$ 11 kHz and $k \approx$ 100 pN/nm. For imaging under ambient (dry) conditions, AC240 cantilevers (Olympus, Japan) with $f \approx$ 65−75 kHz and $k \approx$ 2.0 N/m were used.

Fibrilization Kinetics. Monomeric samples (10 μM) in 0.1 M Tris-HCl pH 8.0 were held at 4 °C in 15 mL capped plastic centrifuge tubes until polymerization was initiated by incubating at 37 °C with shaking at 250 rpm. At various times, the tube was homogenized on a vortex mixer, and a sample was removed for analysis by absorbance at 300 nm (turbidity), DLS, and ThT fluorescence. Before ThT analysis, the sample was centrifuged for 5 min at 12 000 × *g* at room temperature. The soluble supernatant and the resuspended insoluble precipitate were separately analyzed with ThT as above. The soluble protein concentration in the supernatant was measure by absorbance at 280 nm before being used in the ThT assay.

Conflict of Interest: The authors declare the following competing financial interest(s): We have submitted a provisional patent application drawing on this work, but otherwise have no potential financial interest in the work.

Acknowledgment. This work was supported by the RISE program from the UC Davis Office of Research (to M.D.R.P., A.K., A.N., C.S., N.R.H., N.H., A.J.K., K.M.R., M.D.T., G.Y.L., X.C., R.R.P.S., and D.L.C.). The GPU computing cluster used here was sponsored by the US National Science Foundation grants DMR-1207624 and DMR-0844115 (to A.J.K., N.M., N.R.H., K.M.R., D.L.C., and R.R.P.S.). Some AFM images were acquired using the instrument at Keck Spectral Imaging Facility initiated by NSF-MRI and subsequently upgraded under the support of W. M. Keck Foundation. Protein expression, purification, folding, and spectral analysis of SBAFP-m1 was done by MDRP and CS. SBAFP-m1 was designed by MDT. The RGAFP-m1 work was done by A.N. and K.F., with protein design by M.D.R.P. and M.D.T. Computer simulations were designed by N.R.H., and performed by N.R.H., N.M., and A.J.K. MD based height analysis was performed by K.M.R. and N.M. MD based β-sheet content analysis was performed by K.M.R. AFM sample preparations, imaging, and data analysis were done by A.K. and G.Y.L. D.L.C. is the lead principal investigator on the project, supervising the theoretical design and simulation work with R.R.P.S., motivating the selection of proteins, and writing of the paper along with M.D.T.

Supporting Information Available: Additional details about SDS-PAGE results, mass spectroscopy results, kinetics results (via dynamical light scattering), and the RGAFP model structure used in the theoretical height analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

1. Olmsted, J. B.; Borisy, G. G. Microtubules. *Annu. Rev. Biochem.* **1973**, *42*, 507–540.
2. Zhang, S. G. Fabrication of Novel Biomaterials Through Molecular Self-assembly. *Nat. Biotechnol.* **2003**, *21*, 1171–1178.
3. Sara, M.; Sleytr, U. B. Crystalline Bacterial Cell Surface Layers (S-layers): From Cell Structure to Biomimetics. *Prog. Biophys. Mol. Biol.* **1996**, *65*, 83–111.
4. Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75*, 333–366.
5. Seeman, N. C. Nanomaterials Based on DNA. In *Annual Review of Biochemistry*; Kornberg, R. D., Raetz, C. R. H., Rothman, J. E., Thorner, J. W., Eds.; Annual Reviews: Palo Alto, CA, 2010; Vol. 79, pp 65−87.
6. Le, J. D.; Pinto, Y.; Seeman, N. C.; Musier-Forsyth, K.; Taton, T. A.; Kiehl, R. A. DNA-templated Self-assembly of Metallic Nanocomponent Arrays on a Surface. *Nano Lett.* **2004**, *4*, 2343–2347.
7. Sacca, B.; Meyer, R.; Erkelenz, M.; Kiko, K.; Arndt, A.; Schroeder, H.; Rabe, K. S.; Niemeyer, C. M. Orthogonal Protein Decoration of DNA Origami. *Angew. Chem., Int. Ed.* **2010**, *49*, 9378–9383.
8. Liu, J. F.; Uprety, B.; Gyawali, S.; Woolley, A. T.; Myung, N. V.; Harb, J. N. Fabrication of DNA-Templated Te and Bi2Te3 Nanowires by Galvanic Displacement. *Langmuir* **2013**, *29*, 11176–11184.
9. Castro, C. E.; Kilchherr, F.; Kim, D.-N.; Shiao, E. L.; Wauer, T.; Wortmann, P.; Bathe, M.; Dietz, H. A Primer to Scaffolded DNA Origami. *Nat. Methods* **2011**, *8*, 221–229.
10. Mao, C. B.; Solis, D. J.; Reiss, B. D.; Kottmann, S. T.; Sweeney, R. Y.; Hayhurst, A.; Georgiou, G.; Iverson, B.; Belcher, A. M. Virus-based Toolkit for the Directed Synthesis of Magnetic and Semiconducting Nanowires. *Science* **2004**, *303*, 213–217.
11. Lee, Y. J.; Belcher, A. M. Nanostructure Design of Amorphous FePO4 Facilitated by a Virus for 3 V Lithium Ion Battery Cathodes. *J. Mater. Chem.* **2011**, *21*, 1033–1039.
12. Barnhart, M. M.; Chapman, M. R. Curli Biogenesis and Function. *Annu. Rev. Microbiol.* **2006**, *60*, 131–147.
13. Si, K.; Choi, Y. B.; White-Grindley, E.; Majumdar, A.; Kandel, E. R. Aplysia CPEB Can Form Prion-like Multimers in

Sensory Neurons that Contribute to Long-Term Facilitation. *Cell* **2010**, *140*, 421–U179.
14. Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P. R.; Simon, R.; Schubert, D.; al, e. Functional Amyloids As Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science* **2009**, *325*, 328–332.
15. Knowles, T. P. J.; Smith, J. F.; Craig, A.; Dobson, C. M.; Welland, M. E. Spatial Persistence of Angular Correlations in Amyloid Fibrils. *Phys. Rev. Lett.* **2006**, 96.
16. Knowles, T. P. J.; Buehler, M. J. Nanomechanics of Functional and Pathological Amyloid Materials. *Nat. Nanotechnol.* **2011**, *6*, 469–479.
17. Slotta, U.; Hess, S.; Spiess, K.; Stromer, T.; Serpell, L.; Scheibel, T. Spider Silk and Amyloid Fibrils: A Structural Comparison. *Macromol. Biosci.* **2007**, *7*, 183–188.
18. Sullan, R. M. A.; Gunari, N.; Tanur, A. E.; Yuri, C.; Dickinson, G. H.; Orihuela, B.; Rittschof, D.; Walker, G. C. Nanoscale Structures and Mechanics of Barnacle Cement. *Biofouling* **2009**, *25*, 263–275.
19. Scheibel, T.; Parthasarathy, R.; Sawicki, G.; Lin, X. M.; Jaeger, H.; Lindquist, S. L. Conducting Nanowires Built by Controlled Self-assembly of Amyloid Fibers and Selective Metal Deposition. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4527–4532.
20. Reches, M.; Gazit, E. Casting Metal Nanowires Within Discrete Self-assembled Peptide Nanotubes. *Science* **2003**, *300*, 625–627.
21. Sakai, H.; Watanabe, K.; Asanomi, Y.; Kobayashi, Y.; Chuman, Y.; Shi, L. H.; Masuda, T.; Wyttenbach, T.; Bowers, M. T.; Uosaki, K.; Sakaguchi, K. Formation of Functionalized Nanowires by Control of Self-Assembly Using Multiple Modified Amyloid Peptides. *Adv. Funct. Mater.* **2013**, *23*, 4881–4887.
22. Knowles, T. P. J.; Oppenheim, T. W.; Buell, A. K.; Chirgadze, D. Y.; Welland, M. E. Nanostructured Films from Hierarchical Self-assembly of Amyloidogenic Proteins. *Nat. Nanotechnol.* **2010**, *5*, 204–207.
23. Arora, A.; Ha, C.; Park, C. B. Insulin Amyloid Fibrillation at Above 100 degrees C: New Insights into Protein Folding Under Extreme Temperatures. *Protein Sci.* **2004**, *13*, 2429–2436.
24. Baxa, U.; Ross, P. D.; Wickner, R. B.; Steven, A. C. The N-terminal Prion Domain of Ure2p Converts from an Unfolded to a Thermally Resistant Conformation Upon Filament Formation. *J. Mol. Biol.* **2004**, *339*, 259–264.
25. Kardos, J.; Micsonai, A.; Pal-Gabor, H.; Petrik, E.; Graf, L.; Kovacs, J.; Lee, Y. H.; Naiki, H.; Goto, Y. Reversible Heat-Induced Dissociation of beta(2)-Microglobulin Amyloid Fibrils. *Biochemistry* **2011**, *50*, 3211–3220.
26. Hammer, N. D.; Wang, X.; McGuffie, B. A.; Chapman, M. R. Amyloids: Friend or foe? *J. Alzheimer's Dis.* **2008**, *13*, 407–419.
27. Ryu, J.; Park, C. B. High Stability of Self-Assembled Peptide Nanowires Against Thermal, Chemical, and Proteolytic Attacks. *Biotechnol. Bioeng.* **2010**, *105*, 221–230.
28. Sunde, M.; Blake, C. The Structure of Amyloid Fibrils by Electron Microscopy and X-ray Diffraction. *Adv. Protein Chem.* **1997**, *50*, 123–159.
29. Wille, H.; Bian, W.; McDonald, M.; Kendall, A.; Colby, D. W.; Bloch, L.; Ollesch, J.; Borovinskiy, A. L.; Cohen, F. E.; Prusiner, S. B.; Stubbs, G. Natural and Synthetic Prion Structure from X-ray Fiber Diffraction. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16990–16995.
30. Buell, A. K.; Dhulesia, A.; Mossuto, M. F.; Cremades, N.; Kumita, J. R.; Dumoulin, M.; Welland, M. E.; Knowles, T. P. J.; Salvatella, X.; Dobson, C. M. Population of Nonnative States of Lysozyme Variants Drives Amyloid Fibril Formation. *J. Am. Chem. Soc.* **2011**, *133*, 7737–7743.
31. Graether, S. P.; Kuiper, M. J.; Gagne, S. M.; Walker, V. K.; Jia, Z. C.; Sykes, B. D.; Davies, P. L. beta-helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect. *Nature* **2000**, *406*, 325–328.
32. Benbassat, A.; Bauer, K.; Chang, S. Y.; Myambo, K.; Boosman, A.; Chang, S. Processing of the Initiation Methionine from Proteins - Properties of the Escherichia-Coli Methionine Aminopeptidase and its Gene Structure. *J. Bacteriol.* **1987**, *169*, 751–757.
33. Gauthier, S. Y.; Kay, C. M.; Sykes, B. D.; Walker, V. K.; Davies, P. L. Disulfide Bond Mapping and Structural Characterization of Spruce Budworm Antifreeze Protein. *Eur. J. Biochem.* **1998**, *258*, 445–453.
34. Middleton, A. J.; Marshall, C. B.; Faucher, F.; Bar-Dolev, M.; Braslavsky, I.; Campbell, R. L.; Walker, V. K.; Davies, P. L. Antifreeze Protein from Freeze-Tolerant Grass Has a Beta-Roll Fold with an Irregularly Structured Ice-Binding Site. *J. Mol. Biol.* **2012**, *416*, 713–724.
35. Cooper, S.; Khatib, F.; Treuille, A.; Barbero, J.; Lee, J.; Beenen, M.; Leaver-Fay, A.; Baker, D.; Popovic, Z.; Players, F. Predicting Protein Structures with a Multiplayer Online Game. *Nature* **2010**, *466*, 756–760.
36. Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Gotz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*; University of California: San Francisco, CA, 2012.
37. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics Modell.* **1996**, *14*, 33–38.
38. Steinle, A.; Li, P.; Morris, D. L.; Groh, V.; Lanier, L.; Strong, R. K.; Spies, T. Interactions of human NKG2D with its ligands MICA, MICB, and homologs of the mouse RAE-1 protein family. *Immunogenetics* **2001**, *52*, 279–287.
39. Provencher, S. W.; Glockner, J. Estimation of Globular Protein Secondary Structure from Circular-dichroism. *Biochemistry* **1981**, *20*, 33–37.
40. Janes, R. W.; Wallace, B. A. *Modern Techniques in Circular Dichroism and Synchrotron Radiation Circular Dichroism Spectroscopy*; IOS Press: Amsterdam, The Netherlands, 2009.
41. Whitmore, L.; Wallace, B. A. Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and References Databases. *Biopolymers* **2007**, *89*, 392–400.
42. Sreerama, N.; Woody, R. W. Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Anal. Biochem.* **2000**, *287*, 252–260.
43. Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2006**, *1*, 2876–2890.
44. Vassar, P. S.; Culling, C. F. A. Fluorescent Stains, with Special Reference to Amyloid and Connective Tissues. *Arch. Pathol.* **1959**, *68*, 487–498.
45. Levine, H. Thioflavine-T Interaction with Synthetic Alzheimers-disease Beta-amyloid Peptides - Detection of Amyloid Aggregation in Solution. *Protein Sci.* **1993**, *2*, 404–410.
46. Biancalana, M.; Koide, S. Molecular Mechanism of Thioflavin-T Binding to Amyloid Fibrils. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 1405–1412.
47. LeVine, H. Quantification of Beta-sheet Amyloid Fibril Structures with Thioflavin-T. *Amyloid, Prions, Other Protein Aggregates* **1999**, *309*, 274–284.
48. Nielsen, L.; Khurana, R.; Coats, A.; Frokjaer, S.; Brange, J.; Vyas, S.; Uversky, V. N.; Fink, A. L. Effect of Environmental Factors on the Kinetics of Insulin Fibril Formation: Elucidation of the Molecular Mechanism. *Biochemistry* **2001**, *40*, 6036–6046.
49. Ionescu-Zanetti, C.; Khurana, R.; Gillespie, J. R.; Petrick, J. S.; Trabachino, L. C.; Minert, L. J.; Carter, S. A.; Fink, A. L. Monitoring the Assembly of Ig Light-chain Amyloid Fibrils by Atomic Force Microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13175–13179.
50. Kusumoto, Y.; Lomakin, A.; Teplow, D. B.; Benedek, G. B. Temperature Dependence of Amyloid Beta-protein Fibrillization. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12277–12282.

51. Harrison, R. S.; Sharpe, P. C.; Singh, Y.; Fairlie, D. P. Amyloid Peptides and Proteins in Review. *Rev. Physiol., Biochem. Pharmacol.* **2007**, *159*, 1–77.
52. Ferrone, F. Analysis of Protein Aggregation Kinetics. *Methods Enzymol.* **1999**, *309*, 256–274.
53. Bhattacharyya, A. M.; Thakur, A. K.; Wetzel, R. Polyglutamine aggregation nucleation: Thermodynamics of a highly unfavorable protein folding reaction. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15400–15405.
54. Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. An Analytical Solution to the Kinetics of Breakable Filament Assembly. *Science* **2009**, *326*, 1533–1537.
55. Cox, D. L.; Singh, R. R. P.; Yang, S. C. Prion Disease: Exponential Growth Requires Membrane Binding. *Biophys. J.* **2006**, *90*, L77–L79.
56. Chen, S. M.; Ferrone, F. A.; Wetzel, R. Huntington's Disease Age-of-onset Linked to Polyglutamine Aggregation Nucleation. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11884–11889.
57. Humenik, M.; Magdeburg, M.; Scheibel, T. Influence of Repeat Numbers on Self-assembly Rates of Repetitive Recombinant Spider Silk Proteins. *J. Struct. Biol.* **2014**, *186*, 431–437.
58. Berne, B. J. Interpretation of the Light Scattering from Long Rods. *J. Mol. Biol.* **1974**, *89*, 755–758.
59. Northrup, S. H.; Erickson, H. P. Kinetics of Protein-Protein Association Explained by Brownian Dynamics Computer-simulation. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3338–3342.
60. Perutz, M. F.; Finch, J. T.; Berriman, J.; Lesk, A. Amyloid Fibers are Water-filled Nanotubes. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5591–5595.
61. Govaerts, C.; Wille, H.; Prusiner, S. B.; Cohen, F. E. Evidence for Assembly of Prions with Left-handed Beta 3-helices into Trimers. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8342–8347.
62. Stork, M.; Giese, A.; Kretzschmar, H. A.; Tavan, P. Molecular Dynamics Simulations Indicate a Possible Role of Parallel Beta-helices in Seeded Aggregation of Poly-Gln. *Biophys. J.* **2005**, *88*, 2442–2451.
63. Kunes, K. C.; Clark, S. C.; Cox, D. L.; Singh, R. R. P. Left Handed Beta-helix Models for Mammalian Prion Fibrils. *Prion* **2008**, *2*, 81–90.
64. Wasmer, C.; Lange, A.; Van Melckebeke, H.; Siemer, A. B.; Riek, R.; Meier, B. H. Amyloid Fibrils of the HET-s(218−289) Prion Form a Beta Solenoid with a Triangular Hydrophobic Core. *Science* **2008**, *319*, 1523–1526.
65. Rufo, C. M.; Moroz, Y. S.; Moroz, O. V.; Stohr, J.; Smith, T. A.; Hu, X. Z.; DeGrado, W. F.; Korendovych, I. V. Short Peptides Self-assemble to Produce Catalytic Amyloids. *Nat. Chem.* **2014**, *6*, 303–309.
66. Ivnitski, D.; Amit, M.; Rubinov, B.; Cohen-Luria, R.; Ashkenasy, N.; Ashkenasy, G. Introducing Charge Transfer Functionality into Prebiotically Relevant Beta-sheet Peptide Fibrils. *Chem. Commun.* **2014**, *50*, 6733–6736.
67. Khurana, R.; Coleman, C.; Ionescu-Zanetti, C.; Carter, S. A.; Krishna, V.; Grover, R. K.; Roy, R.; Singh, S. Mechanism of Thioflavin-T Binding to Amyloid Fibrils. *J. Struct. Biol.* **2005**, *151*, 229–238.
68. Buchko, G. W.; Ni, S. S.; Robinson, H.; Welsh, E. A.; Pakrasi, H. B.; Kennedy, M. A. Characterization of Two Potentially Universal Turn Motifs that Shape the Repeated Five-residues Fold - Crystal Structure of a Lumenal Pentapeptide Repeat Protein from Cyanothece 51142. *Protein Sci.* **2006**, *15*, 2579–2595.
69. Ni, S. S.; Sheldrick, G. M.; Benning, M. M.; Kennedy, M. A. The 2 Angstrom Resolution Crystal Structure of HetL, a Pentapeptide Repeat Protein Involved in Regulation of Heterocyst Differentiation in the Cyanobacterium Nostoc Sp Strain PCC 7120. *J. Struct. Biol.* **2009**, *165*, 47–52.
70. Pentelute, B. L.; Gates, Z. P.; Tereshko, V.; Dashnau, J. L.; Vanderkooi, J. M.; Kossiakoff, A. A.; Kent, S. B. H. X-ray Structure of Snow Flea Antifreeze Protein Determined by Racemic Crystallization of Synthetic Protein Enantiomers. *J. Am. Chem. Soc.* **2008**, *130*, 9695–9701.
71. Lauersen, K. J.; Brown, A.; Middleton, A.; Davies, P. L.; Walker, V. K. Expression and Characterization of an Antifreeze Protein from the Perennial Rye Grass, Lolium Perenne. *Cryobiology* **2011**, *62*, 194–201.
72. Benach, J.; Chen, Y.; Vorobiev, S. M.; Seetharaman, J.; Ho, C. K.; Janjua, H.; Conover, K.; Ma, L.-C.; Xiao, R.; et al. Crystal structure of ydcK from *Salmonella cholerae* at 2.38 Å resolution. Northeast Structural Genomics target SCR6. http://www.pdb.org/pdb/explore/explore.do?structureId=2PIG.
73. Leinala, E. K.; Davies, P. L.; Doucet, D.; Tyshenko, M. G.; Walker, V. K.; Jia, Z. C. A Beta-helical Antifreeze Protein Isoform with Increased Activity - Structural and Functional Insights. *J. Biol. Chem.* **2002**, *277*, 33349–33352.