JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

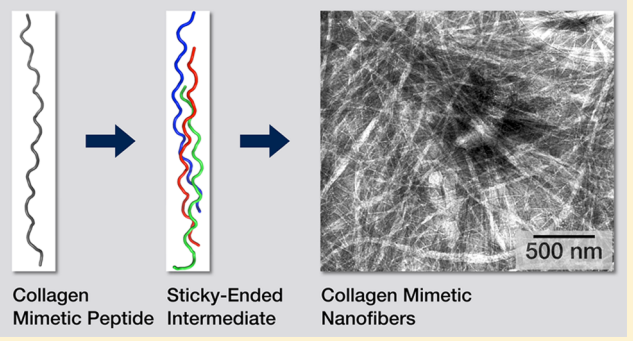
Self-Assembly of Fiber-Forming Collagen Mimetic Peptides Controlled by Triple-Helical Nucleation

Biplab Sarkar,[†] Lesley E. R. O'Leary,[†] and Jeffrey D. Hartgerink*[,†,‡]

[†]Department of Chemistry and [‡]Department of Bioengineering, Rice University, Houston, Texas 77005, United States

Supporting Information

ABSTRACT: Mimicking the multistep self-assembly of the fibrillar protein collagen is an important design challenge in biomimetic supramolecular chemistry. Utilizing the complementarity of oppositely charged domains in short collagen-like peptides, we have devised a strategy for the self-assembly of these peptides into fibers. The strategy depends on the formation of a staggered triple helical species facilitated by interchain charged pairs, and is inspired by similar *sticky-ended fibrillation* designs applied in DNA and coiled coil fibers. We compare two classes of collagen mimetic peptides with the same composition but different domain arrangements, and show that differences in their proposed nucleation events differentiates their fibrillation capabilities. Larger nucleation domains result in rapid fiber formation and eventual precipitation or gelation while short nucleation domains leave the peptide soluble for long periods of time. For one of the fiber-forming peptides, we elucidate the packing parameters by X-ray diffraction.



INTRODUCTION

From building mechanically adaptable body armor in the fish *Arapaima gigas*[1,2] to the changing color and light-sensitivity of reindeer eyes,[3] the highly ordered fibrous network formed by the extracellular protein collagen has evolved to execute a diverse array of physiological functions in every strata of the animal kingdom. In the human body, the collagen network is the structural basis of tissues such as skin, bone, ligaments, and tendon,[4] and collagen content may account for up to three-quarters of the dry weight of these tissues.[5]

Formation of collagen is one of the best examples of multihierarchical assembly in nature. Regardless of the tissue, the fundamental supramolecular unit of collagen assembly is the triple helix, made up of three protein chains interlocked in a right-handed superhelix. The characteristic triple helical folding in collagen is the consequence of its repetitive amino acid sequence $(Xaa-Yaa-Gly)_n$. Glycine is required at every third residue both to provide backbone hydrogen bond donation to the carbonyl of a cross strand amino acid in the Xaa position and also for the small steric bulk presented by its lack of side chain.[6] Amino acids found in the Xaa and Yaa positions have high propensities for being proline (Pro, P) and 4-hydroxyproline (Hyp, O), respectively.[7] These cyclic amino acids help to stabilize the triple helix due to their partial preorganization as a polyproline type II helix, which is the proper conformation necessary to form the triple helix.[8]

Axial Charged Pairs in Collagen Triple Helix. While glycine is required every third amino acid, the selection of amino acids in the Xaa and Yaa positions is more flexible. In particular, the Xaa position is frequently occupied by negatively charged amino acids, whereas the Yaa position is frequently held by positively charged amino acids.[7] On their own, these substitutions of proline and hydroxyproline for charged amino acids result in significant destabilization of the triple helix. However, proper geometric interstrand pairing of oppositely charged amino acids can result in triple helices with thermal stabilities comparable to $(Pro-Hyp-Gly)_n$ sequences.[9−11] We call this geometric relationship an *axial* charged pair, as the side chains of the oppositely charged amino acids lie roughly parallel to the triple helical axis (Figure 1). In order to obtain this geometry a positively charged residue in the Yaa position of the leading strand needs to pair with a negatively charged residue in the Xaa position of the middle strand, one triplet displaced toward the C-terminus. Similarly, the middle strand can couple to the lagging strand and the lagging strand can couple with the leading strand. The helical symmetry makes each of these pairs equivalent. Such axial interactions have been successfully used to design various collagen heterotrimers.[11−14]

Hierarchical Self-Assembly. Higher order supramolecular assembly of collagen triple helices depends on the tissue. For example, in tendon, triple helices continue to assemble into small fibrils and eventually into macroscopic semicrystalline fiber-bundles.[4,15−17] The triple helices pack in a pseudohexagonal fashion, where the lateral distance among neighboring triple helices is roughly the same.[18−20] Researchers have also suggested the presence of a relatively disordered core within an ordered shell of the fiber bundle.[21] Models of collagen packing

Received: May 1, 2014
Published: October 6, 2014

   dx.doi.org/10.1021/ja504377s | J. Am. Chem. Soc. 2014, 136, 14417−14424

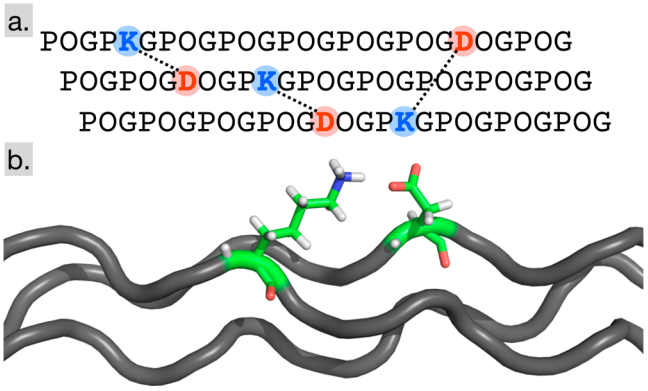


Figure 1. Axial lysine-aspartate salt bridges in collagen triple helix. (a) Rules governing such axial charge pairs. Amino acid residues in each vertical plane are approximately in the same cross section of the triple helix. Single letter amino acid code has been used (O: 4-hydroxyproline). (b) Model of an axial charge pair in a triple helix.

reported to date are mainly based on the fiber diffraction data collected from tendon collagen, which display signatures of both ordered and disordered regions in the fiber bundle.

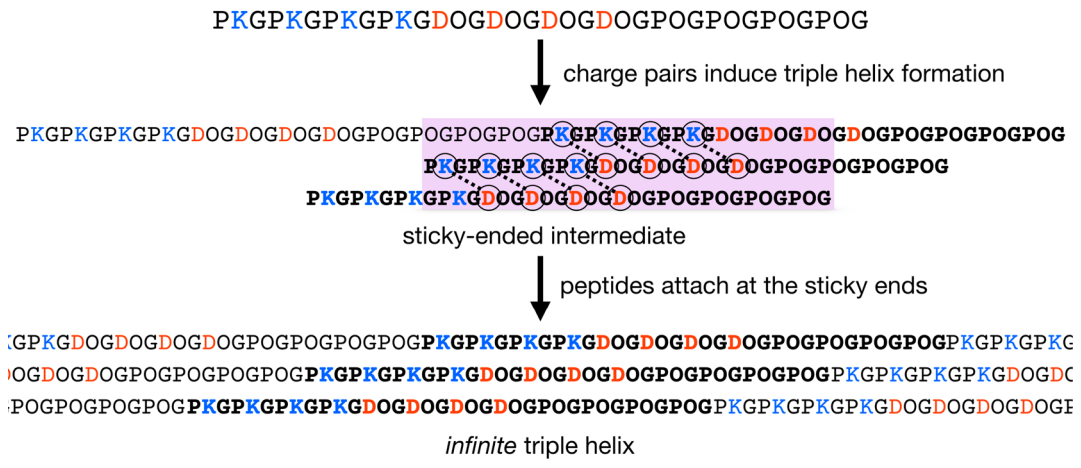
Fibrillar Collagen Mimetic Peptides. Mimicking hierarchical assembly of collagen has long been a desirable target for chemists, both for fundamental understanding of collagen assembly and for potential biomaterials applications. In spite of some important advances toward that target,[22−30] a general design scheme for such biomimetic assembly remains unavailable. We recently reported a synthetic fibrillar collagen mimetic peptide that reproduces several stages of collagen hierarchical assembly including triple helix formation, fiber formation, and gelation.[30] The peptide, $(PKG)_4(POG)_4(DOG)_4$ (referred to here as **F0**), is only 36 amino acids long, in contrast to collagen type I, which is more than 1,000 amino acid residues long. The fiber bundles formed by the peptides **F0** were partially aligned as demonstrated by dried fiber diffraction patterns. Well-resolved fiber diffraction data from aligned triple helices in a native hydrated state of fibrillar collagen mimetic peptides has not been reported so far, and such information may potentially aid our understanding of superhelical twist of triple helices and parallel packing of the triple helices into well-ordered fiber bundles of these peptides.

Blunt-Ended Self-Assembly. The tridomain nature of the sequence of the peptide **F0**, with positive $(PKG)_4$, neutral $(POG)_4$, and negative $(DOG)_4$ regions, was based on an earlier collagen mimetic peptide reported by Rele *et al.*[24] They demonstrated that $(PRG)_4(POG)_4(EOG)_4$ formed D-periodic microfibers. The peptide assembles into a canonical blunt ended triple helix while the positive and negative termini of the helix induce it to pack in an oriented manner, such that two adjacently packed triple helix termini can favorably interact via electrostatic attraction. The packing of the triple helical *protomers* results in fibrillation. We call this fibrillation pathway the *blunt-ended assembly*. Jiang *et al.* have also reported the preparation of elegant two-dimensional nanoscale sheets from similar collagen mimetic peptides, where arginine was replaced by the positively charged amino acid 4-aminoproline.[31]

Sticky-Ended Self-Assembly. Previously, we proposed an alternative fibrillation hypothesis, based on the studies of axial charge pair contacts in collagen triple helix (Figure 1) carried out in our lab.[30] According to this *sticky-ended fibrillation* model, interstrand axial charged pairs formed between the $(PKG)_4$ domain and the $(DOG)_4$ domain establish a large offset in the triple helix and the unsatisfied backbone hydrogen bonds from the exposed glycines create *sticky ends*, not unlike those used in DNA arrays[32−34] or coiled-coil assemblies[35−37] (see Scheme 1). The resulting stagger in the sticky-ended triple helical intermediate facilitates further attachment of peptides at both termini, leading to an *infinite* triple helix. Recently, we demonstrated by an NMR analysis that these types of peptide offsets can indeed be formed when the charged pairs are appropriately designed.[38] We discuss the hypothesis in more detail in the Results and Discussion section.

An important caveat in the comparison of **F0** to the peptide reported by Rele *et al.* is that the structure and fibrillation mechanism is likely dependent on the nature of the charged residues involved. As we have discussed previously,[39] there are significant differences between interstrand charged pairs (lysine-aspartate vs arginine-glutamate) in the collagen triple

Scheme 1. Sticky-Ended Fibrillation Model[a]



[a]The concept of sticky-ended fibrillation is shown above using the peptide $(PKG)_4(DOG)_4(POG)_4$ (**F1**, see Table 1). Initial formation of axial charged pairs may result in a four-peptide intermediate with a 28 amino acid triple helical region (shaded box) and extensive sticky-ends, which are not hydrogen bonded. Further assembly of peptides results in a continuous hydrogen bonding network within the *infinite* triple helix where two-thirds of all possible charged pairs are satisfied (also see Supporting Information Figure S4).

Table 1. Sequence, Solubility, and Morphology of Peptides, Compared with Predicted Fibrillation[a]

	Peptide Sequence	Prediction Blunt-ended Assembly	Prediction Sticky-ended Assembly	Sticky-ended Nucleation Domain (length in aa, 3-peptide, 4-peptide)	Solubility, Morphology
F1	$(PKG)_4(DOG)_4(POG)_4$	no fibrillation	fibrillation	**20, 28**	precipitate,[b] nanofibers
F2	$(POG)_4(PKG)_4(DOG)_4$	no fibrillation	fibrillation	**20, 28**	precipitate,[b] nanofibers
A1	$(DOG)_4(PKG)_4(POG)_4$	no fibrillation	no fibrillation	**4, 20**	soluble,[c] amorphous
A2	$(POG)_4(DOG)_4(PKG)_4$	no fibrillation	no fibrillation	**4, 20**	soluble,[c] amorphous
F0	$(PKG)_4(POG)_4(DOG)_4$	fibrillation	fibrillation	**16, 32**	hydrogel, nanofibers

[a]**F0**–**F2**, fiber-forming peptides; **A1**–**A2**, amorphous peptides. [b]**F1** and **F2** do not precipitate out of solution in high ionic strength buffers. [c]**A1** and **A2** precipitate out of the solution in a few weeks. Single letter amino acid codes are used for peptide sequence nomenclature (O is 4-hydroxyproline). Peptide **F0** has been published before by our lab.[30] Blunt-ended assembly refers to the hypothesis according to Rele *et al.*[24]

helix. Thus, it is possible that the fibrillation model discussed in this work is limited to fibrillar collagen mimetic peptides employing lysine-aspartate and lysine-glutamate charged pairs. The use of arginine has a lower propensity to form specific charge pairs in the triple helix, and thus, while the two hypotheses described here are quite different, either may function in the context of their corresponding amino acid composition.

One challenge with any system designed to self-assemble into a very large construct is that analytical methods to probe the mechanism of assembly are extremely challenging as solution state techniques cannot typically be employed. In this work, we explore fibrillation mechanism through the analysis of a series of closely related self-assembling collagen-like peptides whose sequences have identical composition but different domain arrangement. From our study, we conclude the fiber formation in these tridomain fibrillar collagen mimetic peptides happens through a sticky-ended intermediate, controlled primarily by the kinetics of assembly rather than the stability of the final fiber formed.

PEPTIDE DESIGN

In order to test the viability of both *blunt-ended assembly* and *sticky-ended fibrillation* of tridomain collagen mimetic peptides, we studied four additional peptides in addition to **F0**. In these studies, we have carried out a process of domain swapping where the positive, neutral, and negatively charged domains are scrambled (see Table 1). Among the new peptides synthesized for this study, in **F1** and **F2**, the positive domain precedes the negative domain. In the other two, **A1** and **A2**, the negative domain precedes the positive one.

Of all five peptides, only our original peptide, **F0**, has both charged domains located at the termini. Therefore, according to the blunt-ended assembly hypothesis, none of the four new peptides should form fibers. However, in the sticky-ended assembly scenario, depending on the organization of the given peptide, the initial assembly of the triple helices may lead to different extents of interpeptide glycine-based hydrogen bonding and a different number of axial charged pairs formed, affecting their fiber-forming capability. We compare in the following sections how these peptides differ in their solubility, fiber-forming capability, and long-term crystallinity.

RESULTS AND DISCUSSION

Synthesis and Sample Preparation. Peptides were prepared by solid phase peptide synthesis as described in the Methods section, purified by HPLC, and characterized by mass spectrometry (see Supporting Information for HPLC and MS). Peptides were dissolved in 10 mM phosphate buffer at pH 7 and heated to 85 °C for 15 min after which they were slowly cooled to room temperature. Peptide **F0** subsequently forms a hydrogel, while **F1** and **F2** rapidly precipitate at room temperature (within 1 h for a 1% solution). **F1** and **F2** precipitated out of the solution at a variety of concentrations (0.1–2%, precipitation time is longer at lower concentrations). They also precipitated out of the solution in a variety of buffers, for example, 10 mM phosphate (pH 7), 10 mM Tris (pH 7), 10 mM HEPES (pH 7), and 10 mM MOPS (pH 7), as well as out of a pure aqueous solution (no buffer) after being well dissolved at high temperature. In contrast, peptides **A1** and **A2** remain soluble for weeks.

Solubility and Effect of Ionic Strength. The role of charged pairs in controlling fibrillar self-assembly in **F1** and **F2** is highlighted by the effect of increased ionic strength on their solubility. When dissolved in high-ionic-strength buffers such as phosphate buffered saline (PBS, 10 mM phosphate with 150 mM sodium chloride, pH 7), as well as in saline Tris buffer (10 mM Tris with 150 mM sodium chloride, pH 7) peptides **F1** and **F2** do not precipitate out of solution even after several months. Under these higher ionic strength conditions, charged pair interactions are partially shielded, inhibiting sticky-ended nucleation. The inhibition of sticky-ended nucleation, in turn, prevents nanofiber self-assembly, leading to enhanced solubility. In contrast, when the ionic strength of the buffer is relatively low, for example, in 10 mM phosphate, **F1** and **F2** precipitate out of solution rapidly upon cooling (please see Supporting Information Figure S11) and fiber formation is observed in electron microscopy (see below). **F0** behaves somewhat differently in that it forms a hydrogel under both high and low ionic strengths. However, gel formation is slower at high ionic strength compared to gels formed at low ionic strength. **A1** and **A2** remain soluble under all conditions.

Transmission Electron Microscopy. This dichotomy observed in the solubility of the tested peptides is continued upon analysis by negatively stained transmission electron microscopy (TEM) and optical polarized microscopy (OPM). In TEM, Peptides **F1** and **F2** reveal a highly fibrous

dx.doi.org/10.1021/ja504377s | J. Am. Chem. Soc. 2014, 136, 14417–14424

morphology with several micrometer long fiber bundles (Figure 2a,b) as the dominant nanostructure observed. In contrast,

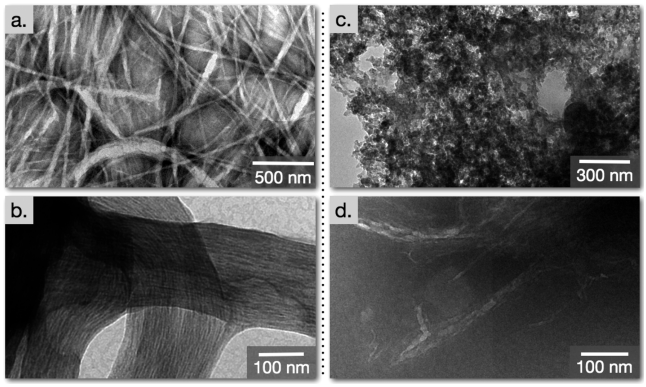


Figure 2. Comparison of the nanoscale morphology of collagen mimetic peptide assembly in negatively stained transmission electron microscopy (TEM). (a) **F1** and (b) **F2** are shown to form fibers, whereas (c) **A1** and (d) **A2** reveal amorphous aggregates.

several-week-old samples of peptides **A1** and **A2** show presence of primarily amorphous aggregates (Figure 2c,d) which may simply be formed upon sample drying.

Optical Polarized Microscopy. The morphology of F*x* and A*x* peptides are different at microscale as well. In optical polarized microscopy (OPM), **F1** and **F2** both were observed to form a birefringent needlelike morphology (Figure 3a,b). On

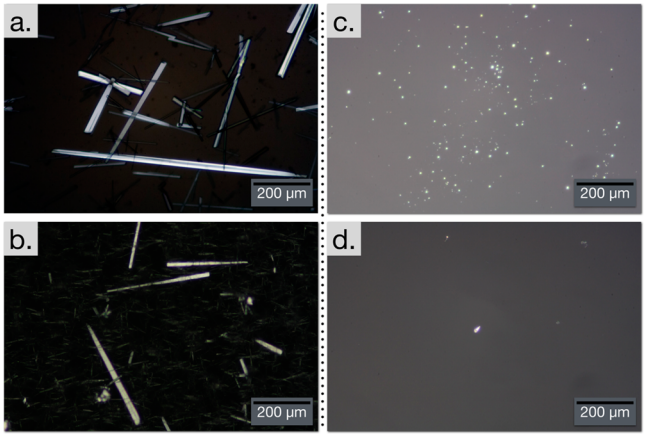


Figure 3. Comparison of the microscale morphology of collagen mimetic peptides in optical polarized microscopy (OPM). (a) **F1** and (b) **F2** form birefringent needles; (c) **A1** and (d) **A2** form amorphous aggregates.

the other hand, when after a few weeks **A1** and **A2** slowly precipitate out of the solution, only amorphous aggregates were detected in OPM (Figure 3c,d). We ascribe the needlelike appearance of **F1** and **F2** microstructures to the inherent high order present in the nanofiber packing. **A1** and **A2** precipitate after several weeks, probably due to nonspecific interactions leading to aggregation. Among the fiber forming peptides, we were able to obtain well-resolved fiber diffraction data from semicrystalline needles of **F1**.

The difference between these two classes of peptides are also manifest in their triple helical stability as measured by circular dichroism experiments. **F1** and **F2** have higher molar residual ellipticity (MRE) values as compared to **A1** and **A2**, indicating a larger triple helical population. Furthermore, **F1** and **F2** have higher melting temperatures compared to **A1** and **A2** (see Figure 4).

Sticky-Ended Fibrillation. The fiber-forming peptides examined in this study are designed to utilize axial charged pair hydrogen bonding to create a sticky-ended approach to triple helix elongation. Scheme 1 illustrates the formation of a sticky ended nucleation for peptide **F1** and its elongation into a collagen-like peptide fiber highlighting the interpeptide axial charged pairs which are formed in the nucleation domain. All five peptides have the exact same overall amino acid composition. Further, in all cases, the infinite triple helices composed of these peptides allow two-thirds of all lysine and aspartate residues to form these favorable axial charged pairs (see Figure S4 in the Supporting Information) and all backbone glycine hydrogen bonding is satisfied in the infinite length approximation (the exception being the fiber tips). Thus, should the fibrillation take place in all four peptides, the thermodynamic stability of the final fiber should be very similar. Despite these similarities, the new set of studied collagen-like peptides could easily be divided into two groups, one (**F1** and **F2**) which has low solubility, forms peptide nanofibers, and forms birefringent ordered needles at the microscale, and the other (**A1** and **A2**) which has good solubility and forms amorphous aggregates at nanoscale and microscale after a few weeks.

The differences between the two groups of peptides are revealed when examining the putative early stages of fiber formation, for example, when only three peptides have associated with one another (Figure 5). For fiber forming peptides, for example, **F1**, the staggered intermediate has a large contiguous triple helical domain (enabling canonical interpeptide hydrogen bonding), which is further stabilized by the formation of interstrand lysine-aspartate axial charged pairs. In contrast, for **A1**, fulfillment of possible interstrand charged pairs leads to a very unstable intermediate having a very short triple helical nucleation region. This remains a low probability structure, and therefore, fiber propagation is not observed.

However, more peptides can associate at the *sticky-ended* termini of the staggered intermediate in the cases of **F1** and **F2** satisfying more hydrogen bonds and thus stabilizing the intermediate. Despite the elongation of the triple helix, the termini remain staggered and thus the triple helix can continue to elongate indefinitely.

The difference in the fibrillation capabilities of these collagen mimetic peptides can be explained by the kinetics of fiber formation. Kinetics of triple helical elongation in this case is limited by the stability of initial staggered intermediates. The number of amino acids forming a contiguous triple helix varies in the sticky-ended triple helical nuclei. Four-peptide staggered assemblies of the fiber forming peptides **F1** and **F2** are seen to have relatively large contiguous triple helical domains (28 amino acid residues) while amorphous peptides have relatively smaller triple helical domains (20 amino acid residues) (see Figure 6). Stability of the nucleating sticky-ended species would depend on the length of the contiguous triple helical domain, referred to as "nucleation region", as canonical interstrand hydrogen bonding can form only in that region. In addition, unlike **A1**−**A2**, all the interstrand charged pairs are located in the triple helical region for **F1** and **F2**; thus, those charged pairs are more likely to contribute to the stability of the intermediate.

Comparison with Blunt-Ended Assembly Scenario. When we compare the fiber-forming capabilities of **F1**, **F2** and

dx.doi.org/10.1021/ja504377s | J. Am. Chem. Soc. 2014, 136, 14417−14424

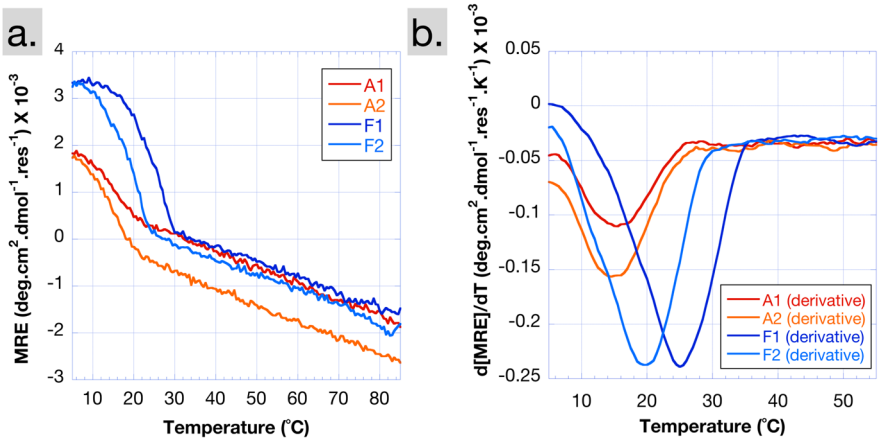


Figure 4. (a) Circular dichroism melting profiles and (b) corresponding derivative plots of the peptides **F1**, **F2** and **A1**, **A2** prepared at a concentration of 0.1% (w/v) in 10 mM phosphate buffer, pH 7 and using a path length of 0.1 cm. See Methods for details.

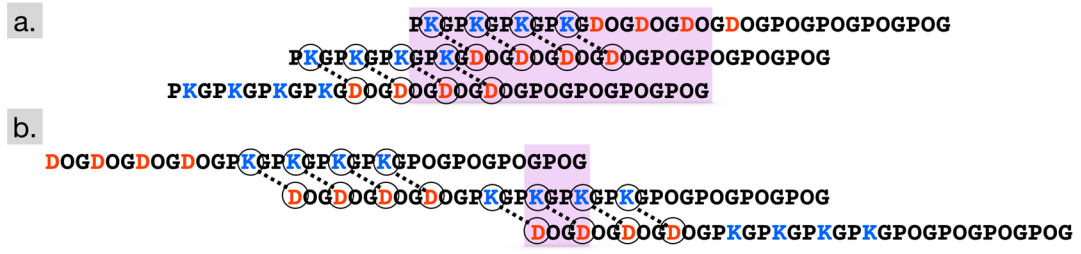


Figure 5. Putative sticky-ended three-peptide assemblies of **F1** and **A1**. (a) The sticky ended structure should be more stable for **F1** (8 possible charged pairs, 25 possible hydrogen bonds, 20 amino acid long triple-helical nucleation domain) than for **A1** (only 4 amino acid long triple helical nucleation domain).

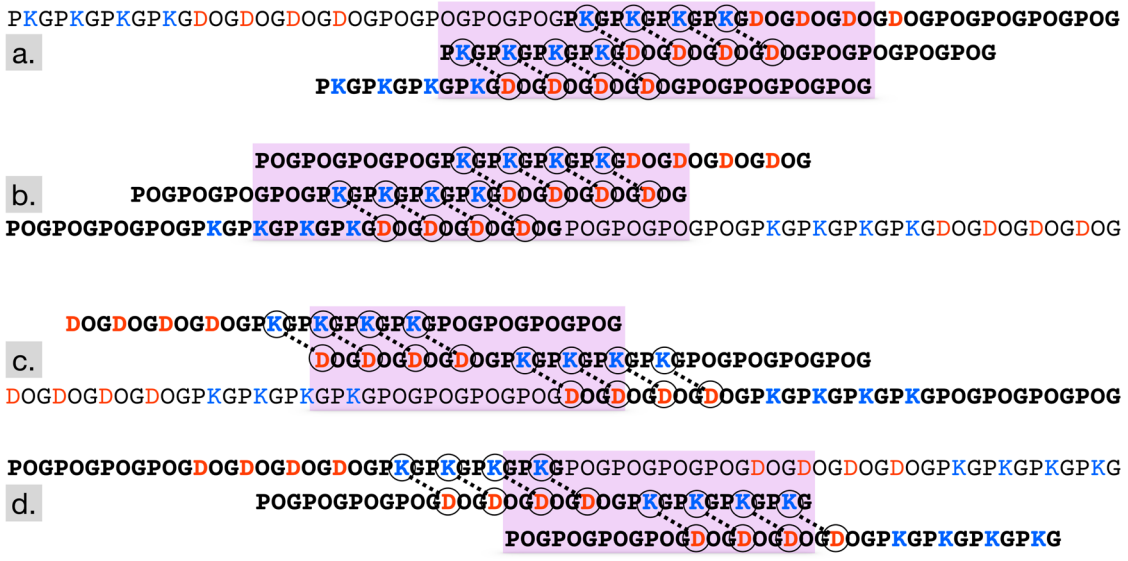


Figure 6. Four peptide nucleation domain of each of the collagen-like peptides studied. For (a) **F1** and (b) **F2**, the triple helical nucleation region (boxed portion) is 28 amino acids long, whereas for (c) **A1** and (d) **A2** the nucleation domain is only 20 amino acids long.

A1, **A2**, the results are not easily explained by the *blunt-ended assembly* hypothesis. None of the four new peptides **F1**, **F2**, **A1**, and **A2** should form nanofibers as all of them have charged domain at only one terminus, not both. We note that peptide **F0** should form nanofibers using either blunt-ended or sticky-ended assembly (see Table 1). However, **F1** and **F2** were observed to assemble into fibers, whereas no fibrillation was seen in **A1** and **A2**. Therefore, we conclude that blunt-ended assembly inadequately accounts for fibrillation observed in these tridomain collagen mimetic peptides containing lysine-aspartate charged pairs. The sticky-ended fibrillation model explains the experimental observations accurately, proving itself to be an effective design methodology for the design of collagen mimetic nanofibers.

X-ray Fiber Diffraction of the Fiber-Forming Collagen Mimetic Peptide F1. A few days after sample preparation, we

dx.doi.org/10.1021/ja504377s | J. Am. Chem. Soc. 2014, 136, 14417−14424

discovered long crystals with high aspect ratio in the aqueous suspension of one of the fiber-forming peptides, **F1**, approaching a centimeter in length. In optical polarized microscopy, the crystals appeared birefringent (Figure 7a and

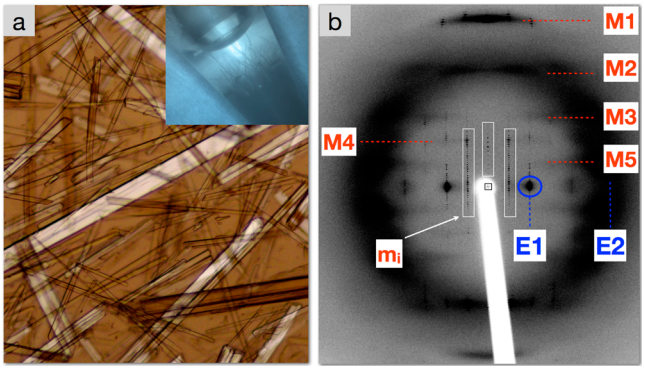


Figure 7. Morphology of the birefringent crystalline needles of **F1** and the fiber diffraction image of a cryogenically cooled crystalline needle. (a) Microcrystals of **F1** observed through cross-polarizers in OPM. Inset shows macroscopic dimension of **F1** crystals grown in an Eppendorf tube. (b) Background-subtracted fiber diffraction image of a cryo-cooled hydrated crystalline needle; the meridional pair of spots M1 (2.8 Å) correspond to periodicity of repeating diffracting unit X–Y–G along the triple helical axis. The other meridional layer lines M2 (4 Å), M3 (6.7 Å), M4 (10 Å), and M5 (20 Å) result from L/n Bragg reflections (n = 5, 3, 2, 1) where L is the triple helical fiber period (20 Å, see Figure S10). The discrete sets of meridional reflections m_i, shown as boxed sets, are due to the axial repeat of the asymmetric unit length (the length of the peptide = 103 Å). The equatorial spot E1 (11 Å at Φ = 90°) corresponds to the lateral spacing between two parallel triple helices in the crystalline bundle.

Supporting Information Figures S6 and S7); thus, there was significant internal order in the needles. Although the quality of the crystals was not sufficient to allow single crystal X-ray diffraction analysis, X-ray diffraction patterns of the cryogenically preserved needles (Figure 7b) collected in a synchrotron facility provided valuable information about the nature of triple-helical packing in the semicrystalline fiber bundles. The diffraction pattern of the hydrated fiber bundles contained both discrete and diffuse spots. The meridionally oriented array of discrete spots (m_i, see Figure 7b) had the periodicity of either 104 or 208 Å, which possibly corresponds to the length of the asymmetric unit along the triple helical axis (the length of the collagen mimetic peptide is approximately 103 Å). As we will see, the pattern of the diffuse spots of the diffraction pattern also hold an interesting set of information.

Alignment and Lateral Packing of Triple Helices in Hydrated Fiber Bundle of F1. The set of meridional bands provide information about the axial periodicity and helicity, and the equatorial reflections are indicative of lateral packing of the protein chains in the triple helix and the packing of triple helices in the fiber. In our system, the triple helices are aligned parallel to the fiber axis, as the crystalline needle was found to be aligned with the meridional 2.8 Å spot, characteristic of triple helices. This arrangement mirrors parallel packing of triple helices in natural collagen, but is dissimilar to the packing of triple helices in collagen mimetic fibers reported by Xu *et al.*[29]

Depending on the Φ-angle, the equatorial pair corresponding to intertriple-helical distance (referred to as **W**) periodically varied from 11 to 15.5 Å (Figure S8). For comparison, in our previously reported dried fiber diffraction of collagen mimetic peptide **F0**, $(PKG)_4(POG)_4(DOG)_4$, the distance between parallel triple helices was reported as 11.5 Å. The degree of alignment in the semicrystalline hydrated fiber bundle of **F1** is much higher than in the dried fiber bundles of **F0**, or dried fiber bundles of **F1** (Figure S9), shown by the fact that the equatorial and meridional spots are well-defined, instead of being smeared out in a ring. This improved directional orientation of the fiber bundles mirrors the highly aligned nature of collagen fibers in tendon, and is reported here for the first time in a synthetic fibrillar collagen system.

We found that the periodic variability of **W** with respect to the crystal orientation angle can be well explained by the arrangement of triple helices in a square lattice (Figure 8). The

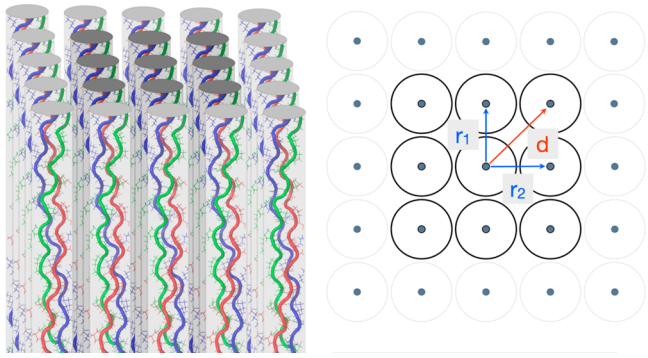


Figure 8. Proposed model for triple helical packing in fibrillar collagen mimetic peptide **F1**. The triple helices pack in a square lattice. The intertriple-helical distance varies periodically with the crystal orientation angle Φ, from 11 to 15.5 Å, reflected in the positions of equatorial diffraction spots E1 and E2. $r_1 = r_2$ = 11 Å, d = 15.5 Å.

arrangement is not precise, as can be evidenced from the diffuse nature of the equatorial spot. This arrangement is different from previously reported pseudohexagonal packing of triple helices in collagen fibers;[18,19] however, the lateral packing parameters are quite similar to those reported recently by Jiang *et al.*[31]

Triple-Helical Period. In contrast to the equatorial spots, the diffuse meridional diffraction spots are fairly wide and correspond to the layer lines resulting from the L/n reflections, where L is the fiber period and n is a positive integer. We observed a pair of strong meridional reflections at 2.8 Å, near-meridional spots of medium intensity at 4.0 Å, and weak near-meridional layer lines at ~6.7 , ~10, and ~20 Å (Figure 7b). All these diffuse and wide meridional reflection are compatible with a fiber period of ~20 Å (see Figure S10). The layer lines correspond to $L/7$, $L/5$, $L/3$, $L/2$ and $L/1$, respectively.

This observation supports the 7/2 triple helical model proposed by Okuyama *et al.*,[40,41] which is supported by recent literature on collagen mimetic peptides.[42−47] We note that none of the collagen mimetic peptides supporting the 7/2 triple helical model self-assemble into nanofibers, and thus their resemblance to collagen is limited to the triple helical level. The idea that collagen has a 20 Å fiber period dates back to the single chain collagen model proposed by Bear and Cohen,[48,49] but fell out of favor in the scientific community after the proposal of the 10/3 triple helical model with a triple helical period of 28.6 Å proposed by Ramachandran and Kartha,[50,51] which was later modified by Rich and Crick.[52,53] Much of the confusion about the collagen molecular structure and triple-

dx.doi.org/10.1021/ja504377s | J. Am. Chem. Soc. 2014, 136, 14417−14424

helical packing arises from the fact that it is very difficult to obtain an unambiguous X-ray diffraction pattern of a very well aligned collagen fiber bundle. Rich and Crick noted, "The final proof of the collagen structure may have to await the production by the polymer chemist of "synthetic collagen" having a simplified amino acid sequence."[53]

CONCLUSION

In this Article, we demonstrate that the nature of the sticky-ended intermediate has dramatic effect on the supramolecular assembly and solubility of collagen-like peptides. We found that larger sticky-ended nucleation domains lead to facile fibrillation. While peptides **F0**−**F2** all self-assemble into collagen-like nanofibers, the organization and morphology of these fibers vary greatly. Peptide **F0** forms a clear hydrogel with good viscoelastic properties, while fibers from peptides **F1** and **F2** phase separate and precipitate despite examining a wide range self-assembly conditions including buffer type and concentration, peptide concentration, and rate of cooling. While we believe we now understand the general rules necessary for fiber formation in tridomain collagen mimetic peptides, controlling the next level of self-assembled hierarchy (the morphology of the formed fibers and the cross-linking, illustrated by hydrogelation or crystallization) is an area that requires further experimentation and insight. Our study shows the hydrated fibers pack into a square lattice of triple helices with a 7/2 helical pitch as a semicrystalline fibrillar bundles of a collagen mimetic peptide. These materials may serve as the basis for synthetic collagen based biomaterials.[54]

METHODS

Peptide Synthesis and Purification. All peptides were synthesized using standard Fmoc chemistry for solid phase peptide synthesis on an Advanced Chemtech Apex 396 multipeptide automated synthesizer, as previously described, at a scale of 0.15 mM on a glycine preloaded Wang resin (loading 0.59 mmol/g). The peptides were cleaved from the resin using a trifluoroacetic acid cocktail containing ethanedithiol, milli-Q water, and triisopropylsilane at a ratio of 37:1:1:1 (v/v), respectively. Both N-terminus and C-terminus were left unprotected. The peptide was triturated with diethyl ether, and recovered by centrifugation. The peptides were purified on a Varian PrepStar220 HPLC using a preparative reverse phase C-18 column. Finally they were dissolved in water, neutralized with 100 mM sodium hydroxide, and dialyzed against distilled water. After dialysis, the solution was frozen overnight and then lyophilized. Once dialyzed, the peptides were analyzed by ESI mass spectrometry on a Bruker microTOF machine. The HPLC chromatograms and mass spectra are given in the Supporting Information.

Sample Preparation. All peptide concentrations were determined by weight. All samples were adjusted to pH 7 with sodium hydroxide prior to the addition of 10 mM sodium phosphate buffer, pH 7, and then annealed for 15 min at 85 °C. Due to the differing time scales of precipitation for each peptide, the samples were incubated for different amounts of time post-annealing before characterization. However, all peptides were allowed to incubate for at least 4 h at room temperature to ensure complete assembly.

Optical Polarized Microscopy. Samples were pipetted up and down to disperse any precipitates and then deposited onto a clean glass microscope slide and covered with a glass coverslip. Samples were viewed on a Nikon Eclipse E400 microscope using a 10×, 20×, or 40× objective lens. Images were taken with a Nikon D7000 camera.

Transmission Electron Microscopy (TEM). TEM Samples were prepared on Quantifoil R1.2/1.3 holey carbon mesh on copper grids. For dry TEM, negative staining techniques were used with phosphotungstic acid (PTA) as the stain. The PTA solution was made at a 2.0% by weight concentration and adjusted to pH ∼ 7 with sodium hydroxide solution. All stains were made biweekly and syringe filtered prior to use to remove any PTA aggregates from the stain. For dry TEM sample preparation, as previously described,[30] the peptide solution was freshly annealed to dissolve all precipitates and then allowed to incubate at room temperature for designated times before TEM sample preparation was begun. After a specific incubation time, the peptide solution was added to the carbon side of a TEM grid, allowed to dry for 1 min, then indirectly blotted with filter paper to remove excess solution. The grid was allowed to dry for 5 min to deposit the peptide onto the grid. After 5 min, the grid was inverted onto an aliquot of PTA solution where it remained for 10 min. Lastly, the grid was removed from the PTA solution and placed carbon side up on filter paper to dry overnight. All dry-TEM imaging was performed on either a JEOL 1230 High Contrast microscope (80 kV) or a JEOL 2010 microscope (200 kV).

Circular Dichroism. All spectra and thermal unfolding studies were performed on a Jasco J-810 spectropolarimeter equipped with a Peltier temperature control system. A quartz Suprasil cell with path length 0.01 cm was used for the spectra measurement (Figure S3), and a quartz cuvette with path length 0.1 cm was used for the melting experiments (Figure 4). Spectra were collected from 190 to 250 nm. The molar residual ellipticity (MRE) $[\theta]$ is calculated from the measured ellipticity using the equation:

$$[\theta] = (\theta m)/(10cln_r)$$

where θ is the ellipticity in mdeg, m is the molecular weight in g/mol, c is the concentration in mg/mL, l is the path length of the cuvette in cm, and n_r is the number of amino acids in the peptide. For the thermal melting experiments, the spectra were monitored at a fixed wavelength, 225 nm, as the temperature was varied from 5 to 55 °C.

X-ray Fiber Diffraction. Crystalline needles of **F1** were pipetted out, and 20% solution of cryo-protectant glycerol in 10 mM phosphate buffer was added to it. One of the large crystals were fished out by using a Hampton CrystalCap HT (SPINE) loop and immediately submerged in liquid nitrogen. The cryogenically cooled crystal samples were sent to X-ray Beamline at Argonne National Laboratory.[55] The sample was exposed to the synchrotron X-ray source (wavelength 0.976 Å) for 3 s per frame, 220 mm away from the detector. The data was collected from $\Phi = 0°$ to $\Phi = 180°$. The software program ADXV was used for data analysis.[56] Background subtraction (radial) was performed (by ADXV) on the diffraction image in Figure 7b to demonstrate the meridional layer lines more clearly.

ASSOCIATED CONTENT

Supporting Information

HPLC traces, ESI mass spectroscopy plots, CD spectra of the peptides, OPM images, and more detailed diffraction data. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

jdh@rice.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Yukimatsu Toh and Prof. Yizhi J. Tao for help with fiber diffraction. This work was funded in part by National Science Foundation (DMR-1206899) and the Robert A. Welch Foundation (Grant No. C1557).

REFERENCES

(1) Yang, W.; Chen, I. H.; Gludovatz, B.; Zimmermann, E. A.; Ritchie, R. O.; Meyers, M. A. *Adv. Mater.* **2013**, *25*, 31−48.

(2) Zimmermann, E. A.; Gludovatz, B.; Schaible, E.; Dave, N. K. N.; Yang, W.; Meyers, M. A.; Ritchie, R. O. *Nat. Commun.* **2013**, DOI: 10.1038/ncomms3634.
(3) Stokkan, K.-A.; Folkow, L.; Dukes, J.; Neveu, M.; Hogg, C.; Siefken, S.; Dakin, S. C.; Jeffery, G. *Proc. R. Soc. B* **2013**, 280 DOI: 10.1098/rspb.2013.2451.
(4) Fratzl, P. *Collagen: Structure and Mechanics*; Springer: New York, 2008.
(5) Shoulders, M. D.; Raines, R. T. *Annu. Rev. Biochem.* **2009**, *78*, 929−958.
(6) Fallas, J. A.; O'Leary, L. E. R.; Hartgerink, J. D. *Chem. Soc. Rev.* **2010**, *39*, 3510−3527.
(7) Persikov, A.; Ramshaw, J.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2000**, *39*, 14960−14967.
(8) Fields, G. B. *Org. Biomol. Chem.* **2010**, *8*, 1237−1258.
(9) Gauba, V.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2007**, *129*, 15034−15041.
(10) Fallas, J. A.; Gauba, V.; Hartgerink, J. D. *J. Biol. Chem.* **2009**, *284*, 26851−26859.
(11) Fallas, J. A.; Hartgerink, J. D. *Nat. Commun.* **2012**, DOI: 10.1038/ncomms2084.
(12) Fallas, J. A.; Lee, M. A.; Jalan, A. A.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2012**, *134*, 1430−1433.
(13) Jalan, A. A.; Hartgerink, J. D. *Biomacromolecules* **2013**, *14*, 179−185.
(14) Jalan, A. A.; Demeler, B.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2013**, *135*, 6014−6017.
(15) Silver, F. H.; Freeman, J. W.; Seehra, G. P. *J. Biomech.* **2003**, *36*, 1529−1553.
(16) Wang, J. H. C. *J. Biomech.* **2006**, *39*, 1563−1582.
(17) Buehler, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12285−12290.
(18) Hulmes, D. J. S.; Miller, A. *Nature* **1979**, *282*, 878−880.
(19) Fraser, R. D. B.; Macrae, T. P.; Miller, A.; Suzuki, E. *J. Mol. Biol.* **1983**, *167*, 497−521.
(20) Orgel, J. P. R. O.; Irving, T. C.; Miller, A.; Wess, T. J. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 9001−9005.
(21) Hulmes, D. J. S.; Wess, T. J.; Prockop, D. J.; Fratzl, P. *Biophys. J.* **1995**, *68*, 1661−1670.
(22) Paramonov, S.; Gauba, V.; Hartgerink, J. *Macromolecules* **2005**, *38*, 7555−7561.
(23) Kar, K.; Amin, P.; Bryan, M. A.; Persikov, A. V.; Mohs, A.; Wang, Y. H.; Brodsky, B. *J. Biol. Chem.* **2006**, *281*, 33283−33290.
(24) Rele, S.; Song, Y. H.; Apkarian, R. P.; Qu, Z.; Conticello, V. P.; Chaikof, E. L. *J. Am. Chem. Soc.* **2007**, *129*, 14780−14787.
(25) Cejas, M.; Kinnney, W.; Chen, C.; Vinter, J.; Almond, H.; Balss, K.; Maryanoff, C.; Schmidt, U.; Breslav, M.; Mahan, A.; Lacy, E.; Maryanoff, B. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 8513−8518.
(26) Yamazaki, C.; Asada, S.; Kitagawa, K.; Koide, T. *Biopolymers* **2008**, *90*, 816−823.
(27) Kar, K.; Ibrar, S.; Nanda, V.; Getz, T. M.; Kunapuli, S. P.; Brodsky, B. *Biochemistry* **2009**, *48*, 7959−7968.
(28) Hsu, W.; Chen, Y.-L.; Horng, J.-C. *Langmuir* **2012**, *28*, 3194−3199.
(29) Xu, F.; Li, J.; Jain, V.; Tu, R. S.; Huang, Q.; Nanda, V. *J. Am. Chem. Soc.* **2012**, *134*, 47−50.
(30) O'Leary, L. E. R.; Fallas, J. A.; Bakota, E. L.; Kang, M. K.; Hartgerink, J. D. *Nat. Chem.* **2011**, *3*, 821−828.
(31) Jiang, T.; Xu, C.; Liu, Y.; Liu, Z.; Wall, J. S.; Zuo, X.; Lian, T.; Salaita, K.; Ni, C.; Pochan, D.; Conticello, V. P. *J. Am. Chem. Soc.* **2014**, *136*, 4300−4308.
(32) Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C. *Nature* **1998**, *394*, 539−544.
(33) Mao, C. D.; Sun, W. Q.; Seeman, N. C. *J. Am. Chem. Soc.* **1999**, *121*, 5437−5443.
(34) LaBean, T. H.; Yan, H.; Kopatsch, J.; Liu, F. R.; Winfree, E.; Reif, J. H.; Seeman, N. C. *J. Am. Chem. Soc.* **2000**, *122*, 1848−1860.
(35) Pandya, M. J.; Spooner, G. M.; Sunde, M.; Thorpe, J. R.; Rodger, A.; Woolfson, D. N. *Biochemistry* **2000**, *39*, 8728−8734.
(36) Papapostolou, D.; Smith, A. M.; Atkins, E. D. T.; Oliver, S. J.; Ryadnov, M. G.; Serpell, L. C.; Woolfson, D. N. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10853−10858.
(37) Woolfson, D. N. *Biopolymers* **2010**, *94*, 118−127.
(38) Jalan, A. A.; Jochim, K. A.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2014**, *136*, 7535−7538.
(39) Russell, L. E.; Fallas, J. A.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2010**, *132*, 3242−4243.
(40) Okuyama, K.; Takayanagi, M.; Ashida, T.; Kakudo, M. *Polymer J.* **1977**, *9*, 341−343.
(41) Okuyama, K. *Connect. Tissue Res.* **2008**, *49*, 299−310.
(42) Kramer, R. Z.; Venugopal, M. G.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *J. Mol. Biol.* **2000**, *301*, 1191−1205.
(43) Kramer, R. Z.; Bella, J.; Brodsky, B.; Berman, H. M. *J. Mol. Biol.* **2001**, *311*, 131−147.
(44) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2002**, *11*, 262−270.
(45) Kawahara, K.; Nishi, Y.; Nakamura, S.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *Biochemistry* **2005**, *44*, 15812−15822.
(46) Schumacher, M.; Mizuno, K.; Bachinger, H. P. *J. Biol. Chem.* **2005**, *280*, 20397−20403.
(47) Schumacher, M. A.; Mizuno, K.; Bachinger, H. P. *J. Biol. Chem.* **2006**, *281*, 27566−27574.
(48) Bear, R. S. *Adv. Protein Chem.* **1952**, *7*, 69−160.
(49) Cohen, C.; Bear, R. S. *J. Am. Chem. Soc.* **1953**, *75*, 2783−2784.
(50) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593−595.
(51) Ramachandran, G. N. *Nature* **1956**, *177*, 710−711.
(52) Rich, A.; Crick, F. H. *Nature* **1955**, *176*, 915−916.
(53) Rich, A.; Crick, F. H. *J. Mol. Biol.* **1961**, *3*, 483−506.
(54) Kumar, V. A.; Taylor, N. L.; Jalan, A. A.; Hwang, L. K.; Wang, B. K.; Hartgerink, J. D. *Biomacromolecules* **2014**, *15*, 1484−1490.
(55) http://ls-cat.org/; Northwestern University: Argonne, IL.
(56) Arvai, A. S. *ADXV - A program to display X-ray diffraction images*; La Jolla, CA; http://www.scripps.edu/tainer/arvai/adxv.html.