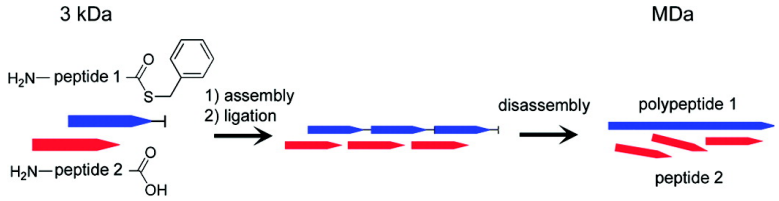
Article

Self-Assembled Templates for Polypeptide Synthesis

Maxim G. Ryadnov, and Derek N. Woolfson

J. Am. Chem. Soc., **2007**, 129 (45), 14074-14081• DOI: 10.1021/ja072960s • Publication Date (Web): 20 October 2007



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 5 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 10/20/2007

Self-Assembled Templates for Polypeptide Synthesis

Maxim G. Ryadnov*,‡ and Derek N. Woolfson*,‡,§

Contribution from the School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK, and Department of Biochemistry, School of Medical Sciences, University of Bristol, University Walk, Bristol, BS8 1TD, UK

Received May 4, 2007; E-mail: Max.Ryadnov@bristol.ac.uk; D.N.Woolfson@bristol.ac.uk

Abstract: The chemical synthesis of polypeptide chains >50 amino acids with prescribed sequences is challenging. In one approach, native chemical ligation (NCL), short, unprotected peptides are connected through peptide bonds to render proteins in water. Here we combine chemical ligation with peptide self-assembly to deliver extremely long polypeptide chains with stipulated, repeated sequences. We use a self-assembling fiber (SAF) system to form structures tens of micrometers long. In these assemblies, tens of thousands of peptides align with their N- and C-termini abutting. This arrangement facilitates chemical ligation without the usual requirement for a catalytic cysteine residue at the reactive N-terminus. We introduced peptides with C-terminal thioester moieties into the SAFs. Subsequent ligation and disassembly of the noncovalent components produced extended chains ≥10 μm long and estimated at ≥3 MDa in mass. These extremely long molecules were characterized by a combination of biophysical, hydrodynamic, and microscopic measurements.

Introduction

The chemical synthesis of proteins has been a goal for chemists and biochemists for about a century.[1−5] The introduction of solid-phase peptide synthesis (SPPS)[6,7] by Merrifield in the 1960s revolutionized the field and nowadays offers an accessible and affordable route to the production of peptides and small proteins. However, the synthesis of polypeptide chains of >50 amino acids in length by SPPS remains challenging.[4] The preparation of such polypeptides has already opened and will likely continue to open up new lines of research in chemical and synthetic biology, polymer physics and biomaterials research.

A rapidly growing subfield in protein synthesis, namely chemoselective ligation, at least partly addresses this issue, particularly in regard to the preparation of functional proteins.[4,8] The approach relies on the coupling, or ligation, of unprotected polypeptide fragments. This is a semisynthetic method performed in water. In this way, the influence of any folded structure in the polypeptide chains can be impaired by employing denaturing conditions or used to help drive couplings in a so-called conformationally or folding-assisted manner.[8−13] For instance, early examples of the latter, and of direct relevance to the study reported herein, include work from the Ghadiri and Chmielewski groups in which a leucine-zipper peptide templates the co-assembly and subsequent ligation of two half-peptides.[10,13] In both cases, the N-terminal half-peptide has a C-terminal thioester, and the N-terminus of the C-terminal peptide is cysteine. These so-called peptide ligases or self-replicating systems render short peptides that match the length of the covalent template.

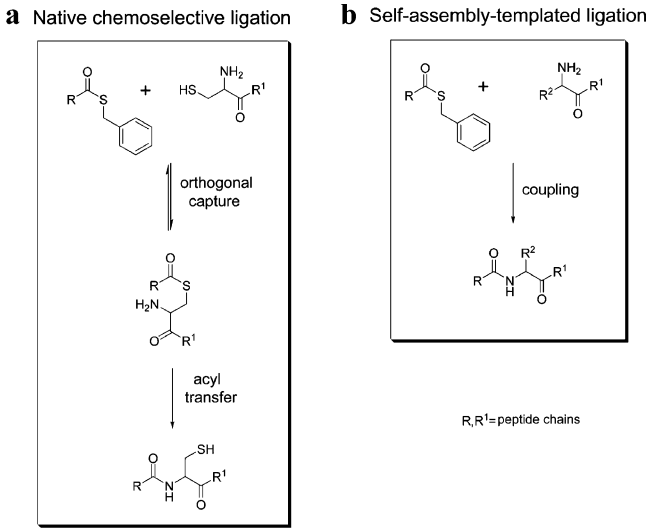
Different chemistries can be employed for ligating peptides and proteins.[14] However, those that produce standard peptide bonds are of interest in the synthesis of native proteins. In this vein, the laboratories of Kent[15] and Tam[16] have developed native chemoselective ligation (NCL) (Scheme 1a), which utilizes the concept of entropic activation originally proposed by Kemp.[17] This is a spontaneous two-step orthogonal coupling method: peptides are first coupled via a nonamide capture reaction, usually realized through the reaction of a C-terminal thioester on the first fragment, and the side-chain thiol of an N-terminal cysteine residue on the second, followed by an intramolecular S-to-N acyl transfer to give the native peptide bond (Scheme 1a). Thus, NCL requires a cysteine residue either present or engineered at what becomes the residue following the new peptide bond (Scheme 1a). To broaden the amino-acid type at

‡ School of Chemistry.
§ Department of Biochemistry.

(1) Fischer, E. *J. Chem. Soc. Trans.* **1907**, *91*, 1749−1765.
(2) Albericio, F. *Curr. Opin. Chem. Biol.* **2004**, *8*, 211−221.
(3) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91−118.
(4) Wilken, J.; Kent, S. B. *Curr. Opin. Biotechnol.* **1998**, *9*, 412−426.
(5) Du Vigneaud, V.; Ressler, C.; Swan, J. M.; Roberts, C. W.; Katsoyannis, P. G.; Gordon, S. *J. Am. Chem. Soc.* **1953**, *75*, 4879−4880.
(6) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149−2154.
(7) Gutte, B.; Merrifield, R. B. *J. Am. Chem. Soc.* **1969**, *91*, 501−502.
(8) Dawson, P. E.; Kent, S. B. H. *Ann, Rev. Biochem.* **2000**, *69*, 923−960.
(9) Beligere, G. S.; Dawson, P. E. *J. Am. Chem. Soc.* **1999**, *121*, 6332−6333.
(10) Severin, K.; Lee, D. H.; Kennan, A. J.; Ghadiri, M. R. *Nature* **1997**, *389*, 706−709.
(11) Camarero, J. A.; Pavel, J.; Muir, T. W. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 347−349.
(12) Proudfoot, A.; Rose, K.; Wallace, C. *J. Biol. Chem.* **1989**, *264*, 8764−8770.
(13) Yao, S.; Ghosh, I.; Zutshi, R.; Chmielewski, J. *J. Am. Chem. Soc.* **1997**, *119*, 10559−10560.
(14) Lemieux, G. A.; Bertozzi, C. R. *Trends Biotechnol.* **1998**, *16*, 506−513.
(15) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776−779.
(16) Tam, J.; Lu, Y.; Liu, C.; Shao, J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12485−12489.
(17) Kemp, D. S. *Biopolymers* **1981**, *20*, 1793−1804.

10.1021/ja072960s CCC: $37.00

Scheme 1. Chemoselective Ligation Mechanisms. (a) Native Chemical Ligation Requiring a Catalytic Cysteine. (b) Proximity-Driven Chemical Ligation Using Self-Assembled Templates without the Necessity for Cysteine



this position, however, alternatives have been developed that employ auxiliaries that can be appended to the N-terminus of the second fragment, facilitate the reaction, and then be removed.[18,19] Conformationally assisted ligations have also been reported. For example, using chymotrypsin inhibitor 2, Beligere and Dawson demonstrate that stable protein folds can function as scaffolds directing the ligation of loops.[9] Structurally amorphous loops are preferred, as these do not impose secondary structure restrictions. However, the rate of such reactions without cysteine is quite low, as the orthogonal capture step is eliminated,[9] which makes ligation highly entropy driven and proximity dependent.

Here we show that cysteine-free chemoselective ligation can be effected by peptide self-assembly to prepare structurally defined polypeptides of prescribed sequences and of MDa in size. We exemplify this using a Self-Assembling Fiber (SAF) system,[20] which comprises two leucine-zipper sequences of *de novo* design, SAF-p1 and SAF-p2a (Table 1 and Figure 1b).[21] These are designed to co-assemble in an offset manner to give a sticky ended dimer.[20] Further end-to-end association of the dimers is programmed to promote longitudinal elongation and form an extended dimeric coiled coil; that is, two noncovalently linked chains of α-helical peptides that supercoil around one another to form a rope, or, as described herein, protofibrils. The peptides within each chain are aligned in a head-to-tail fashion, with the adjacent C- and N-termini almost certainly linked noncovalently through a CO_2^--to-NH_3^+ salt bridge[22] (Figure 1). The SAF peptides are short, both 28-residues long, and amenable to standard Fmoc-based SPPS. The assembly of the peptides has been characterized both in solution, primarily using circular dichroism (CD) spectroscopy, which confirms the designed underlying α-helical structure, and by electron microscopy (EM) as follows. The expected dimensions of the target structure are 2 nm across and microns long. In practice, however, the protofibrils associate laterally to form thickened fibers 70 ± 20 nm wide and tens of microns long (Figure 2a).[21,22] This ripening leads to stabilized fibers that are very straight (persistence lengths ≥10 μm) and which exhibit nanoscale internal and external order. These structural features suggest tight and conserved packing of the protofibrils and the peptides that comprise them.[22] We have shown that the peptides can be engineered as SAF-compatible *specials* to introduce additional features into the fiber such has branches, kinks, breaks, cross-links, and functional appendages.[23−25]

Here, we report the use of SAFs to template the end-to-end assembly and direct the subsequent ligation of peptides with

Table 1. Peptide Sequences Used in the Study

Peptide	Sequence[a]
SAF-p1	KIAALKQKIASLKQEIDALEYENDALEQ
SAF-p1E	KIAALKQKIASLKQEIDALEYENDALEQ-COSBn
SAF-p2	KIRRLKQKNARLKQEIAALEYEIAALEQ
SAF-p2E	KIRRLKQKNARLKQEIAALEYEIAALEQ-COSBn
STeP	KIAALKQKIAALKQEIAALEYEIAALEQ
STePE	KIAALKQKIAALKQEIAALEYEIAALEQ-COSBn

[a] COSBn = thiobenzyl ester.

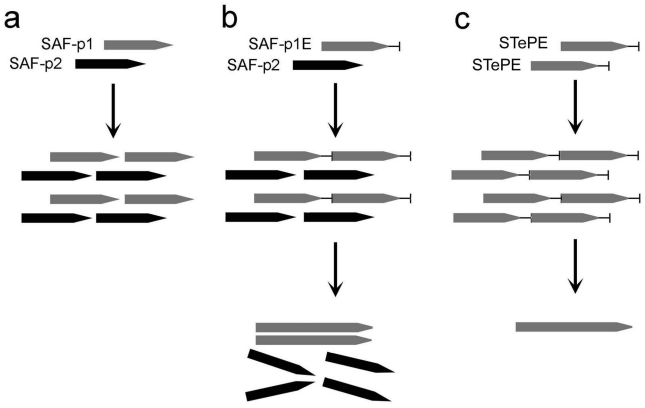


Figure 1. Sticky-ended peptide self-assembly and the possibilities for chemical ligation. (a) SAF peptides combine to give a sticky ended dimer that propagate longitudinally into protofibrils and laterally to form fibers. (b) SAF-p1 thioester (SAF-p1E) templated by its standard partner (SAF-p2) reacts with another SAF-p1E in a head-to-tail fashion to produce SAFE polypeptides chains. (c) In a new system, STeP, the sticky ended dimmer, is formed by a single peptide sequence, which, as a thioester (STePE), self-ligates into a polypeptide chain. Peptides are shown with arrows pointing in the N-to-C direction. For clarity, only two protofibrils are shown; in reality, many protofibrils bundle to form the matured fibers.

(18) Offer, J.; Boddy, C. N. C.; Dawson, P. E. *J. Am. Chem. Soc.* **2002**, *124*, 4642−4646.
(19) Low, D. W.; Hill, M. G.; Carrasco, M. R.; Kent, S. B. H.; Botti, P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 6554−6559.
(20) Pandya, M. J.; Spooner, G. M.; Sunde, M.; Thorpe, J. R.; Rodger, A.; Woolfson, D. N. *Biochemistry* **2000**, *39*, 8728−8734.
(21) Smith, A. M.; Banwell, E. F.; Edwards, W. R.; Pandya, M. J.; Woolfson, D. N. *Adv. Funct. Mater.* **2006**, *16*, 1022−1030.
(22) 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.
(23) Ryadnov, M. G.; Woolfson, D. N. *J. Am. Chem. Soc.* **2005**, *127*, 12407−12415.
(24) Ryadnov, M. G.; Woolfson, D. N. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 3021−3023.
(25) Ryadnov, M. G.; Woolfson, D. N. *J. Am. Chem. Soc.* **2004**, *126*, 7454−7455.

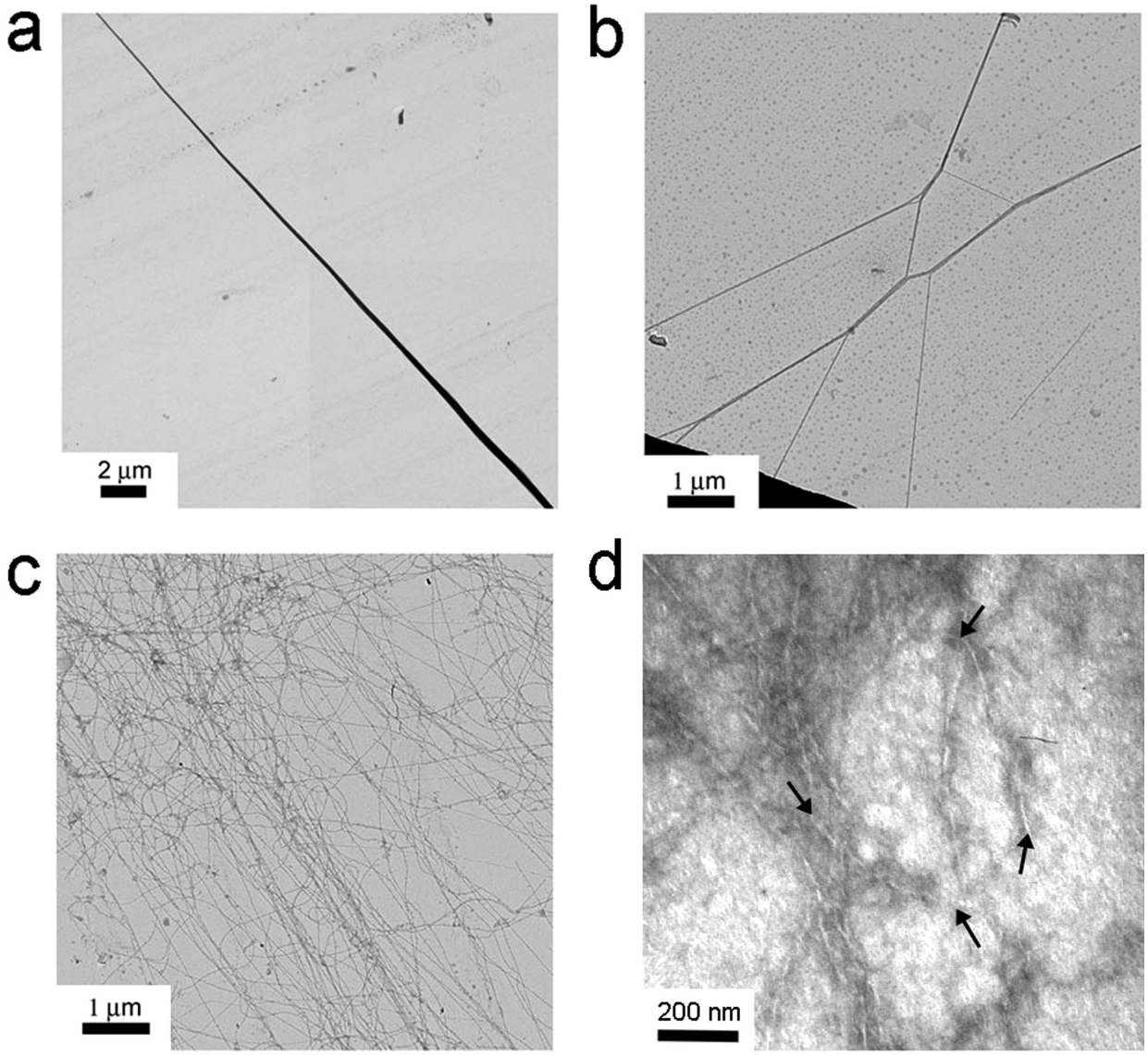


Figure 2. Visualization of the unligated and ligated self-assembled peptide fibers. Electron micrographs of standard, nonligated SAFs (a). SAFE fibrils prepared and maintained at pH 7.4 (b), and titrated to pH 8.5 (c). SAFE fibrils assembled by doping the thioester-containing SAF-p2E into mixtures of the standard SAF peptides (in the ratio 0.01:1:1) (d). In part d, the ends of some fibers are visible as highlighted with arrows.

C-terminal thioesters (Scheme 1b). In this system the templating of the reactive peptides fixes them conformationally and brings the reacting termini into close proximity (Figure 1). The effect is similar to that seen in topotactic reactions performed in crystalline or solid substances where the key factor is minimized molecular dynamics.[26] We propose that the paracrystalline nature of the SAFs, which is reflected in high nanoscale ordering both on the surfaces of and within the fibers,[22] facilitates the ligation of the C-terminal thioesters and abutting N-termini to effect a topotactic polymerization. As we demonstrate through a combination of circular dichroism (CD) spectroscopy, electron microscopy (EM), and mass spectrometry (MS), this can be used to synthesize polypeptides in the kDa to MDa mass range. These fold as stabilized α-helices that extend from hundreds of nanometers to tens of microns in length. To our knowledge, this is the first such synthesis of linear polypeptides comprising proteinogenic amino acids in a defined sequence.

(26) Kaupp, G. *Curr. Opin. Solid State Mater. Sci.* **2002**, *6*, 131−138.

Results

Design Rationale. The aforementioned features of the SAF system, namely the head-to-tail alignment of peptides and the proximity of their termini, the high degree of order in the fibers, and the ability to incorporate nonstandard peptides, suggested to us that the SAF system was perfectly set up for cysteine-free chemical ligation. To test this, we synthesized a series of SAF peptides with C-terminal thioesters for chemoselective ligation, Table 1. Each peptide thioester is a self-ligator, as it can potentially couple to both of its abutting peptides, which are copies of the same peptide. As we demonstrate, the proximity of the termini and the order in the fibers allow intermolecular ligation without the prior orthogonal capture step and thus without the need for an N-terminal cysteine.

The Individual Thioester Peptides Do Not Assemble. As controls to test for any unintended nonconformationally assisted ligation, the individual benzyl thioesters of SAF-p1 and SAF-p2, i.e., SAF-p1E and SAF-p2E, respectively, were first incubated separately under our standard conditions for peptide

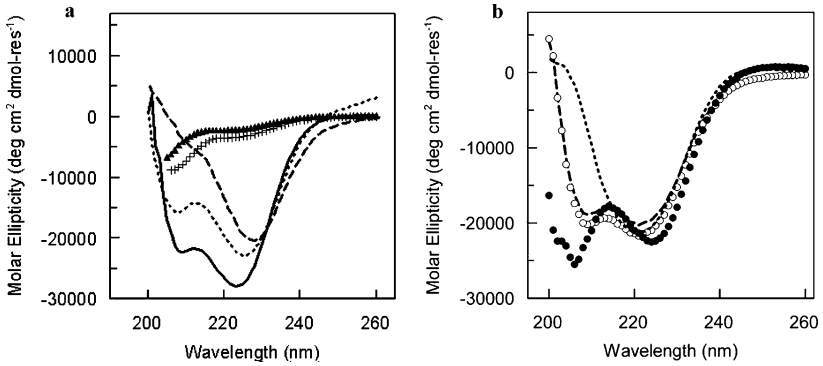


Figure 3. Secondary structure of the SAF-based fibrillar assemblies. (a) CD spectra for SAF-p1E (black triangles), SAF-p2E (crosses), SAF-p1/SAF-p2 (solid line), SAF-p1E/SAF-p2 (dashed line), and SAF-p1/SAF-p2E (dotted line) at pH 7.4. (b) SAF-p1E/SAF-p2 (white circles) and SAF-p1/SAF-p2E (black circles) at pH 8.5, postmelt and recooled SAF-p1E/SAF-p2 (dashed line), and SAF-p1/SAF-p2E (dotted line) at pH 7.4.

fiber assembly (100 μM peptide, 20 °C). This was done at both pH 7.4 and pH 8.5, the former being standard, while the latter was used to test further for unintended ligation, as increased pH is known to accelerate normal NCL reactions.[16] Under both conditions, neither peptide was folded as judged by circular dichroism (CD) spectroscopy (Figure 3a). Moreover, there was no evidence of ligation products by mass spectrometry. These were important controls, though the results are not surprising as (1) the standard peptide concentration is ten times lower than that used typically in normal NCL,[9,17,18,27] and (2) neither thioester SAF peptide has an N-terminal cysteine normally required to accelerate ligation. Nevertheless, the controls were repeated at a concentration more typically used in NCL, i.e., 1 mM in each peptide.[9,17,18,27] These gave similar results to the experiments performed at the lower concentrations, confirming that neither assembly nor ligation of the cysteine-free thioester peptides occurred to appreciable degrees.[8−10]

Preparing, Visualizing, and Characterizing Fibers and Ligated Polypeptides. To begin testing templated assembly and ligation, SAF-p1E was mixed with its complementary standard partner, SAF-p2, in equimolar ratio. The peptides readily assembled under the aforementioned conditions to give micron-length fibers (SAFEs) as revealed by transmission electron microscopy (TEM) (Figure 2b). Morphologically, the SAFEs were different from the standard SAFs: they were thinner and tended to wrap one around another, which, in some cases, led to the formation of loose networks (Figure 2b). This is perhaps not surprising, as the SAF system presents a potentially permissive supramolecular background in which slight changes in the component peptides can lead to changes in the fibers.[23] In the case of the thioester peptides, such changes have the chance of becoming 'locked-in' through ligation.

Nonetheless, as judged by CD spectroscopy, SAFEs retained the characteristic underlying α-helical structure of the SAF system.[20] Spectra of the standard SAFs had minima near 208 and 222 nm typical of the α-helical conformation. However, the 222 nm band was slightly red-shifted, and that at 208 nm slightly reduced in intensity (Figure 3a).[20] These features indicate light scattering from nano-to-mesoscale particles/fibers.[20] Similarly, CD spectra for SAFE showed the 208 and 222 nm bands, but with further red shifts of the latter (Figure 3a). As far as comparisons of such spectra can be made, the spectra for SAFE were intermediate between those for the standard, linear SAFs and recently described interconnected SAFs.[23] The latter were characterized by a complete disappearance of the 208 nm band, consistent with very large assemblies, which were indeed observed by TEM.[23]

Furthermore, at least partial ligation of SAFE under standard SAF conditions is supported by comparing the aforementioned CD spectra with those recorded after thermal denaturation and recooling of SAFE (Figure 3b). For example, postmelt fibrils from the SAF-p1E/SAF-p2 mixture remained α-helical, whereas the corresponding CD spectra for the SAF-p1/SAF-p2E mixture indicated a switch to β-structure (Figure 3b). This suggests that at least some polymerization of SAF-p1E and SAF-p2E occurred, respectively, and that the resulting polypeptides have subtly different properties at least with respect to thermal denaturation.

Concomitant Ligation and Fiber Disassembly. If it were not for the difficulties with thermal denaturation of peptides and proteins alluded to above, heating would be one way to drive ligation. NCL is usually performed at pH 5−7.5.[15,16,27] However, Tam et al. showed that thioester-mediated ligations accelerate with increasing pH.[16] Thus, to achieve complete ligation at room temperature, preparations of SAFE were preassembled for 24 h and then titrated to pH 8.5. From the CD spectra, the resulting structures were still α-helical; indeed they were more typically α-helical than the pH 7.4 preparation of either the standard (SAF) or ligated (SAFE) fibers (Figure 3b); that is, the distortions from light scattering were reduced considerably. Interestingly, and consistent with this, the pH increase led to a dramatic change in the morphology of fibers as observed by TEM (Figure 2). Specifically, the structures treated at pH 8.5 were approximately ten and three times thinner than SAF and SAFE, respectively. The resulting *fibrils* also appeared to be more flexible and entangled than the standard SAFs, which are rigid and linear. This made estimates of fibril lengths uncertain. Further titrations up to pH 10 revealed no further changes in morphology. Moreover, similar results were obtained using a single sharp pH increase from 7.4 to 10, "pH shock", suggesting that pH 8.5 is the cutoff for complete ligation. Consistent with previous work from this laboratory, no fibers were observed at pH ≥ 8 for standard SAF mixtures.[21]

Summarizing the above, the CD and TEM data combined strongly support the ligation of thioester monomers into extended polypeptide chains as designed. We attribute the thinning of the fibers after the pH changes to the disassembly of the noncovalently linked, standard SAF components of

(27) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. *J. Am. Chem. Soc.* **1997**, *119*, 4325−4329.

SAFEs. This raises the question, why the resulting fibers are not one helix thick, i.e., ~ 1 nm? (While uniform, the fibril widths observed, are still thicker than this at 7 ± 2 nm.) Moreover, why do these bundles remain α-helical? Clearly, it is thermodynamically more favorable for the ligated chains to form extended and bundled α-helices than dissociate or unfold completely. Presumably, this persistent α-helicity and the bundling are due to (1), the extended canonical α-helical coiled-coil pattern of hydrophobic and polar residues along the polypeptides, which favors dimerization and would give 2 nm structures; plus (2) subsequent, though possibly different, inter-superhelical electrostatic interactions that cement and thicken the standard SAF fibers.[21]

Another potential cause of bundling could be interchain cross-linking, which might result from the reaction of side-chain amino groups (of the lysine residues in the sequences) with the thioesters. Given the considerable order in the SAFs and that consecutive C- and N-termini must be tightly abutted,[22] we think that this is highly unlikely. Nonetheless, we attempted a control experiment with N-acetylated SAF-p1E (acSAF-p1E) to test the regioselectivity of the reaction for α-amino groups. However, mixing acSAF-p1E with its uncapped standard partner, SAF-p2, did not render fibers at equimolar ratios. (This is completely consistent with our recent work on the interrogation of the matured standard fibers.)[22] In the control experiment, some amorphous and much smaller aggregates were seen with lower ratios of acSAF-p1E in mixtures with SAF-p1 and SAF-p2, but these dissolved at pH 8 (Figure S5). Furthermore, no ligated or cross-linked peptides were detected at either ratio by mass spectrometry.

Similarly, and also as control experiments, our attempts to effect carbodiimide-mediated ligation of both SAF and SAFE were unsuccessful. Carbodiimides, and in particular the water-soluble 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide (EDC), can be used as cross-linking reagents to couple carboxyl groups and primary amines. Because C- and N-termini are tightly packed in the templates, one might consider that they could be coupled simply using carbodiimides, that is, provided the templates remain intact under the experimental conditions used. In our experiments, however, the addition of EDC caused the disassembly of the templates, with no signs of ligated peptides by mass spectrometry.

What Are the Lengths and Masses of the Polypeptide Chains Produced, and Can These Be Tailored? Given the micron lengths of the SAFE fibrils formed after complete ligation, it is tempting to speculate that we have generated synthetic polypeptides of repeated sequence ≥30 000 residues long and ≥3 MDa in mass. This would be larger than any preceding synthetic polypeptides of defined sequence, or of any natural protein. This approximate calculation of the lower limit runs as follows: the individual SAF peptides are ~4 nm in length when configured as an α-helix; therefore, even for fibers and fibrils as short as 4 μm in length (a conservative length from the EM images of Figure 2), ~1000 peptides would be aligned in these structures; the SAF peptides are 28 residues long and ~ 3kDa in mass; thus, after ligation, the polypeptides could be ~30 000 residues and ~3 MDa. This analysis is subject to a number of caveats, particularly that the fibrils formed may be aggregates of shorter polypeptides.[28] Experimental characterization of such long polypeptides is difficult. Nevertheless, we attempted to explore this as follows.

In size-exclusion chromatography (SEC) the ligated peptides eluted in the void volume. Similarly, matrix assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) mass spectrometry failed to detect any traces of the ligated SAFE products. Dynamic light scattering (DLS), which is often used to obtain an average size of macromolecules by relating their size to their hydrodynamic radii, revealed an average size of 15 μm for the polypeptides eluted in the void volume of the SEC. Although this result is consistent with the TEM data, the DLS was not consistent from sample to sample. As mentioned above, the propensity for the fibrils to wrap around one another complicates the SEC and DLS further.[28] For these reasons, we took another approach to tackle the problem.

We set out to decrease the size of the ligated polypeptides by employing mixtures of the thioester and standard SAF peptides at different ratios. In these experiments, the standard peptides effectively act as chain terminators in the chemical-ligation polymerization. For example, at ratios as little as 0.01:1:1 (thioester:standard:standard partner) shorter SAFEs were produced (Figure 2d and Figure S3). In these images the ends of fibrils are better defined. Some of the shortened fibers observed were ~200 nm, which equates to polypeptides potentially of ~1500 amino acids and molecular masses of ~150 kDa, and the lengths of most did not exceed 0.5 μm, corresponding to SAFEs possibly up to ~400 kDa in mass. At this low ratio of reactive to nonreactive peptides, a distribution of ligated products from dimer up is expected, with dimers being the most abundant. Many of these would not be visible in TEM. Consistent with this, however, MALDI-TOF mass spectrometry gave a broad distribution of *m*/*z* values, dominated by peaks at the lower end of the expected spectrum (Figure 4 and Figure S1): indeed, the main peak at 6630 corresponds to a SAF-p2E dimer, that at 9957 a trimer, and so on up to peaks that are consistent with an 14- and 15-mers; though peaks above this and up to ~ 60 kDa are evident (Figures 4 and S1).

A Self-Templating Peptide (STeP) System. As noted above, both the SAFs and SAFEs, which are two-peptide systems, bundle to give thickened fibers. This is driven and stabilized by inter-protofibril interactions, which can be rationally engineered to some extent.[21] In an attempt to reduce such potential protofibril−protofibril interactions, a self-assembling single-peptide (STeP) was designed with a largely alanine exterior (Table 1). Others have shown that the sticky end assembly can be engineered into single-peptide systems, and that these adopt stable helical fibers that can be thinner, though also morphologically less well-defined.[29−31] STeP was expected to self-associate to form a sticky ended dimer and thence self-templated fibers (Figure 1c). For an unmodified peptide with a normal C-terminus, CD spectroscopy (Figure 5a) and TEM (Figure 6) confirmed that STeP formed α-helical fibers, with many of the characteristics of the dual-peptide SAFs: notably, the peptide

(28) Paramonov, S. E.; Gauba, V.; Hartgerink, J. D. *Macromol.* **2005**, *38*, 7555−7561.
(29) Zimenkov, Y.; Conticello, V. P.; Guo, L.; Thiyagarajan, P. *Tetrahedron* **2004**, *60*, 7237−7246.
(30) Zimenkov, Y.; Dublin, S. N.; Ni, R.; Tu, R. S.; Breedveld, V.; Apkarian, R. P.; Conticello, V. P. *J. Am. Chem. Soc.* **2006**, *128*, 6770−6771.
(31) Wagner, D. E.; Phillips, C. L.; Ali, W. M.; Nybakken, G. E.; Crawford, E. D.; Schwab, A. D.; Smith, W. F.; Fairman, R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12656−12661.

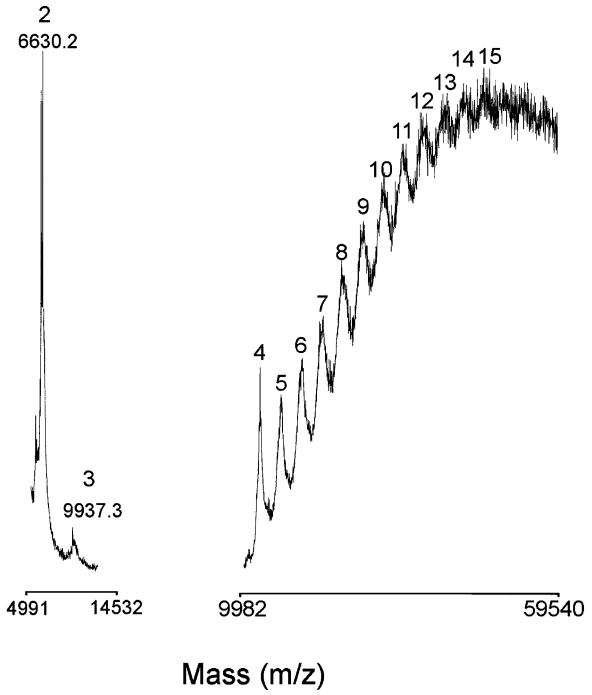


Figure 4. Doping a thioester peptide into a standard SAF background. MALDI-TOF mass spectrum for a variant of SAFE assembled from SAF-p2E added in a 0.01:1:1 ratio with the standard SAF peptides. The resulting series of dimer (expected mass, 6630 ((2 × 3324) − 18), trimer (expected mass, 9936 ((3 × 3324) − 36), and higher oligiomers of SAF-p2. For clarity, only *m*/*z* values for dimer and trimer are shown.

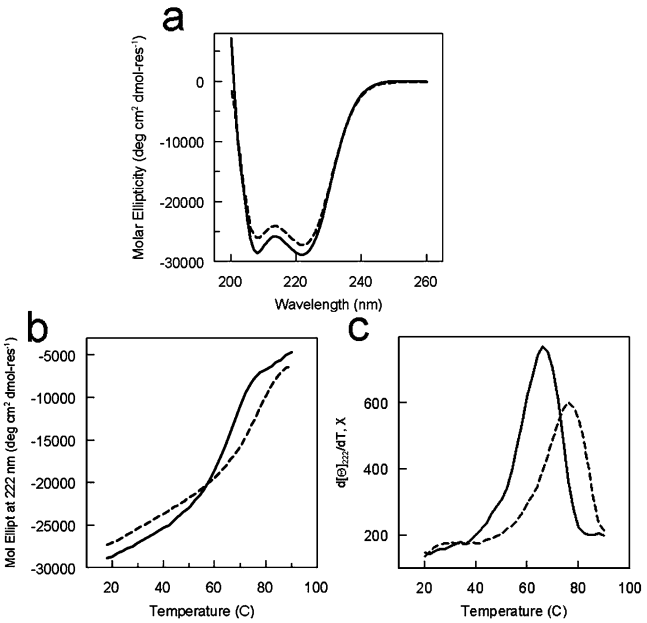


Figure 5. Secondary structure and stability of the STeP and STePE fibrils. (a) CD spectra for STeP (solid line) and STePE (dashed line). (b) Thermal unfolding curves for the STeP recorded by following the CD signal at 222 nm as a function of temperature. (c) First derivatives of the curves from b used to determine the midpoints of thermal unfolding. The keys for b and c are the same as that for a.

assembled to α-helical, extended fibrous structures, which, though thinner and less well-defined than the SAFs, bundled to form a variety of associated fibers, tapes, and splintered structures (Figures 6a and S4).

In contrast, the thioester version of STeP (STePE) produced shorter and thinner fibrils at neutral pH (Figure 6b). No assemblies were found under basic conditions for both STeP and STePE by TEM, whereas complete disappearance of SAFE could only be achieved with high salt (1 M KF). Overall, this supports impaired interfibril interactions and, as a result, a weaker background of STeP fibers compared to that of the SAFs.

An important question in the SAFE and STePE systems is whether covalently linking successive peptides stabilizes the fibers? This is more easily addressed in the STeP system because (1) the aforementioned structural switching seen in one of the SAFE systems upon thermal denaturation, and (2) all of the monomers should be linked in STePE. To test this, the CD signals for preformed STeP and STePE fibrils were monitored as a function of temperature. The curves obtained were sigmoidal, indicating cooperative unfolding transitions as expected for assembled and thermodynamically defined structures, Figure 5b; the transitions were reversible. The midpoints of the curves, taken from their first derivatives, Figure 5c, were 66 °C and 76 °C, respectively. Thus, ligation of the peptides stabilizes the assemblies even though the fibrils observed are shorter and thinner than those observed for the nonligated fibers.

In continuation of the mass spectrometry experiments described for SAFE in the STePE system, we first attempted to observe either monomers or oligomers from a preassembled STePE. This gave no discernible peaks, indicating that the majority of the monomers had been incorporated into fibers and then ligated to form very large polymers. In a second experiment, preassembled STePE was treated with the protease chymotrypsin. Each STeP monomer contains a single tyrosine residue (Figure 7, Table 1). Tyrosines are substrates for chymotrypsin, which cleaves the peptide bond after the aromatic residues. Thus, chymotrypsin cleavage of ligated STePE should produce a single, permuted peptide, STeP*, with a mass of 3068 Da; alternatively, unligated STePs would give at least two peaks at 2313 and 773 Da. MALDI-TOF analysis of chymotrypsin-treated STePE yielded a strong peak at 3091 Da corresponding to the $[STeP^* + Na]^+$ molecular ion, Figure 7b, and neither peak expected if digestion of the unligated peptide had occurred was observed, Figure S2. Peaks at lower *m*/*z* values were found, but these could be rationalized by chymotrypsin cleaving after leucine residues in STePE, which is a known side reaction of the protease, Figure S2. Thus, these mass spectrometry data are fully consistent with ligation of STeP to give STepE polymers.

Discussion

In summary, we have shown that peptide folding and self-assembly can be combined with chemoselective ligation to synthesize extremely long polypeptide chains of prescribed, albeit repeated, sequence. The approach described focuses on achieving high effective local concentrations of reactants by restricting their spatial orientation. In this respect, the method resembles solid-state topochemical reactions, which primarily benefit from minimized molecular dynamics,[26] and is analogous to proximity-driven peptide coupling performed by ribosomes.[32] In our system, the organization required is achieved through the rational design of a self-assembling peptide template;

(32) Rodnina, M. V.; Beringer, M.; Wintermeyer, W. *Trends Biochem. Sci.* **2007**, *32*, 20−26.

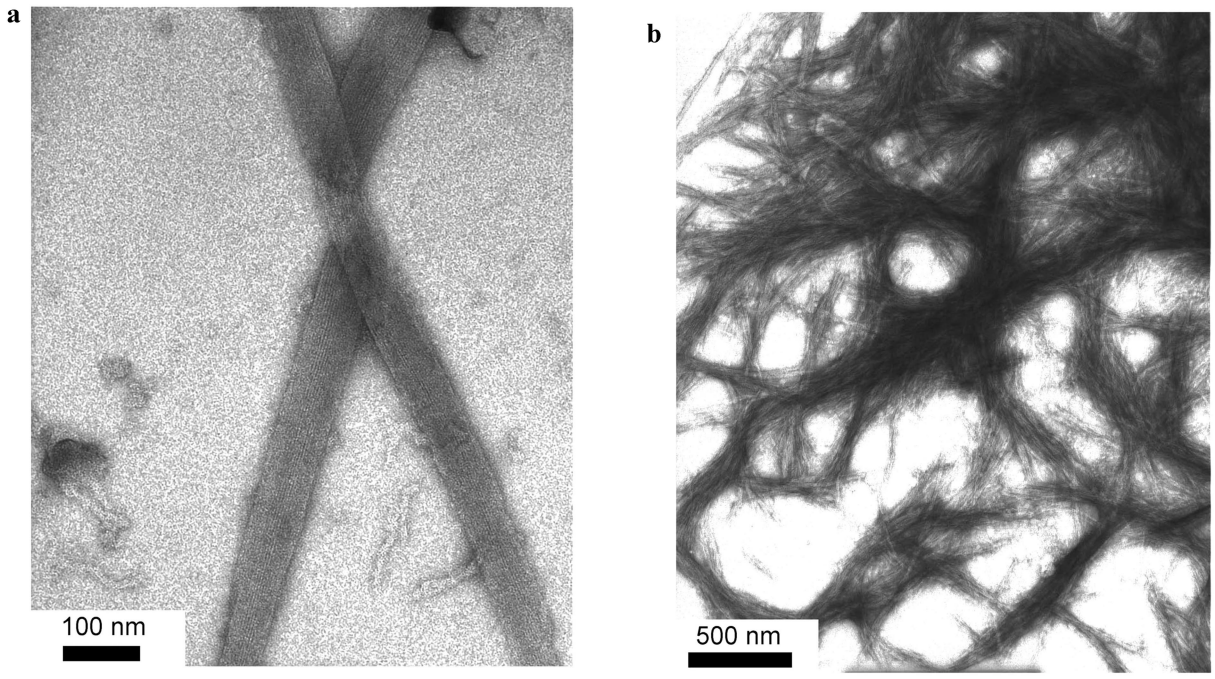


Figure 6. Assemblies formed by the self-templating peptide (STeP). Electron micrographs for STeP (a) and STePE (b) fibrils formed at neutral pH.

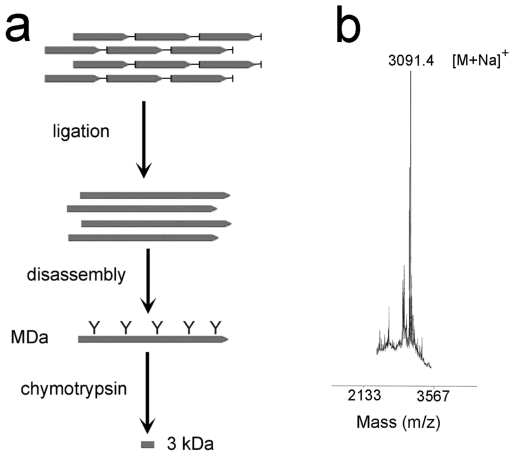


Figure 7. Assembly and digestion of the STePE. (a) Cartoon for the proposed mode of assembly of the STePE and how these might be digested by chymotrypsin. (b) Mass spectrum for chymotrypsin-digested STePE.

namely, a noncovalently assembled peptide fiber. Specifically, we have engineered a coiled coil-based fiber using understood sequence-to-structure relationships for peptide folding and assembly. The fibers are formed by chemically synthesized short peptides, which self-organize in a head-to-tail fashion to form supramolecular helices.[20−25] They are paracrystalline and exhibit hierarchical order on the nanoscale.[22] Specifically, the peptide extremities are noncovalently linked, which drives their longitudinal alignment into pseudo polypeptide chains. By using peptides with C-terminal thioesters, the termini of the assembled peptides can be linked (ligated) to form true polypeptide chains.

In addition to adding to methods available in peptide and protein synthesis, it is interesting to speculate to what uses these extremely long polypeptides might be put. They could, for example, be used in basic protein chemistry and physics experiments to understand better the physical properties of extended polypeptide chains either 'unfolded' or folded as extended α-helices; or they might be used in cell biology to generate novel polypeptide-based materials with repeated sequences that encode cell-binding and capture motifs.

Materials and Methods

Peptide Synthesis and Mass Spectrometry. Peptides were assembled on a PS3 automatic synthesizer (Protein Technologies, Tucson) and a Liberty Peptide Synthesis System (CEM Microwave Techniology, Matthews) using standard Fmoc/*t*Bu solid-phase protocols and HBTU/DIPEA as coupling reagents. The C-termini of the peptides were converted to thioesters on a Rink Amide MBHA resin using orthogonal chemistry of glutamate residues in conjunction with standard allyl chemistry protocols.[33] Fmoc-Glu-OAl served as a starting material, and the N-terminal amino acid was Boc-protected. Following removal of the allyl group, benzyl mercaptan was coupled to the α-carboxyl of a glutamate using 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide (EDC). The resulting thiobenzyl ester-terminated peptides were cleaved from the resin by TFA-TIS-HSBn (95:3.5:1.5) mixture.

Peptides were purified by RP-HPLC and confirmed by mass spectrometry (ESI and MALDI-TOF). MS $[M + H]^+$: SAF-p1, *m*/*z* 3173 (calcd), 3174 (found); SAF-p2, *m*/*z* 3324 (calcd), 3324 (found); SAF-p1E, *m*/*z* 3280 (calcd), 3280 (found); SAF-p2E, *m*/*z* 3431 (calcd), 3432 (found); acSAF-p1E, *m*/*z* 3322 (calcd), 3322 (found); SteP, *m*/*z* 3068 (calcd), 3069 (found); STePE, *m*/*z* 3175 (calcd), 3175 (found).

Fiber Assembly, Ligation, and Digestion. Fibers were assembled as described elsewhere.[23] Typically, 200 μL samples (100 μM of each peptide) were incubated overnight in filtered (0.22 μm) aqueous 10 mM MOPS, pH 7.4−7.6, room temperature. To enhance ligation, the preparations were titrated with 1 M NaOH up to pH 8.5 unless stated otherwise. Carbodiimide-mediated couplings were attempted under the standard assembly conditions for SAFs using EDC (1.1 molar equiv) over 15−30 min. In all cases, ligated peptides were either freeze-dried and resuspended in TFA/water (1:1) or 1 M KF or isolated by SEC prior to MALDI-TOF analysis (Figure S1, Supporting Information). Where stated, the peptides were freeze-dried and resuspended in 10

(33) Kates, S. A.; Daniels, S. B.; Albericio, F. *Anal. Biochem.* **1993**, *212*, 303−310.

mM phosphate buffer (pH 7.4, 100 μL) prior to further incubation with chymotrypsin (1−3 μL, 1 mg/mL) for 1−3 h at 37 °C. After the incubation, 20 μL of TFA or 0.1 N HCl was added to the preparations. Two microliter fractions were analyzed by MALDI-TOF (Figure S2, Supporting Information). Standard fiber samples (SAF and STeP) were used as controls throughout.

Circular Dichroism Spectroscopy. Circular dichroism spectroscopy was performed on a JASCO J-810 spectropolarimeter fitted with a Peltier temperature controller as described elsewhere.[23] All measurements were taken in ellipticities in mdeg and converted to molar ellipticities ([θ], deg cm^2 dmol res^{-1}) by normalizing for the concentration of peptide bonds. Thermal denaturation curves were recorded at 1−2 °C intervals using 1-nm bandwidth, with the signal averaged for 16 s and with a 2 °C/min ramp rate.

Transmission Electron Microscopy. Following incubations, 8 μL drops of peptide solutions were applied to carbon-coated copper specimen grids (Agar Scientific) and dried. The grids were stained with 0.5% uranyl acetate (8 μL) for 10−20 s and examined in a JEOL JEM 1200 EX MKI microscope at the accelerating voltage of 100 kV. Images were digitally acquired with a fitted camera (MegaViewII). Additional images are shown in Figure S3−5 in Supporting Information.

Size-Exclusion Chromatography. Ligated peptides were analyzed using a PL-aquagel-OH mixed column (Polymer Laboratories) with mass range of 100 to 10 M. The column was calibrated with PEO standards. Phosphate (10 mM, pH 7.4), 20 mM HEPES/100 mM KCl (pH 8), or 8 M aqueous urea all containing 0.02% sodium azide were used as elution buffers.

Photon Correlation Spectroscopy. Ligated peptides resuspended in water to final concentrations of 0.5−1 mg/mL were analyzed on a Malvern 4800 Autosizer. Hydrodynamic radii were obtained through the fitting of autocorrelation data using the manufacturer's software.

Acknowledgment. We thank the BBSRC of the UK for financial support (grants 04676 and E20126), and Prof. Steve Mann and the Woolfson Group at Bristol for valuable discussions.

Supporting Information Available: Further characterization data. This material is available free of charge via the Internet at http://pubs.acs.org.

JA072960S