Bilayer Fibril Formation by Genetically Engineered Polypeptides: Preparation and Characterization

Natalya I. Topilina,[†] Seiichiro Higashiya,[†] Narender Rana,[‡] Vladimir V. Ermolenkov,[†] Christopher Kossow,[‡] Autumn Carlsen,[‡] Silvana C. Ngo,[†] Christopher C. Wells,[†] Eric T. Eisenbraun,[‡] Kathleen A. Dunn,[‡] Igor K. Lednev,[†] Robert E. Geer,[‡] Alain E. Kaloyeros,[‡] and John T. Welch*[,†]

Department of Chemistry and College of Nanoscale Science and Engineering, The University at Albany, State University of New York, Albany, New York 12222

Received November 25, 2005; Revised Manuscript Received February 7, 2006

A de novo, genetically engineered 687 residue polypeptide expressed in *E. coli* has been found to form highly rectilinear, β-sheet containing fibrillar structures. Tapping-mode atomic force microscopy, deep-UV Raman spectroscopy, and transmission electron microscopy definitively established the tendency of the fibrils to predominantly display an apparently planar bilayer or ribbon assemblage. The ordered self-assembly of designed, extremely repetitive, high molecular weight peptides is a harbinger of the utility of similar materials in nanoscience and engineering applications.

Introduction

Molecular self-assembly is critical to the success of nanoscale molecular devices ranging from molecular nanoelectronic assemblies[1,2] to biologically functional nanostructures.[3] Successful approaches can employ biomimicry by utilizing biomaterial[1] self-assembly to form one-dimensional (1D) ordered, nanotemplates potentially suitable for 1D charge transfer.[4−10] Polypeptides that fold via established rubrics,[11] undergo intermolecular self-assembly,[12] and present an array of functionalities are attractive and chemically robust candidates such as templates. β-Sheet forming sequences, folded polypeptides that readily coacervate and thereby significantly increase the size of an autonomously formed assemblage, are of particular interest. Fibrillar structures, including well-known amyloid fibrils,[13−15] frequently involve face-to-face β-sheet interactions with individual strands oriented perpendicular to the fibril axis.[16] Amyloid fibril formation has been reported from assemblies composed of β-sheet forming polypeptides comprised of as few as 8 to 50 amino acids.[17−20] We have found that a de novo designed polypeptide sequence consisting of 687 amino acid residues in a regular repeat, γ-turn containing structure[21,22] self-assembles to form rigid, planar fibrils.[23] Intersheet aggregation of the antiparallel β-sheet structures with charged edges has been found to be effective in promoting aqueous solubility[24] but counterintuitively promotes controlled aggregation as assessed on defined surfaces.[25] Formation of rigid, planar fibrils such as these is atypical for amyloid fibrils. However, the observed fibrils, consisting of large polypeptide molecules with a defined architecture, could be considered to be derived from an array of prismoid blocks, where each unit consists of 84 antiparallel beta strands. The resultant 3.5 nm × 38 nm building blocks present precisely located functionality in the final assembly. Utilization of this genetically engineered approach for the production of large monodisperse polypeptide block copolymers, with precise dimensional and sequence control, affords access to a variety of materials that may have biomolecular or bioelectronic applications.

Experimental Section

Materials. Restriction endonucleases *Bam*HI, *Eco*RI, and *Bsa*I, T4 polynucleotide kinase were purchased from New England Biolabs (Beverly, MA). DNA Ligation Kit and benzonase were purchased from Fermentas Inc. (Hanover, MD) and Novagen, Inc. (Madison, WI), respectively. Inoue ultracompetent and electrocompetent cells of XL1-Blue (Stratagene, La Jolla, CA) and DH5αF′ (Invitrogen, Carlsbad, CA) were prepared according to standard methods.[26] Competent cells of BLR(DE3) and BLR(DE3)pLysS were purchased from Novagen, Inc. (Madison, WI). BLR(DE3)pLysSRARE was prepared from BLR-(DE3) transformed by pLysSRARE isolated from Rosetta(DE3)pLysS (Novagen, Inc., Madison, WI) and the frozen stock of the chemical competent cells (containing 10% glycerol to prevent autolysis upon thawing) was prepared by standard method[26] and stored at −80 °C. Plasmids pUC18 and pET-28a-c were obtained from Bayou Biolabs (Harahan, LA) and Novagen, Inc. (Madison, WI), respectively. Plasmids and DNA fragments separated by agarose gel electrophoresis were purified using QIAprep Spin Miniprep Kit and QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA). Western blotting was done using SuperSignal West HisProbe kit (Pierce, Rockford, IL). Large-scale purification of polyhistidine-tagged repetitive polypeptides was affected with Ni-NTA Superflow affinity columns (Qiagen Inc., Valencia, CA).

Synthetic DNA Monomers and Adaptive Sequences. 20 mM of each pair of coding DNA, **1** (8Y), **2** (8E), **3** (8H), and **4** (8K), respectively, were mixed and then the 5′-hydroxyl groups were phosphorylated by T4 polynucleotide kinase at 37 °C for 30 min to 1 h. Deactivation of T4 polynucleotide kinase was done at 80 °C for 20 min and then the solution was slowly cooled to ambient temperature for annealing. The 5′ end of each of the adaptive oligonucleotide sequences **5b**, **6a**, **7b**, and **8a** (non-palindromic 5′ *Ban*I extruding ends) was phosphorylated for 30 min to 1 h at 37 °C, respectively, and then the kinase was deactivated at 80 °C for 20 min. To each of the solution was added the complimentary oligonucleotides **5a**, **6b**, **7a**, and **8b**, respectively, to reach a final concentration of 20 mM. The mixture was annealed to give **5** (A1-1), **6** (A1-2), **7** (A2-1), and **8** (A2-2),

* To whom correspondence should be addressed. E-mail: jwelch@uamail.albany.edu.

[†] Department of Chemistry.

[‡] College of Nanoscale Science and Engineering.

10.1021/bm0509016 CCC: $33.50
Published on Web 03/09/2006

Figure 1. DNA design and sequences.

respectively. The resulting "inside-phosphorylated" sequences were combined appropriately to give respectively the adaptive sequences **5** and **6** (A1) and **7** and **8** (H6-H6) at a final concentration of 1 mM. See Figure 1.

Preparation of Recipient Vectors. pUC18 (10 μg) and pET-28a-c (1 μg) were double-digested with *Bam*HI-*Eco*RI and *Nco*I-*Bam*HI, respectively, under the recommended conditions and then purified by agarose gel electrophoresis. Excised gel containing the desired fragments was purified using QIAquick Gel Extraction Kit and eluted with TE (pH 8.0, 100 μL) to give pUC18/*Bam*HI-*Eco*RI for efficient amplification of desired oligomerized units or eluted with TE (pH 8.0, 40 μL) to give pET-28/*Nco*I-*Bam*HI for expression of the coding sequences, respectively.

Block Copolymerization of 1, 2, 3, and 4 for Construction of 32YEHK Repetitive Polypeptides. First, **1** (8Y) and **2** (8E) were condensed with adapters **5** and **6** (A1) and the desired fragments were purified by agarose gel electrophoresis. Purified fragments were then cloned into pUC18/*Bam*HI-*Eco*RI. Plasmids harboring dimerized DNA coding sequences, 16YE (**9**), 16YY, 16EY, and 16EE, were identified. Similarly, plasmids harboring 16HK (**10**) were identified from ligation of **3** (8H) and **4** (8K) along with 16HH, 16KH, and 16KK. An equimolar amount of **9** and **10** was digested by *Bsa*I and the resulting mixture of fragments of 16YE and 16HK units was purified and ligated with adapter constructs **5** and **6**. Desired fragments of the adapted tetramers were isolated, purified, and cloned into pUC18/*Bam*HI-*Eco*RI. Plasmids harboring 32YEHK (**11**) were identified along with 16YE2 (32YEYE), 32KHYE, and 16KH2 (32KHKH). **11** was digested by *Bsa*I and the resulting 32YEHK fragment was purified, oligomerized with **5** and **6**, and then cloned into pUC18/*Bam*HI-*Eco*RI. The plasmid harboring 7 repeats of 32YEHK (**12**, 32YEHK7) was identified. **12** was digested by *Bsa*I and the resulting 32YEHK7 fragment was oligomerized in the presence of **7** and **8** (H6-H6) and cloned into pET-28a/*Bam*HI-*Eco*RI. A plasmid library harboring 7, 14, 21 repeats of 32YEHK repetitive coding sequences bearing the H6-H6 adapter sequence could be prepared by this method. It is important to note that prolonged ligation usually resulted in significant degradation of vectors and inserts. The representative scheme from DNA concatamerization to peptide expression is shown in Figure 2.

Polypeptide Expression. pET-28-derived expression vectors harboring coding sequences were used to transform expression hosts, BLR-(DE3)pLysS or BLR(DE3)pLysSRARE (42 °C for 90 s and then 4 volume of SOC, 37 °C for 1 h). The transformants were selected on LB agar plates containing chloramphenicol (34 μg/mL), kanamycin sulfate (50 μg/mL), and 1% glucose without tetracycline. Selected colonies were inoculated into 2xYT (30 mL for 0.8 L of culture) containing kanamycin (50 μg/mL), chloramphenicol (34 μg/mL), and 1% glucose. The overnight precultures were used to inoculate desired amounts of expression culture (2xYT containing kanamycin (10 μg/mL) and chloramphenicol (34 μg/mL). The cultures were grown over $OD_{600} = 1.0$ and polypeptide expression was induced by the addition of IPTG to a final concentration of 1 mM. Cells were harvested after 4 h by centrifugation (3500*g*, 20 min at 4 °C). The cell pellets were resuspended in water (15 mL per 1 L of cell culture) and stored at −20 °C. Expression and purification of polypeptides were analyzed by SDS-PAGE and western-blotting using SuperSignal West HisProbe kit.

Polypeptide Purification and Chemical Modification. *Method A.* The frozen cells were lysed by freeze−thaw sonication for 20 min in the presence of benzonase (2 μL per 1 L culture) and PMSF (phenylmethansulfonyl fluoride, 2 mM). The cells were then allowed to incubate for 30 min at RT. Urea (11 g per 1 L culture, 8 M at final concentration) was added, and the sample was incubated on a boiling water bath for 4 h with occasional mixing. In the course of denaturation, the lysine residues were converted to homocitrulline by urea. The resulting solution was centrifuged for 30 min at 15000*g* at 20 °C and the supernatant was diluted 2-fold with 10 mM imidazole, 8 M urea, PBS, or H_2O with 10 mM imidazole and then was recentrifuged for 1 h at 25000*g* at 20 °C.

The supernatant was applied to a Ni-NTA column. The column was washed with 40 mL portions of 10 and 20 mM imidazole in 8 M urea and PBS (phosphate buffered saline, 0.1 M phosphate buffer (pH 7.4), 0.5 M NaCl at final concentration) and then was eluted with 300 and 500 mM imidazole solutions. Most of the desired polypeptide was found in 300 mM fraction.

Method B. To the incubated lysate prepared by the same method as Method A guanidine hydrochloride (14 g per 1 L culture, 6 M at final concentration) was added, and the sample was incubated on a boiling water bath for 1 h with occasional mixing in the course of denaturation. The resulting solution was centrifuged for 30 min at 15000*g* at 20 °C and the supernatant was diluted 2-fold with H_2O and then recentrifuged for 2 h at 25000*g* at 20 °C.

The supernatant was applied to a Ni-NTA column. The yield of polypeptide was 30 mg per 1 L of cell culture.

MALDI-TOF-MS: calculated for $A_{253}E_{21}G_{338}H_{33}K_{21}Y_{21}$ (deformylmethionylated) 50639; found 50722. Carbamylated (urea-boiled) polypeptide was prepared from polypeptide-containing 8 M urea eluent by incubating the solution for 3 h on a boiling water bath. MALDI-TOF-MS: calculated for $A_{253}E_{21}G_{338}H_{33}K_{21}Y_{21}$+21 carbamylation (CHNO), 51542; found 51490.

The eluent containing the carbamylated polypeptide was dialyzed against doubly distilled H_2O for 2−5 days at 4 °C using a dialysis membrane with a 3500 Da molecular weight cutoff. After dialysis, formation of a gelatinous phase within the solution was observed. The gelatinous phase was separated by centrifugation at 15000*g* for 45 min at 4 °C and subsequent decantation of the supernatant. The polypeptide aggregates were rendered more soluble by adding to the gel an equal volume of water followed by intensive vortex mixing. An aqueous phase containing a soluble polypeptide fraction could be separated from the remaining gelatinous phase by centrifugation at 15000*g* for 45 min at 4 °C. The concentration of the polypeptide in this extract was determined by UV−vis spectroscopy using the tyrosine absorption maxima at 276 nm. This solution was used for all further studies.

Resonance Raman Measurements. A recently built, deep-UV, Raman instrument is described in detail elsewhere.[27] Briefly, the fourth harmonic (195 nm, ~2 mW) of the *Indigo S* laser system (coherent) is used to generate Raman scattering from a 0.6-mm diameter thermostatically controlled sample solution stream. A custom-built subtractive

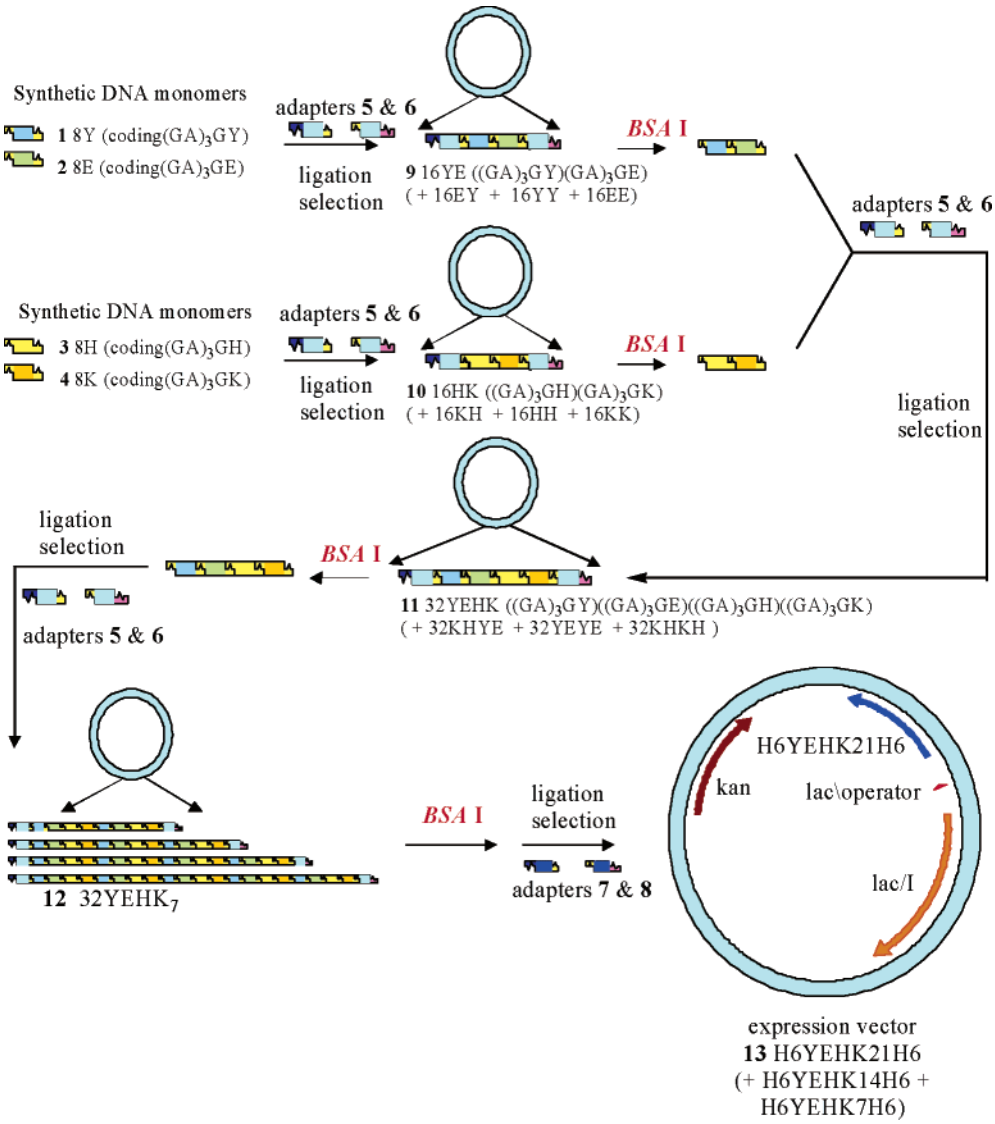

Figure 2. Construction of repetitive DNA sequences and expression.

double spectrograph equipped with a Roper Scientific *Spec-10:400B* CCD camera (liquid nitrogen cooled) is utilized for recording Raman spectra. The spectral resolution of the system is 4 cm^{-1}. Spectra are analyzed using GRAMS/AI (7.01) software.

Atomic Force Microscopy (AFM) Measurements. Sample Preparation: After dialysis, aqueous solutions of the polypeptide were used at a pH of 7.0 for further studies. A concentration of 0.12 mg/mL was determined using UV−visible spectroscopy.

Fifty microliters of the aqueous solution was dropped on a 12 mm × 12 mm sample of highly oriented pyrolytic graphite (HOPG) (Veeco Instruments). The polypeptide solution so deposited on HOPG was incubated at room temperature for 30 min and then removed by a micropipet. The surface was dried with a slow stream of nitrogen before imaging.

AFM Imaging Conditions. AFM imaging was performed in resonant tapping mode under ambient conditions with a MultiMode microscope (Digital Instruments) using a Nanoscope IIIa control system. A *J* scanner was used with a lateral range of ~125 μm. Images were collected using Nanoscope III software version 4.42r8 in height and phase mode simultaneously. The standard silicon TESP cantilevers (nominal spring constant, 40 N/m, resonance frequency of about 300 kHz, and tip radius less than 10 nm purchased from BudgetSensors) were used for imaging. The force was minimized during the imaging by choosing a setpoint corresponding to more than 90% of the free oscillation amplitude of about 12 nm. Typical scan frequency was 1 Hz and images were collected at a resolution of 512 × 512 points. The tip convolution of about 8 nm for features 1.4 nm in height was calculated by imaging the 1.4 nm diameter gold nanoparticles generally used for TEM calibration. Offline Nanoscope image analysis software (version 5.12r3) was used for image analysis. The images were flattened and no further image processing was performed.

Transmission Electron Microscopy (TEM) Measurements. Specimens were prepared for transmission electron microscope observation by placing a droplet of the peptide-containing solution (as above) on a standard carbon-coated support grid. After 30 min solution was removed and the specimen then stained using uranyl acetate to accentuate the morphology of the structures. Bright field TEM images were recorded on a film in a JEOL 200CX operating at 80 kV. All images were obtained at focus, using an objective aperture to introduce contrast into the images rather than defocus.

Results and Discussion

The β-sheet forming repetitive polypeptide blocks were designed with selected amino acids at the turn positions in order to facilitate control of coacervation and formation targeted supramolecular structures. The defined volumes and uniquely modified edges of the molecular subunits are controlled by the selection of turn-inducing amino acids. The resultant nanostruc-

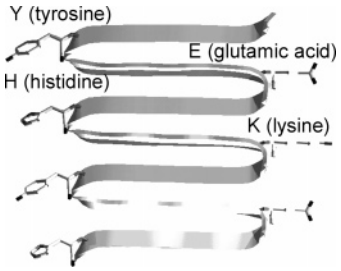

Figure 3. Structural model of the YEHK peptide construct.

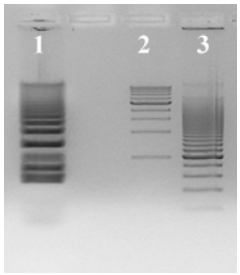

Figure 4. Representative agarose gels from oligomerization and isolation of populations of oligomers with enrichment of the longer fragments. Lane 1: H6YEHKH6 oligomers; lane 2: high molecular weight markers; lane 3: low molecular weight markers.

tures are therefore amenable to the introduction of functional (e.g., charge transport) moieties with potential applications in nanoelectronic devices. The construct consisting of repetitive polypeptides with 32 amino acid repeats, $(GA)_3GY(GA)_3GE(GA)_3GH(GA)_3GK$ (32YEHK), was designed based on the pioneering work of Tirrell[28] (Figure 3).

Genetic Coding. Inspired by Tirrell's findings,[29,30] DNA sequences coding β-sheet forming polypeptides, with GA repeats as the β-strand forming motif, were prepared and oligomerized in a head-to-tail manner.[31,32] Two linking DNA constructs were prepared for recursive oligomerization/block-copolymerization **5** and **6** and expressions **7** and **8** employing a modified method of Urry.[33] The adapter constructs **5** and **6** contain recognition sites for a type IIs restriction endonuclease *Bsa*I that facilitates recovery of the oligomerized units for recursive concatenation/block-copolymerization without inclusion of undesired DNA.[34−39] Expression adapters **7** and **8** were designed to attach a reactive amino acid appendage to the β-sheet, e.g., an oligohistidine tract without an intervening fusion sequence.

Both adaptive DNA sequences possess cohesive ends for ligation and cloning into commercially available vectors such as pUC18 and pET-28. This approach contrasts with published work[33] where the digestion of the oligomerized DNA by restriction endonucleases to generate cohesive ends was required. Utilization of concatenation/block copolymerization in the presence of adaptive DNA sequences containing the appropriate recognition sites of type II and IIs restriction endonucleases enabled, respectively, cloning and regeneration of assembled DNA units. Repetitive coding sequences could be prepared in a reproducible and predictive manner without using special cloning vectors. This strategy also suppresses the intramolecular cyclization of multimers that is problematic when longer sequences are constructed.[34] As a demonstration of feasibility, the repetitive 32YEHK-coding sequences were prepared by ligation of the synthetic oligonucleotides coding **1** (8Y) and **2** (8E) in the presence of adaptive sequences **5** and **6**. Oligonucleotides with the expected mobility on an agarose gel electrophoretogram (Figure 4) were purified and incorporated into a high-copy number, selection vector that could subsequently be cleaved by restriction digest.

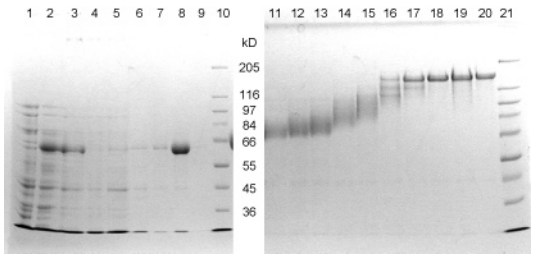

Figure 5. 8% SDS PAGE gel electrophoretic analysis of induction, purification, and modification of H6YEHK21H6. Lane 1, before induction; 2, 4 h induction; 3, whole cell lysate; 4, flow through 1; 5, flow through 2; 6, 10 mM imidazole; 7, 20 mM imidazole; 8, 300 mM imidazole; 9, 500 mM imidazole; 8, M urea PBS; 10 and 21, molecular weight marker (SIGMAMARKER Wide Molecular Range); 11, 300 mM fraction heated at 98 °C for 1 min; 12, 2 min; 13, 3 min; 14, 4 min; 15, 5 min; 16, 7 min; 17, 10 min; 18, 20 min; 19, 60 min; 20, 180 min.

Recombinant vectors harboring the (16YE)-coding sequence **9** were selected in *E. coli* and digested by *Bsa*I to recover the YE-coding oligonucleotide fragment (48 base pairs). The 16HK-coding fragment **10** was obtained similarly. These two fragments were subsequently joined in the presence of the adaptive DNA sequences **5** and **6**.

Utilizing this strategy, a recombinant vector harboring a 32YEHK-coding sequence was selected and then digested with *Bsa*I to recover the 32YEHK-encoding DNA fragment **11**. The 32YEHK fragments were oligomerized in the presence of **5** and **6** in a repetition of the aforementioned process. The resulting concatenated DNA fragment, e.g., an oligonucleotide tract of 7 repeats of 32YEHK (32YEHK7), was isolated. The 7-repeat sequence was further oligomerized in the presence of **7** and **8**, which add hexahistidine tracts on both N- and C-terminals of polypeptide products upon expression, and incorporated into an expression vector. In this process a library of expression vectors harboring 7, 14, and 21 repeats of 32YEHK-coding sequences was fashioned for subsequent expression in the appropriate *E. coli* hosts.

Peptide Isolation and Functionalization. On purification the desired β-sheet polypeptide (H6YEHK21H6) was isolated (Figure 5). The negatively charged glutamic acid (E) and positively charged lysine (K) residues originally were introduced at adjacent turn sites to facilitate antiparallel β-sheet assembly by salt bridge formation (see Figure 3). It was found however that when the crude cell lysate was denatured with a urea solution prior to affinity chromatography, the lysines were carbamylated to homocitrulline residues.[40] This modification did not seem to influence peptide folding; folding to form β-sheet containing structures usually occurred during the course of dialysis. Denaturation of the lysate with guanidinium hydrochloride solutions obviated this conversion. Confirmation of the carbamylation process can be detected by gel electrophoretic analysis (Figure 5). In a time course experiment, the polypeptide isolated from a guanidinium denatured solution was heated in the presence of 8 M urea. In lanes 11−20, it can be seen that this treatment led to a gradual increase in polypeptide mobility until all the lysine residues were converted to homocitrulline as established by MALDI-TOF-MS. It was found that polypeptide constructs containing homocitrulline residues exhibited a more pronounced propensity to undergo intermolecular aggregation to form fibrils than material prepared where the conversion was avoided. Therefore, our studies have focused on these homocitrulline containing peptides.

The peptide synthesis strategy described above enables the following: (a) use of conventional expression vectors, such as

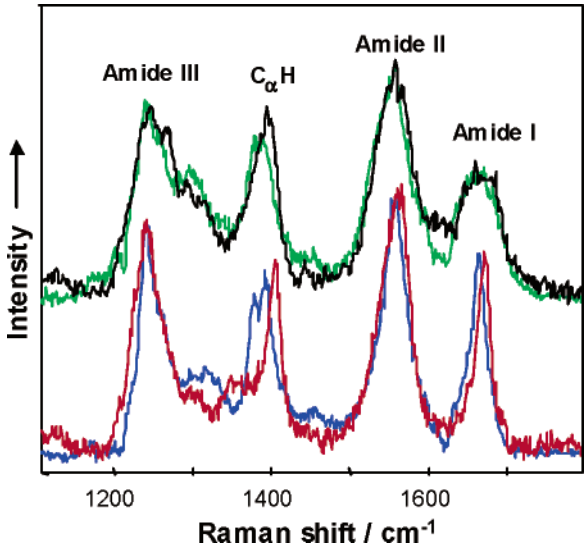

Figure 6. 195 nm excitation RR spectra of H6YEHK21H6 backbone obtained at 22 °C (blue curve) and 90 °C (green curve). Tyrosine contribution was subtracted. Red and black curves represent the RR spectra of poly-L-lysine (PLL) β-sheet and PLL random coil, respectively.

pUC18; (b) construction of a library of repetitive DNA sequences with higher multimer numbers; and (c) flexible incorporation of the adaptive DNA sequences, e.g., **5** and **6** or **7** and **8**, that are required for amplification and oligomerization of DNA constructs and for introduction of the completed DNA tract into an expression vector.

Characterization of H6YEHK21H6 by Deep UV Resonance Raman Spectroscopy. Far-UV CD spectroscopy is the conventional technique for the determination of protein secondary structure and for characterization of protein structure. With H6YEHK21H6 CD spectra did indicate the reversible melting of the β-sheet conformation on heating and subsequent incubation at room temperature. However, traditional deconvolution programs did not provide good approximations of H6YEHK21H6 secondary structure transformations. Since the CD phenomenon results from coupling of magnetic and electric dipole transition moments of coupled electronic transitions, stacked structures such as those resulting from the tyrosine and/or histidine side chains (Figure 3) facilitate coupling of the dipole transition moments and therefore may contribute to the anomalously intense CD signals observed.

In contrast to CD, the Raman scattering phenomenon is based on short-range interactions. Consequently, regular stacking of chromophores might result in no substantial contribution to the additive Raman scattering of individual chromophores. As a result, it was found that spectroscopic investigation of the polypeptide folding structure in solution that employed deep-UV Raman spectroscopy[27,41] was particularly informative. Illumination of the polypeptide with 195 nm light, near the amide chromophore $\pi-\pi^*$ electronic transition (absorption at ~190 nm), resonantly enhances Raman scattering from the polypeptide backbone. Consequently, the Raman spectrum, dominated by amide bands, represents a vibrational signature of secondary structural elements.[42,43] Figure 6 shows resonance Raman (RR) spectra of H6YEHK21H6 (blue and green lines) at 22 and 90 °C. A RR spectrum of tyrosine was subtracted from both these spectra to eliminate the contribution of the aromatic amino acid side chain. The β-sheet conformation dominates the H6YEHK21H6 secondary structure at room temperature as revealed by comparison of its RR spectrum with that of poly-L-lysine (red lines in Figure 6) in a predominantly

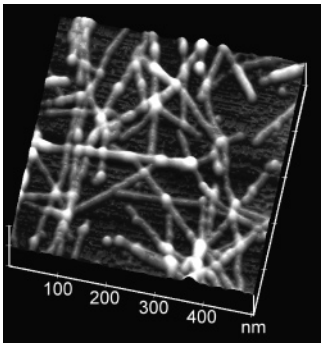

Figure 7. TM-AFM topograph of H6YEHK21H6 fibrils on HOPG.

β-sheet conformation (poly-L-lysine β-sheet assemblage was obtained by incubation at 52 °C in a pH 11.3 solution for 3 h).[42] At 90 °C, the H6YEHK21H6 RR spectrum closely resembles that of a poly-L-lysine random coil (black lines in Figure 6 recorded at room temperature and pH 4.0), indicating the temperature-induced melting of the H6YEHK21H6 β-sheet structure. Discrepancies between H6YEHK21H6 and poly-L-lysine spectra in the C_α-H bending mode region are attributed to differences in the amino acid residue sequences. The observed temperature-induced structural changes in the H6YEHK21H6 were found to be reproducibly reversible.

Characterization of H6YEHK21H6 by AFM. The spectroscopic study of polypeptide folding was followed by an investigation of the controlled self-assembly of the polypeptide on a defined surface. For that purpose, purified H6YEHK21H6 in doubly distilled water (DDW) was employed for solvent-based deposition on highly oriented pyrolytic graphite (HOPG) or polycrystalline Ni surfaces to investigate β-sheet self-organization. Deposition from the liquid phase was followed by DDW rinsing and dry-N_2 drying protocols. Topographic investigations utilized tapping-mode atomic force microscopy (TM-AFM). Figure 7 displays a TM-AFM topographic image of the HOPG surface following H6YEHK21H6 deposition. In contrast, the pre-deposition HOPG surface exhibited an atomically flat surface marked by the absence of any discernible features.

The most notable feature in Figure 7 is the presence of highly linear fibrillar H6YEHK21H6 structures. These structures are stable under ambient conditions and exhibit no conformational change after extended storage times. The fibril lengths range from tens to thousands of nanometers. The H6YEHK21H6 fibril width is uniform with an average value of 15 ± 2 nm as determined prior to deconvolution.

Figure 8 displays a probability density plot of fibril thickness derived from an analysis of 79 fibril samplings derived from multiple deposition experiments. Observed fibril thicknesses as determined by TM-AFM did not exhibit a continuum of thicknesses as might be expected from lamellar stacking.[44] Fibril thickness varies in discrete increments of approximately 0.8 ± 0.2 nm. Figure 8 clearly shows a substantially higher probability for fibrils with a thickness of 1.7 ± 0.2 nm, a thickness not inconsistent with a loose bilayer or "ribbon" configuration[20] (Figure 9). The observed fibril thickness increments are consistent with thickness estimates from the computationally derived H6YEHK21H6 polypeptide β-sheet structure. We speculate that the intrasheet interactions and the disposition of the hexahistidinyl tracts at the H6YEHK21H6 termini strongly favor a bilayer configuration. Nonetheless, we are cognizant of the complexity of turn formation and the implication of turn motif on amphiphilicity. Differing turn motifs can be assumed by β-hairpins as reflected in the number of amino acids per

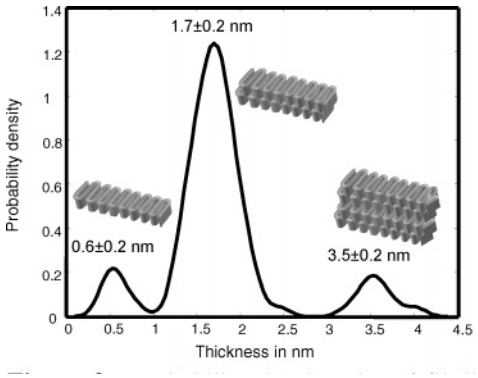

Figure 8. Probability density plot of fibril thickness derived from AFM topographic data of H6YEHK21H6 fibrils. 79 samplings were employed to generate the distribution profile.

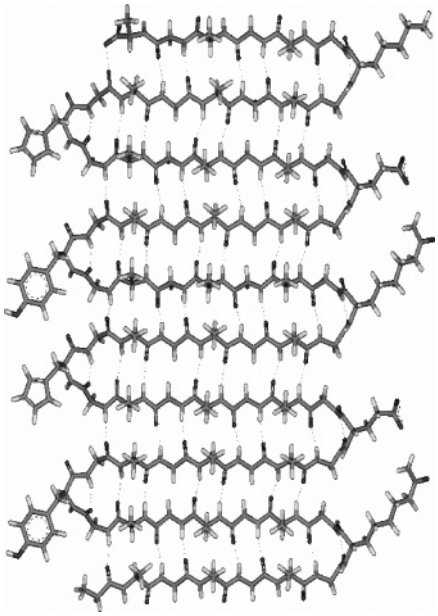

Figure 9. YEHK repeat units with γ-turns forming antiparallel β-sheet structure. Width (turn-strand turn) of computed model, 3.4−3.8 nm; thickness, 0.72 nm.

turn and the number of hydrogen bonds formed between the distal strands[45,46] with a consequence that the faces of the β-sheet may be differentiated by the orientation of the methyl groups of constituent strands toward a single face of the sheet. It has been previously shown that closely related constructs poly-[$(AG)_3EG$], poly[$(AG)_3YG$], and poly[$(AG)_3KG$] form amphiphilic β-sheet structures via the intermediacy of γ-turns to redirect the β-strand subunits.[44] If a γ-turn containing structure is assumed for H6YEHK21H6 as shown in Figure 9, the amphiphilic character of the resultant β-sheet would be consistent with the observed bilayer formation. The nonpolar alanine methyl bearing faces would be shielded from the solvent, while the glycyl derived moieties would be exposed to the aqueous environment.

Characterization of H6YEHK21H6 by TEM. To confirm this structure and further investigate H6YEHK21H6 fibril morphology, transmission electron microscope (TEM) studies were carried out. H6YEHK21H6 was deposited on carbon-coated support grids. Deposition protocols were similar to those for HOPG substrates. A uranyl acetate stain was utilized on post-dried samples to enhance edge contrast of the fibril assemblies in the TEM. Figure 10 displays resulting TEM micrographs. The highly linear assemblies in the upper image of Figure 10 agree qualitatively with the TM-AFM topographs

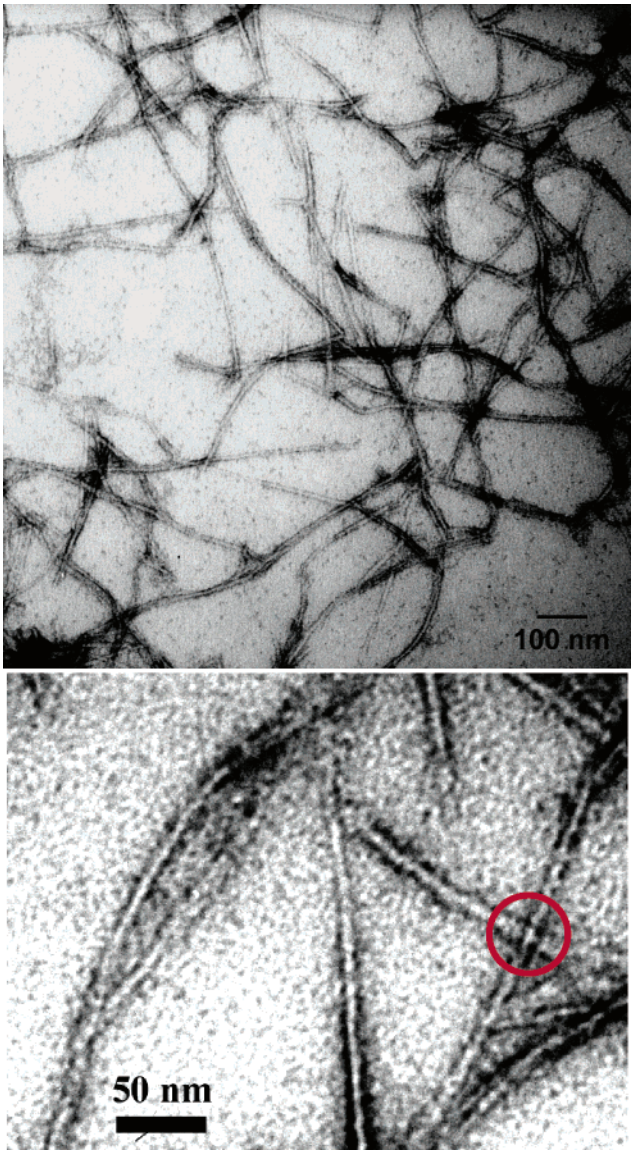

Figure 10. TEM micrographs of H6YEHK21H6 fibrils on carbon-coated Cu grids.

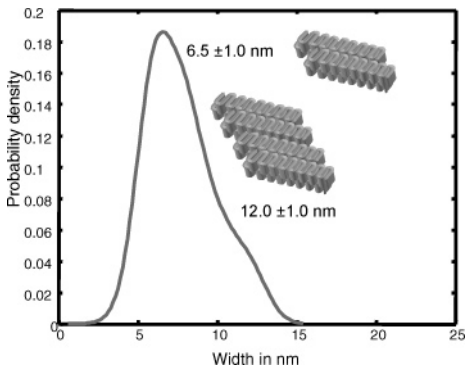

Figure 11. Probability density plot of fibril width derived from TEM of H6YEHK21H6 fibrils. 161 samplings were employed to generate the distribution profile.

of Figure 7. The lower image of Figure 10 displays a higher magnification micrograph of the fibril assemblies. On analysis of the TEM data, the average width of the fibrillar structures was found to be about 6.5 ± 1 nm as shown in Figure 11.

An average measured fibril width of 6.5 ± 1 nm was predominant for all TEM data studied. This is in approximate agreement with the width of a structure comprised of two

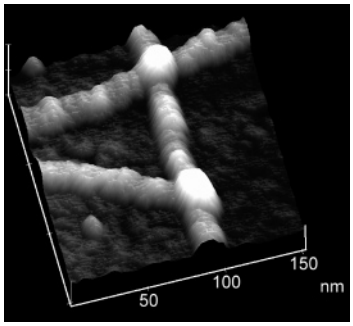

Figure 12. AFM topography of assembled domains of H6YEHK21H6 on HOPG.

β-sheets side-by-side where the hydrophobic turn groups (Tyr and His) of the two sheets pack closely (see width estimates in Figure 9).

As expected, TM-AFM data yielded a much larger average fibril width of 15 ± 2 nm due to tip-convolution effects. Calibration of the tip-convolution of the DI Nanoscope III AFM utilized Au nanospheres with a nominal diameter of 1.4 nm, chosen for the similarity of the sphere diameter to the fibril thickness. This calibration revealed a tip-convolution-induced feature broadening of 8 ± 1 nm and therefore resulted in very good agreement between the TEM and TM-AFM fibril-width measurements.

Fibril Formation. A higher resolution TM-AFM micrograph of crossed fibrils is shown in Figure 12. Note the apparent increase in the fibril thickness at the intersections of the fibrils. A similar crossing phenomenon is evidenced in the lower image of Figure 10 (red circle) where the distinct edges of two overlapping fibrils are seen via TEM. Both these sets of image data imply that the fibrils form in solution prior to deposition on the substrate. This contrasts with the surface-templated assembly of low molecular weight polypeptide β-sheets on HOPG that have been reported recently.[19,47,48] The surface disposition of the fibrillar structures presented herein more closely resemble observations by Marini et al.,[17] although the polypeptide β-sheets reported in that work consisted of only eight residues, nearly 2 orders of magnitude fewer than the polypeptides considered here.

It is important to note that β-sheets frequently exhibit a right-handed twist. An increase in the number of strands with a concomitant increase in number of H-bonds should improve sheet stability; however, the twist induces dis-symmetry in the interaction of side chains from neighboring strands.[49] Cooperative β-sheet formation has been reported in the formation of helical ribbons based upon the self-assembly of the octapeptide FKFEFKFE[50] where the KE and F domains lead to self-assembly of the individual molecular strands into a ribbon with an extended β-sheet geometry. Over longer times, the ribbons tend to form super helices around the initial strand.[17] It has also been reported that low molecular weight peptides assemble into elongated β-sheets containing tapes which then dimerize to form helical ribbons and subsequently aggregate to form fibrils by forming twisted lamellae.[20] In this work, helical assemblies were not commonly observed for H6YEHK21H6 in contrast to either of the above materials or the products described in ref 17.

Figure 13 depicts a possible assemblage of four H6YEHK21H6 molecules which can accommodate the observed dimensions of the fibrils that apparently form in solution. The amphiphilic character induced by the involvement of the γ-turn in reversing strand direction is consistent with the facile solution formation

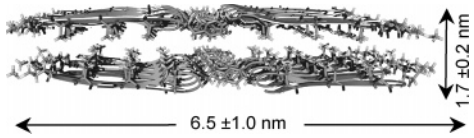

Figure 13. In a tentative structural model for the most common fibrils observed both the dimensions of those fibrils and the amphiphilic nature of the proposed antiparallel β-sheet can be accommodated by the interaction of four polypeptide molecules.

of bilayer thick fibrils. By design the β-sheet forming H6YEHK21H6 molecule has relatively hydrophobic and hydrophilic edges dominated by the turn groups; as a consequence, in the proposed model the more hydrophobic surfaces and turn groups of the molecules are shielded from the aqueous environment. Further experiments to validate this hypothesis are underway.

Conclusion

A β-sheet forming structure has been designed de novo which not only remains soluble in water but also aggregates to form on surfaces the non-templated fibrillar aggregates displayed in Figures 7, 10, and 12. While fibril formation is fundamental to several disease states, e.g., amyloidosis,[16,51] previous investigations of amyloid fibril formation examined assemblies composed of β-sheets with a relatively small number of 8 to 50 linearly arranged residues,[17−19] as compared to the synthetic oligomeric structure consisting of 687 residues in the regular repeating β-turn containing structure reported herein.

The novel self-assembly motifs described not only have implications for the understanding of fundamental polypeptide aggregation processes central to diseases, such as familial amyloid polyneuropathy (FAP), Alzheimer's, or bovine spongiform encephalopathy (BSE), but also inform the design of nanoscale architectures with potential utility in molecular device engineering. The description of such novel self-assembly processes has prompted our investigation of the utility of these materials as templates for ordered molecular assembly on defined surfaces, precisely the assembly required by nanoelectronic devices. Yet, these same processes are critical in elucidating the fundamental mechanisms of intermolecular protein association where stable β-sheet derived structures are central to the etiology and pathogenesis of disease.

Acknowledgment. This research was supported by the Microelectronics Advanced Research Corporation (MARCO), the Defense Advanced Research Programs Agency (DARPA), and the New York State Office of Science, Technology and Academic Research (NYSTAR) through the Interconnect Focus Center (IFC) and the Materials, Structures, and Devices (MSD) Center. MALDI and ESI mass spectrometry were performed at the Mass Spectrometry Facility at the Department of Chemistry and Chemical Biology, Rensselaer Polytechnic Institute, that were supported by NSF Grants CHE-0091892 and CHE-0078056 and we owe our appreciation to Prof. Dmitri Zagorevski for his helpful advice with mass spectroscopy. We acknowledge the kind assistance of Prof. Jeffrey L. Travis of the Department of Biological Science, University at Albany, and Prof. William A. Samsonoff at Wadsworth Center, New York State Department of Health, for their kind help with TEM imaging. In addition, it is a pleasure to recognize the contributions to this work of Mr. Kenneth S. Bousman, Profs. Richard P. Cunningham, Charles P. Scholes, and Dr. Evelina Loghin.

CDV

References and Notes

(1) Collier, C. P.; Wong, E. W.; Belohradsky, M.; Raymo, F. M.; Stoddart, J. F.; Kuekes, P. J.; Williams, R. S.; Heath, J. R. *Science* **1999**, *285*, 391−394.
(2) Wang, W.; Lee, T.; Reed, M. A. *Phys. Rev.* **2003**, *B68*, 35416.
(3) Scheibel, T.; Parthasarathy, R.; Sawicki, G.; Lin, X.-M.; Jaeger, H.; Lindquist, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4527−4532.
(4) Rochefort, A.; Martel, R.; Avouris, P. *Nano Lett*. **2002**, *2*, 877.
(5) Braun, E.; Eichen, Y.; Sivan, U.; Ben-Yoseph, G. *Nature* **1998**, *391*, 775−778.
(6) Thurn-Albrecht, T.; Schotter, J.; Kastle, G. A.; Emley, N.; Shibauchi, T.; Krusin-Elbaum, L.; Guarini, K.; Black, C. T.; Tuominen, M. T.; Russell, T. P. *Science* **2000**, *290*, 2126−2129.
(7) Kirsch, R.; Mertig, M.; Pompe, W.; Wahl, R.; Sadowski, G.; Böhm, K. J.; Unger, E. *Thin Solid Films* **1997**, *305*, 248−253.
(8) Kim, M. L.; Wong, K. W.; Mann, S. *Chem. Mater.* **1999**, *11*, 23−26.
(9) Mirkin, C. A.; Taton, T. A. *Nature* **2000**, *405*, 626−627.
(10) Mbindyo, J. K.; Reiss, B. D.; Martin, B. R.; Keating, C. D.; Natan, M. J.; Mallouk, T. E. *Adv. Mater.* **2001**, *13*, 249−254.
(11) Daggett, V.; Fercht, A. *Nat. Rev. Mol. Cell Biol*. **2003**, *4*, 497−502.
(12) Ellgaard, L.; Helenius, A. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 181−191.
(13) Guijarro, J. I.; Sunde, M.; Jones, J. A.; Campbell, I. D.; Dobson, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4224−4228.
(14) Chiti, F.; Webster, P.; Taddei, N.; Clark, A.; Stefani, M.; Ramponi, G.; Dobson, C. M**.** *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3590−3594.
(15) Fändrich, M.; Fletcher, M. A.; Dobson, C. M. *Nature* **2001**, *410*, 165−166.
(16) Kelly, J. W. *Curr. Opin. Struct. Biol.* **1998**, *8*, 101−106.
(17) Marini, D. M.; Hwang, W.; Lauffenburger, D. A.; Zhang, S.; Kamm, R. D. *Nano Lett.* **2002**, *2*, 295−299.
(18) Yang, G.; Woodhouse, K. A.; Yip, C. M. *J. Am. Chem. Soc*. **2002**, *124*, 10648−10649.
(19) Kowalewski, T.; Holtzman, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3688−3693.
(20) Aggeli, A.; Nyrkova, I. A.; Bell, M.; Harding, R.; Carrick, L.; McLeish, T. C. B.; Semenov, A. N.; Boden, N. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 11857−11862.
(21) Tirrell, D. A.; Tirrell, J. G.; Mason, T. L.; Fournier, M. J. In *Bioorganic Chemistry: Peptides and Proteins*; Hecht, S. M., Ed.; Oxford University Press: New York, 1998; pp 446−472, 519−521.
(22) Van Hest, J. C. M.; Tirrell, D. A. *Chem. Commun*. **2001**, 1897−1904.
(23) Hamada, D.; Yanagihara, I.; Tsumoto, K. *Trends Biotechnol*. **2004**, *22*, 93−97.
(24) Richardson, J. S.; Richardson, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2754−2759.
(25) Zhang, S.; Holmes, T.; Lockshin, C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3334−3338.
(26) Sambrook, J.; Russel, D. W. *Molecular Cloning*, 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 2001.
(27) Lednev, I. K.; Ermolenkov, V. V.; He, W.; Xu, M. *Anal. Bioanal. Chem.* **2005**, *381* (2), 431−437.
(28) Cantor, E. J.; Atkins, E. D. T.; Cooper, S. J.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *J. Biochem.* **1997**, *122*, 217−225.
(29) Yoshikawa, E.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *Macromolecules* **1994**, *27*, 5471−5475.
(30) Cantor, E. J.; Atkins, E. D. T.; Cooper, S. J.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *J. Biochem.* **1997**, *122*, 217−225.
(31) Higashiya, S.; Ngo, S. C.; Bousman, K. S.; Jin, X.; Welch, J. T.; Cunningham, R. P.; Eisenbraun, E. T.; Geer, R. E.; Kaloyeros, A. E. *Polym. Prepr. (Am. Chem. Soc., Div. Polym. Chem.)* **2003**, *44*, 679−680.
(32) Ferrari, F. A.; Richardson, C.; Chambers, J.; Causey, S. C.; Pollock, T. J.; Capello, J.; Crissman, J. W. U.S. Patent 5,243,038, 1993.
(33) McPerson, D. T.; Xu, J.; Urry, D. W. *Protein Expression Purif.* **1996**, *7*, 51−57.
(34) Won, J.-I.; Barron, A. E. *Macromolecules* **2002**, *35*, 8281−8287.
(35) McMillan, R. A.; Lee, T. A.; Conticello, V. P. *Macromolecules* **1999**, *32*, 3643−3648.
(36) McMillan, R. A.; Conticello, V. P. *Macromolecules* **2000**, *33*, 4809−4821.
(37) Zhou, Y.; Wu, S.; Conticello, V. P. *Biomacromolecules* **2001**, *2*, 111−125.
(38) Goeden-Wood, N. L.; Conticello, V. P.; Muller, S. J.; Keasling, J. D. *Biomacromolecules* **2002**, *3*, 874−879.
(39) Meyer, D. E.; Chilkoti, A. *Biomacromolecules* **2002**, *3*, 357−367.
(40) Zhen, L.; Janusz, P. *J. Proteome Res.* **2004**, *3*, 567−571, and references therein.
(41) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc*. **1999**, *121*, 8074−8086.
(42) Chi, Z.; Chen, X. G.; Holtz, J. S.; Asher, S. A. *Biochemistry* **1998**, *37*, 2854−2864.
(43) Copeland, R. A.; Spiro, T. G. *Biochemistry* **1987**, *26*, 2134−2139.
(44) Parkhe, A. D.; Cooper, S. J.; Atkins, E. D. T.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *Int. J. Biol. Macromol*. **1998**, *23*, 251−258.
(45) de Alba, E.; Jimenez, M. A.; Rico, M. *J. Am. Chem. Soc.* **1997**, *119*, 175−183.
(46) Sibanda, B. L.; Blundell, T. L.; Thornton, J. M. *J. Mol. Biol.* **1989**, *206*, 759−77.
(47) Brown, C. L.; Aksay, I. A.; Saville, D. A.; Hecht, M. H. *J. Am. Chem. Soc.* **2002**, *124*, 6846−6848.
(48) Kogan, M. J.; Dalcol1, I.; Gorostiza, P.; Lopez-Iglesias, C.; Pons, M.; Sanz, F.; Ludevid, D.; Giralt, E. *J. Mol. Biol.* **2001**, *312*, 907−913.
(49) Lacroix, E.; Kortemme, T.; Lopez de la Paz, M.; Serrano, L. *Curr. Opin. Struct. Biol*. **1999**, *9*, 487−493.
(50) Hwang, W.; Marini, D. M.; Kamm, R. D.; Zhang, S. *J. Chem. Phys*. **2003**, *118*, 389−397.
(51) Tan, S. Y. *Histopathology* **1994**, *25*, 403−414.

BM0509016