Engineering nanoscale order into a designed protein fiber

David Papapostolou†, Andrew M. Smith‡§, Edward D. T. Atkins⁶, Seb J. Oliver⁵, Maxim G. Ryadnov⁴, Louise C. Serpell⁵, and Derek N. Woolfson¹‡‡

¹School of Chemistry, University of Bristol, Bristol BS 8 1TS, United Kingdom; ²Department of Biochemistry, School of Life Sciences, University of Sussex, Falmer BN 1 9QG, United Kingdom; ³Department of Physics, University of Bristol, Bristol BS 8 1LT, United Kingdom; ⁴Department of Biochemistry, University of Bristol, Bristol BS 8 1TD, United Kingdom; and ⁵Department of Physics, University of Sussex, Falmer BN 1 9QH, United Kingdom

Edited by Alan R. Fersht, University of Cambridge, Cambridge, United Kingdom, and approved May 21, 2007 (received for review January 30, 2007)

We have established a designed system comprising two peptides that coassemble to form long, thickened protein fibers in water. This system can be rationally engineered to alter fiber assembly, stability, and morphology. Here, we show that rational mutations to our original peptide designs lead to structures with a remarkable level of order on the nanoscale that mimics certain natural fibrous assemblies. In the engineered system, the peptides assemble into two-stranded α-helical coiled-coil rods, which pack in axial register in a 3D hexagonal lattice of size 1.824 nm, and with a periodicity of 4.2 nm along the fiber axis. This model is supported by both electron microscopy and x-ray diffraction. Specifically, the fibers display surface striations separated by nanoscale distances that precisely match the 4.2-nm length expected for peptides configured as α-helices as designed. These patterns extend unbroken across the widths (>50 nm) and lengths (>10 μm) of the fibers. Furthermore, the spacing of the striations can be altered predictably by changing the length of the peptides. These features reflect a high level of internal order within the fibers introduced by the peptide design process. To our knowledge, this exceptional order, and its persistence along and across the fibers, is unique in a biomimetic system. This work represents a step toward rational bottom-up assembly of nanoscored fibrous biomaterials for potential applications in synthetic biology and nanobiotechnology.

Author contributions: D.P. and A.M.S. contributed equally to this work; D.P., A.M.S., and D.N.W. designed research; D.P., A.M.S., and L.C.S. performed research; M.G.R. contributed new reagents/analytic tools; D.P., A.M.S., E.D.T.A., S.J.O., L.C.S., and D.N.W. analyzed data; and D.P., E.D.T.A., L.C.S., and D.N.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Ferely available online through the PNAS open access option.

Abbreviations: AM, ammonium molybdate; SAF, self-assembling-fiber; EM, electron microscopy; TEM, transmission EM; UA, uranyl acetate; WiAX, wide-angle x-ray diffraction.

Present address: Material Science Centre, University of Manchester, P.O. Box 88, Manchester M60 1QD, United Kingdom.

†To whom correspondence should be addressed. E-mail: d.n.woolfson@bristol.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0700801104.

© 2007 by The National Academy of Sciences of the USA

www.pnas.org/cgi/doi/10.1073/pnas.0700801104

PNAS | June 26, 2007 | vol. 104 | no. 26 | 10853–10858
also stabilize them with respect to thermal denaturation. Here, we describe detailed structural studies of the leucine-zipper building blocks of the SAF-p1/SAF-p2 design. Inspired by proposed mechanisms for the higher-order assembly of natural fibrous proteins (29–32), we argued that this might best be achieved by optimizing potential interprotofibril Coulombic (charge–charge) interactions, that is, rather than by introducing hydrogen-bonding interactions, which would require specific orientation of the partner protofibrils. To this end, we modeled contiguous copies of the SAF-p1 and SAF-p2 sequences (Table 1) as an extended dimeric coiled coil (26). [Leucine zippers, and coiled-coil proteins in general, share a seven-residue sequence repeat of hydrophobic (H) and polar (P) amino acids, HPPHPPP often designated abedefg. This directs the folding and subsequent assembly of amphipathic α-helices to form the so-called coiled coils. The first- and second-generation SAF peptides described herein have four heptads or 28 residues.] Inspection of the model suggested that two aspartic-acid side chains at consecutive b sites in the heptad repeat of SAF-p1 formed negatively charged pairs that would wind around the surface of the protofibril (Table 1). Therefore, to introduce potentially complementary positively charged pairs, we redesigned SAF-p2 to incorporate two arginine residues also spaced seven residues apart at two consecutive e sites. This resulted in peptide SAF-p2a (Table 1), which should combine with SAF-p1 to give protofibrils with matching acidic and basic patches on their surfaces (26). We refer to the SAF-p1:SAF-p2a combination as the second-generation SAFs.

**Rational Peptide Redesign.** To further stabilize lateral fiber assembly in the SAFs, we sought to enhance potentially complementary features presented on the surface of the leucine zipper building blocks of the SAF-p1/SAF-p2 design. Inspired by proposed mechanisms for the higher-order assembly of natural fibrous proteins (29–32), we argued that this might best be achieved by optimizing potential interprotofibril Coulombic (charge–charge) interactions, that is, rather than by introducing hydrogen-bonding interactions, which would require specific orientation of the partner protofibrils. To this end, we modeled contiguous copies of the SAF-p1 and SAF-p2 sequences (Table 1) as an extended dimeric coiled coil (26). [Leucine zippers, and coiled-coil proteins in general, share a seven-residue sequence repeat of hydrophobic (H) and polar (P) amino acids, HPPHPPP often designated abedefg. This directs the folding and subsequent assembly of amphipathic α-helices to form the so-called coiled coils. The first- and second-generation SAF peptides described herein have four heptads or 28 residues.] Inspection of the model suggested that two aspartic-acid side chains at consecutive b sites in the heptad repeat of SAF-p1 formed negatively charged pairs that would wind around the surface of the protofibril (Table 1). Therefore, to introduce potentially complementary positively charged pairs, we redesigned SAF-p2 to incorporate two arginine residues also spaced seven residues apart at two consecutive e sites. This resulted in peptide SAF-p2a (Table 1), which should combine with SAF-p1 to give protofibrils with matching acidic and basic patches on their surfaces (26). We refer to the SAF-p1:SAF-p2a combination as the second-generation SAFs.

**Biophysical Characterization.** As judged by CD spectroscopy (spectrum not shown) and EM (see below), SAF-2pa combined with SAF-p1 to form fibers that were at first sight similar to those originally observed for SAF-p1 and SAF-p2 (11). Consistent with an improved design, the critical concentration for fiber formation by SAF-p1/SAF-p2a improved to ~30 μM compared with ~60 μM for the original design, and the second-generation fibers were thicker (~70 ± 20 nm) and better defined (Fig. 2E). Furthermore, the temperatures up to which fibers could be assembled also improved: previously, SAF-p1:SAF-p2 mixtures yielded fibers only below room temperature and assembly was best at 5°C; the new SAF-p1:SAF-p2a design, however, could be assembled at up to 22°C. The solution-phase biophysical experiments are described in detail in ref. 26.

**Nanoscale Order in the Matured Fibers.** The change in thermal-unfolding behavior between the first- and second-generation design fibers prompted us to look for differences in their structure and morphology by negative-stain TEM. Like the first-generation fibers (Fig. 2 A and B; ref. 11), the redesigned fibers were linear, nonbranched, and extended for many microns. However, the latter were thicker, which, along with the improved thermal stability, is fully consistent with the redesign principles.

### Table 1. SAF peptide sequences

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAF-p1</td>
<td>K ILALKK AQALKQ GQAALQK GAALQK</td>
</tr>
<tr>
<td>SAF-p2</td>
<td>K ILALKE AQALKE GQAALKE GAALKE</td>
</tr>
<tr>
<td>SAF-p2a</td>
<td>K IRALKQ LNARLQ IAALYE IAALEQ</td>
</tr>
<tr>
<td>Ac-SAF-p1-NH2</td>
<td>Ac-K ILALKQ AQALKQ GQAALQK GAALQK</td>
</tr>
<tr>
<td>Ac-SAF-p2a-NH2</td>
<td>Ac-K IRALKQ LNARLQ IAALYE IAALEQ-NH</td>
</tr>
<tr>
<td>SAF-p1-ext</td>
<td>K ILALKK AQALKQ GQAALQK GAALQK</td>
</tr>
<tr>
<td>SAF-p2a-ext</td>
<td>K IRALKQ LNARLQ IAALQE IAALQE</td>
</tr>
</tbody>
</table>

Differences between successive designs are highlighted by bold italics.
In addition, and curiously, the second-generation fibers showed clear evidence of order above that in the first-generation structures (compare Fig. 2 A and C). Specifically, the second-generation fibers showed well defined periodic light and dark banding patterns (referred to here as ‘striations’) perpendicular to the long fiber axis. These patterns persist along the entire lengths (approximately tens of microns) of the fibers and extend completely across their widths perpendicular to the long fiber axes (Fig. 2C).

Like the reports for 2D crystals of truncated prion proteins (33) and fibrous collagens (34), the dark striations probably result from the uranyl acetate (UA) acting as a positive stain, interacting with anionic moieties. This is supported by the observation in our system that copious washing of the stained fibers with water on the carbon TEM grids removed background (negative) stain, but failed to remove the striations. Consistent with this finding, similar patterns with the same spacing of striations (see below) were observed for fibers stained with sodium phosphatase buffer and stained with a solution of phosphotungstic acid (a cationic stain, like UA) and also with the anionic stains ammonium molybdate (AM) and dimolybdate. Interestingly, similar patterns with comparable relative widths of the light and dark bands were observed for fibers stained with AM followed by washing with distilled water and a second staining with UA [see supporting information (SI) Appendix, that is, rather than two overlapping patterns or one with wider dark bands. A possible explanation of this observation is that the anionic and cationic peptide moieties likely to be interacting with UA and AM are proximal.

**The Striation Spacings Match the Lengths Expected for Folded Peptides.** From our earlier studies (11), X-ray diffraction from drawn fibers of the first-generation SAFs indicates that the SAF peptide assemble as designed: that is, each peptide adopts an α-helical conformation and these pair to form twisted two-stranded coiled-coil structures as first postulated by Francis Crick (35). In addition, the axes, or long directions, of these coiled coils align within the fiber along the long fiber axis itself. These architectural features are also evident in the diffraction patterns from fibers prepared from the redesigned peptides as discussed below.

Therefore, it is reasonable that the striations observed by EM relate to features that repeat in the peptides and/or along the coiled-coil superhelical axis. Consistent with this concept, Fourier transform analysis of individual fibers from the positive-stain electron micrographs revealed that the striations were spaced 4.2 nm apart (SD 0.13 nm over 71 measurements). This experimental measurement closely matches the distance that a 28-residue SAF peptide would cover in a fully folded α-helical coiled-coil conformation (i.e., 28 × 0.148 nm = 4.144 nm, as the rise per residue in a coiled coil is 0.148 nm), which also corresponds to ~1/4 of the supercoil pitch for a two-stranded coiled coil (36).

Together, the x-ray and TEM data suggest that the striations highlight the longitudinal repeat of the peptides along the fiber main axis. Furthermore, as individual striations cover the entire width of the fibers without interruption, it is likely that the helical peptides are also in register across the widths of the fiber.

To test the correspondence between peptide length and the separation between striations further, a fifth heptad was added to the second-generation design (Table 1) extending the peptides to 35 aa, which should increase the length of each individual helix to ~5.2 nm. Fibers assembled from these third-generation peptides were imaged by TEM and striations were observed (Fig. 2E and F). Consistent with our design and prediction, Fourier transform analysis returned a striation length of 5.2 nm (SD 0.07 nm over 23 measurements).

Other groups have described longitudinal stripes on negatively stained images of designed fibers (37, 38) and natural proteins (39). It is important to emphasize the second- and third-generation SAFs show identical striation patterns upon positive and negative staining. Thus, we believe that any signal from longitudinal stripes of negatively stained SAFs, if any exist on the SAFs at all, is too weak to compare with the lateral striations.

**Internal Nanoscale Order.** The wide-angle x-ray diffraction (WAX) pattern from an oriented, partly dried fiber of second-generation SAFs is shown in Fig. 3. The overall x-ray diffraction features are similar to those predicted (35), and seen (40) for α-helical coiled-coil conformations, and, in particular, double-stranded coiled coils.
The prominent second-order layer line (marked by L in Fig. 3) bears a relationship to the relatively strong 0.54-nm signal: 0.54 nm is 108(4) nm, and the electron diffraction pattern is consistent with the observed 0.54-nm orthogonal striations observed in the fibers by using real-space EM imaging.

For completeness, we note that the true crystallographic periodicity of a slowly twisting coiled-coil structure of this kind is in the region 15–18 nm (35). Furthermore, in our case the coiled coil is not continuous, but will have scattering dislocations approximately every 4 × 1.036 = 4.144 nm from the termini of the peptides that make up the fiber. In both cases, however, the relatively weak scattering expected would require more-sensitive images and would have to be sought in the ultra-low to low-angle x-ray diffraction regions.

A Molecular Model for the SAF Structure. Putting the above observations and analyses together, we derived a model for the structure of the SAFs in which canonical two-stranded α-helical coiled coils are aligned along the long axis of the fibers and packed hexagonally around this axis (Fig. 4). In addition, from other experiments that indicate the fibers are polar, the coiled coils are aligned parallel, i.e., in one direction along the fiber (42).

To compare this model and the observed experimental data, a simulated diffraction pattern (Fig. 3) was generated by using the program CLEARER (43) and the above unit cell dimensions and model coordinates for a unit cell. The positions of the diffraction signals in the simulated and experimental diffraction patterns match very closely (Fig. 3). However, the experimental data do not allow a detailed molecular description of the coiled coils with accurate positions of side chains; therefore, the intensities in the experimental and calculated patterns will not necessarily compare well. Nonetheless, this simulation strongly supports the overall arrangement of coiled coils in the model, namely, longitudinal registered coiled-coil protofibrils arranged laterally on a hexagonal lattice (Fig. 4).

Origin of the Positive Staining and the Resulting Striations. The apparent colocation of the anionic and cationic stains observed in the TEM experiments (even with the inherent low resolution of ~2 nm with these methods) possibly suggests that they are highlighting proximal oppositely charged moieties in the fibers. One possibility is that the C-terminal extremities of each peptide that abut in the design, i.e., UA could be interacting with the carboxylates of the C terminus, and AM could be interacting with the primary amines of the N termini. To probe this possibility, we synthesized capped versions of SAF-1 and SAF-2, Ac-SAF-1-NH2 and Ac-SAF-2a-NH2, in which the terminal charges were replaced by neutral amide bonds (Table 1). Unfortunately, the mixture of both capped peptides did not render fibers. When the individual capped peptides were mixed with their complementary uncapped peptide (i.e., Ac-SAF-1-NH2 with SAF-2a, and Ac-SAF-2a-NH2 with SAF-1) fibers were observed, but only in preparations incubated at 4°C. This destabilization is fully consistent with our design principles; i.e., that the ends of the (uncapped) peptides are close enough in space to abut and form stabilizing intermolecular salt bridges (Fig. 1) (11). Perhaps surprisingly, however, the fibers containing one capped peptide were still striated (see SI Appendix), and the distances
between striations was the same as that in normal second-generation fibers. Thus, it is unlikely that the stains are interacting with the peptide extremities, at least not with the N- and C-terminal charges.

Another possibility is that the stains are interacting with the charged side chains on the surfaces of the folded peptides. Two groups of such side chains can be considered: first, there are those at the e and g positions of the heptad repeats (Table 1), which help direct and cement the SAF-p1:SAF-p2 heterodimers. However, these are common to both the first-, second-, and third-generation SAFs, and striations are observed only for the second and third generations. Second, there is the additional quartet of aspartic acid and arginine residues introduced between the first- and second-generation designs. In this case, cationic UA stain could interact with one or both of the aspartic acid residues of SAF-p1 and the anionic AM stain could interact with the complementary arginine residues of SAF-p2a. These groups of charged residues are within 1 nm of each other along the coiled-coil axis. Thus, within the resolution of the stained TEM, it is reasonable to consider these groups of residues as one cluster of negative and positive charge repeated every peptide length along the fibers. We favor this model for the origin of positive staining to give lateral striations. However, the removal of this cluster, i.e., effectively partially reversing the first- to second-generation experiment, drastically reduces the stability of and order in the fibers, and it is difficult to see how this hypothesis could be tested by mutagenesis alone.

Conclusion
To summarize our studies and results, we have built on our previous design of a SAF system considerably. Specifically, after noting that the first-generation fibers were thickened we sought to explore and promote this by engineering improved protofibril–protofibril interactions. This rational redesign results in second-generation fibers that are indeed thickened further and stabilized. Moreover, these redesigned fibers show intriguing nanoscale features in positive-stain TEM and considerable internal nanoscale order as judged by WAX. The features observed by TEM are striations orthogonal to the long fiber axis. These extend coherently along the whole length (≥10 μm) of the fibers and run uninterrupted across the widths of the fibers (≥50 nm). The separation between the striations precisely matches the lengths of the peptide building blocks as originally designed (11), that is, configured as α-helices and assembled end to end along the length of the fiber. In further support of this arrangement, WAX reveals contiguous coiled coils along the long axis of the fibers, which are packed hexagonally across the widths of the fibers. A resulting molecular model for the structure has been derived and validated by back-calculating the WAX pattern.

We propose that the structure and order in the SAFs arises through the aforementioned protofibril–prototibril interactions introduced in the peptide design process: weak, complementary interactions are repeated with helical symmetry along the long axis of the fibers; these combine cooperatively to cement prototibril–prototibril interactions throughout the fibers. As a result, the fibers are thickened, stabilized, and highly ordered.

To our knowledge, the level of organization that we observe has not been reported for any other biomimetic fibrous peptide or protein system. The ultrastructure observed is comparable, at least qualitatively, to that seen in natural protein fibers, for instance, in certain natural protein fibers, including some collagen, fibrins, and lamins (27, 34, 44, 45). In these cases, this external order reflects high internal organization within these highly evolved systems. Our fibers, however, are only the second and third iterations in a rational design process. This unexpected mimicry in a relatively simple and designed binary peptide system is intriguing and may shed light on both natural and synthetic self-assembly processes in general. In comparison to other designed α-helix-based fibers, the SAFs display considerably more order in their assembled structures (7, 17, 46); although we note longitudinal stripes are observed by TEM in one system (38), and hexagonal packing of helical fibrils has also been demonstrated by WAX in another (12). In addition, because ours is a dual-peptide system, whereas the others rely on just one peptide to self-assemble, we have more control over the assembly process (42). This ability to control assembly brings further utility to the system as it allows additional functionalized peptides to be added to self-assembly mixtures. In turn, it allows fibers morphology and function to be engineered (7, 25, 42, 47–49). In view of current attention being given to the generation of biosprirals for applications in nanobiotechnology (9, 46, 50), our findings offer promise for nanoscale engineering of self-assembling systems from the bottom up in water; for example, the directed incorporation of functionalized peptides regularly spaced along the fibers can be envisaged from the observations presented here.

Materials and Methods
Peptide Synthesis. Peptides were made chemically and purified as described (11, 42). The final products were identified by MALDI-TOF MS (Micromass, Manchester, U.K.), MS [M + H]+: SAF-p1, m/z 3174 (calc), 3175 (found); SAF-p2, m/z 3128 (calc), 3129 (found); SAF-p2a, m/z 3325 (calc), 3326 (found); SAF-p1-ext, m/z 3927 (calc), 3928 (found); SAF-p2a-ext, m/z 4079 (calc), 4080 (found); Ac-SAF-p1-NH₂, m/z 3215 (calc), 3214 (found); and Ac-SAF-p2a-NH₂, m/z 3367 (calc), 3367 (found).

EM. Samples for TEM were prepared as described (11), except phosphotungstic acid-stained fibers that were grown in phosphate buffer and stained with a 1% solution of the stain in water were used. Fiber samples for the third-generation peptides and the capped versions of the second-generation peptides were prepared similarly except that fibrollogenesis was performed overnight at 37°C and 5°C, respectively. In all cases, the final

![Hexagonal-packing model for the designed SAF coiled coils as deduced from the x-ray diffraction pattern (Fig. 3). (A) Schematic view of a fiber cross-section. The coiled-coils are packed in an hexagonal lattice of size 1.824 nm (a-axis). (B) Computer model showing two dimensions of the hexagonal lattice. (C) Additional 3D schematic section illustrating the proposed organization of protofibrils in the fibers. The black sections orthogonal to the fiber axis represent an EM stain interacting with the fibers with a periodicity of 4.2 nm, as seen with the 2-nm resolution inherent to the technique.](image-url)
concentration of each SAF peptide was 100 μM. Micrographs were taken at ×1,000 and ×40,000 by using software from Gatan (Pleasanton, CA). Images were converted to the Image2000 Software format (51) and then examined with XIMDISP (52). Fourier transforms of selected boxed, padded, and floated regions were interpolated and calculated interactively to search for diffraction spacings. A catalase calibration specimen (Agar Scientific, Essex, U.K.) was processed in the same way as the samples to confirm the calibration of the microscope at all magnifications.

X-Ray Fiber Diffraction. The x-ray fiber diffraction pattern for the SAF-p1:SAF-p2a assembly was recorded for a sample prepared under the stretched frame procedure (53). A droplet of the fresh solution containing 700 μM of each peptide was placed between the ends of two wax-filled capillaries distant from 1.5 mm from each other and allowed to dry slowly at 22°C, yielding a dried rod composed of partially aligned fibers. X-ray fiber diffraction images were collected using a CuKα rotating anode source (wavelength 0.15418 nm) and a R-AXISIV detector (Rigaku, Tokyo, Japan). The sample was kept at room temperature and placed at a distance of 300 mm to the detector, then exposed for 10 min over a rotation of 0.5° angle.

Model Building. Atomic coordinates for a SAF-p1:SAF-p2a two-stranded coiled were generated by using the program MAKECCSC kindly provided by Gerald Ofer (personal communication) with standard parameters for a double-stranded coiled coil (24): pitch, 15 nm; radius of the supercoil, 0.47 nm; helical rise per amino acid unit, 0.1495 nm. Successive side chains were placed in their favored conformations as observed in a side-chain rotamer library created from a database of dimeric coiled coils of known 3D structure (ref. 22 and D.P., Gerald Ofer, and D.N.W., unpublished work). The output pdb file consisted of a 140-residue dimeric coiled coil with the amino acid sequences of SAF-p1 (chain A) and SAF-p2a (chain B) repeated five times each. PyMOL (www.pymol.org) was used to generate the symmetry related objects (hexagonally packed coiled coils) based on the experimentally measured unit-cell dimensions.

Calculation of Simulated X-Ray Diffraction Pattern from Model Coordinates. A simulated diffraction pattern was calculated from the modeled coordinate file by using the program CLEARER (43). The coordinates were arranged within a unit cell with dimensions a = b = 1.824 nm, c = 1.036 nm, α = β = 90°, γ = 120°, and the long axis of the fiber was oriented parallel to c. The simulated diffraction pattern was calculated by adding the intensities from all reciprocal lattice points, taking account of the fiber-axis direction, crystallite dimensions, disorder, and diffraction geometry. Diffraction settings were identical to the experimental settings.

We thank Gerald Ofer, Sam MacDonald, and John Squire for valuable discussions. A.M.S. received studentship support from the Engineering and Physical Sciences Research Council. D.N.W. and L.C.S. were supported by Biotechnology and Biological Sciences Research Council Grant BB4676. S.J.O. was supported by Medical Research Council Grant G0309584.