

High-resolution NMR characterization of a spider-silk mimetic composed of 15 tandem repeats and a CRGD motif

Glendon D. McLachlan,^{1*} Joseph Slocik,² Robert Mantz,² David Kaplan,³ Sean Cahill,⁴ Mark Girvin,⁴ and Steve Greenbaum¹

¹Department of Physics and Astronomy Hunter College, City University of New York, New York, NY 10065

²Air Force Research Laboratory, Wright-Patterson AFB, Wright Paterson, OH 45433

³Department of Biomedical Engineering, Chemical and Biological Engineering, Bioengineering and Biotechnology Center, Tufts University, Medford, MA 02155

⁴Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY 10461

Received 25 June 2006; Revised 20 October 2008; Accepted 21 October 2008

DOI: 10.1002/pro.12

Published online 2 December 2008 proteinscience.org

Abstract: Multidimensional solution NMR spectroscopic techniques have been used to obtain atomic level information about a recombinant spider silk construct in hexafluoro-isopropanol (HFIP). The synthetic 49 kDa silk-like protein mimics authentic silk from *Nephila clavipes*, with the inclusion of an extracellular matrix recognition motif. 2D ¹H-¹⁵N HSQC NMR spectroscopy reveals 33 cross peaks, which were assigned to amino acid residues in the semicrystalline repeat units. Signals from the amorphous segments in the primary sequence were weak and broad, suggesting that this region is highly dynamic and undergoing conformational exchange. An analysis of the deviations of the ¹³C_α, ¹³C_β, and ¹³CO chemical shifts relative to the expected random coil values reveals two highly α -helical regions from amino acid 12–19 and 26–32, which comprise the polyalanine track and a GGLGSQ sequence. This finding is further supported by ϕ -value analysis and sequential and medium-range NOE interactions. Pulsed field gradient NMR measurements indicate that the topology of the silk mimetic in HFIP is nonglobular. Moreover, the 3D ¹⁵N-NOESY HSQC spectrum exhibits few long-range NOEs. Similar spectral features have been observed for repeat modules in other polypeptides and are characteristic of an elongated conformation. The results provide a residue-specific description of a silk sequence in nonaqueous solution and may be insightful for understanding the fold and topology of highly concentrated, stable silk before spinning. Additionally, the insights obtained may find application in future design and large-scale production and storage of synthetic silks in organic solvents.

Keywords: multidimensional NMR; circular dichroism (CD); nuclear overhauser effect (NOE); hexafluoro-isopropanol (HFIP); silk mimetic

Introduction

The exceptional physical properties of silks have been attributed to the molecular architecture of the polypeptide chain and processing of silk in solution prior

to fiber formation.^{1,2} The silk-fiber's tensile strength coupled with its low density, stretch ability, toughness, and biodegradable feature makes this an attractive polymer with material science and medical applications.¹ Therefore, the design of new and robust materials with a wide range of applications might be realized from duplicating the molecular design of silk proteins.^{3–7}

Native silk fibroins from silkworms and spiders are block copolymers, consisting of semicrystalline blocks covalently linked by amorphous modules.^{8–11}

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: AFOSR.

*Correspondence to: Glendon D. McLachlan, Department of Physics and Astronomy Hunter College, City University of New York, New York, NY 10065. E-mail: gmclachl@hunter.cuny.edu.

N-terminus 1 MHHHH HHSSG LVPRG SGMKE TAAK FERQH 30
31 MDSPD LGTDD DDKAM ADIGS **CRGD** 54
55 **(TSGRG GLGGQ GAGAA AAAGG AGQQG YGGLG SQG)**₁₅ 550
551 **TSRGDC** GSEFELRRQA CGRTRAPPPP PLRSGC 581 C-terminus

Figure 1. The primary sequence of the recombinant 49-kDa CRGD protein. The bold typed residues in brackets comprise the 33-amino acid repeat and the RGD motif is located adjacent to the N-terminus and C-terminus.

Although the composition of the primary sequence in the amorphous segments is dominated by hydrophilic amino acids, the semicrystalline units are predominantly hydrophobic and are primarily alanines and glycines. This minimal molecular design comprising a repetitive primary sequence, the secondary structural motifs, and the tertiary arrangement of the polypeptide chain results in a macromolecular silk-fiber with an exceptional array of material properties. Particularly, the major ampullate gland silk of *Nephila clavipes* has shown exceptional toughness, elasticity, and stability over a range of conditions.^{11,12} However, a complete description of the molecular interactions in native silks which account for these properties is incomplete. Therefore, silk-like peptides bearing the morphological features of the natural molecule have provided the bulk of information relating to local structure and fiber formation.

The Cystine-Arginine-Glycine-Aspartate (CRGD) silk mimetic's design is based on the consensus sequence of *Nephila clavipes* dragline silk combined with an extracellular matrix protein sequence (integrin recognition motif) covalently linked at the termini.^{5,7} The RGD motif is an epitope for osteoblast binding. The primary structure of the engineered CRGD-silk mimetic consists of a semicrystalline block forming a 33-amino acid sequence repeated 15 times, covalently linked to an amorphous sequence and the CRGD motif spliced between these sequences (see Fig. 1). Practical applications of the CRGD-silk mimetic include functioning as a platform for bone cell nucleation and growth, tissue engineering for cartilage repair, stem cell regeneration, and promoting silica composite growth and assembly in vitro, after functionalization with a silica regulating peptide.^{5,7,13,14} In addition to industrial applications, insights gained from studying silk-like mimetics can also provide valuable clues relevant to the processing of authentic silk proteins in vivo. Of particular interest are the proper folding and annealing of the prespun, authentic, and recombinant silk polymer. For example, silk from *Nephila edulis* is subjected to a series of chemical treatments and dehydration before fiber formation.^{15,16} The ionic fluxes within the silk glands will also alter the local structure and aggregation state of the highly concentrated spinning dope, as the environment in the spinning channel is altered. The mature fiber that results is highly ordered, rich in β -sheet content, and linked by amorphous stretches.¹⁷ Interestingly, the recent findings

that silk fibroins can form amyloid fibrils in vitro attest to the necessity for tight control of the processing in vivo, since premature formation of these fibrillar structures that are rich in β -sheet content would be deleterious to the spider and limiting in the case of mass production in industry.¹⁸ Similar schemes might be employed by silkworms as well, given that the physiological apparatus employed by spiders for silk processing has been delineated in *Bombyx mori*.^{19,20} Moreover, despite the differences in primary sequences between silk from spiders and silkworms, both have been suggested to exhibit similar secondary structural features.²¹

Hexafluoro-isopropanol (HFIP) is an organic solvent employed in artificial spinning and the regeneration of the silk fiber, and results in fibers with similar structural and mechanical features as the native fiber.^{22–26} Moreover, the difficulties associated with solubilizing hydrophobic sequences at concentrations relevant for NMR measurements are diminished by low polarity solvents such as HFIP. However, the mechanism by which fluorinated alcohols stabilizes secondary structures in proteins and peptides is not completely agreed upon, and includes stabilization of hydrogen bonds for α -helix formation, formation of micelle-like structures around hydrophobic surfaces, and thermodynamic and hydrophobic effects.^{27–30} In this report, HFIP is employed as a solubilizing, nonaqueous solvent, which might shed light on the conformational propensities of the silk-like protein.

The engineered silk mimetic offers the opportunity to incorporate features that minimize some of the experimental difficulties encountered with studying native silks, while retaining much of the molecular characteristics of the natural polymer. For example, truncated synthetic peptides with enhanced solubility have been used widely to study the salient features of silks, and have resulted in a number of the proposed molecular models.³¹ In this study, we present a description of a 49-kDa silk-like protein in a nonaqueous solvent at atomic resolution. Our results might have industrial applications and, more importantly, shed light on a salient conformation of the silk polypeptide chain during processing in nonaqueous environments, which are necessary for proper fiber formation.

Results and Discussion

2D ¹H-¹⁵N HSQC spectrum of CRGD in 100% HFIP

The conformational polymorphism exhibited by silk proteins is facilitated by a balance between hydrophobic and electrostatic interactions, which is regulated in vivo by a series of chemical changes, including pH gradient, K⁺, Na⁺, Cl⁻, and P ionic flux and dehydration of the silk solution before fiber formation.^{15,16} In vitro, the conformation of the CRGD-silk mimetic is dependent on the local chemical environment and undergoes

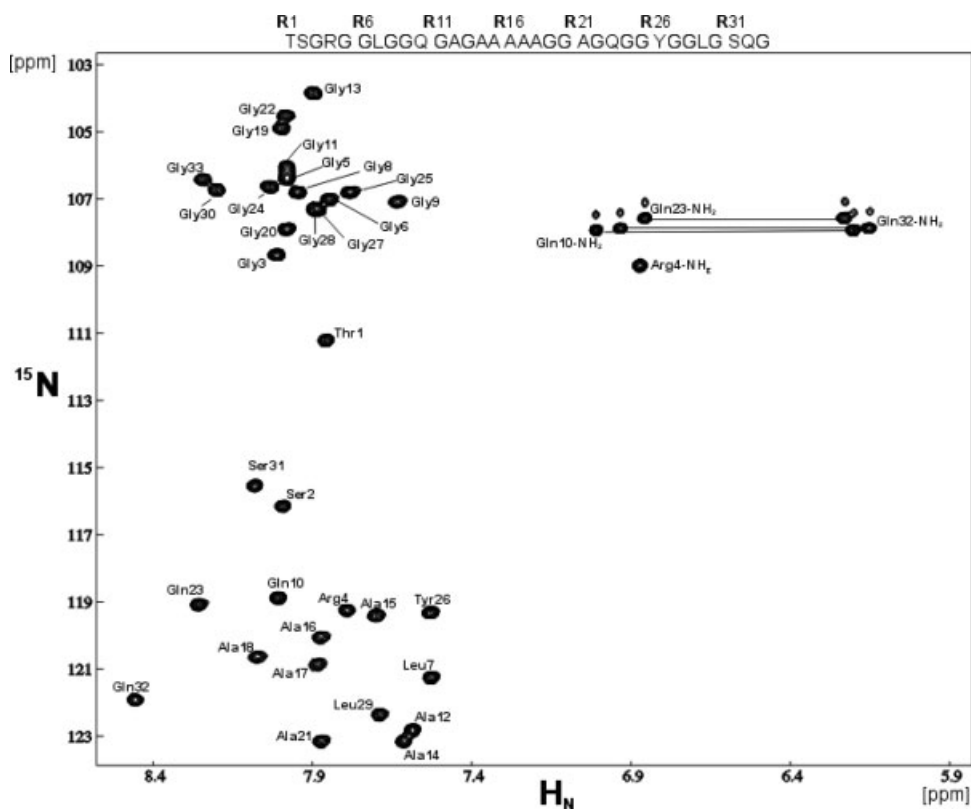


Figure 2. 2D ^1H - ^{15}N HSQC spectrum of the 49 kDa CRGD-15 repeat protein in 100% HFIP and 25°C. The 33 cross peaks, which are visible in the 2D-HSQC spectrum were assigned to the 33-amino acid repeats. The cross peaks are labeled according to the three-letter convention for amino acids and the residues labeled 1–33 correspond to the repeating residues starting at T55 as seen in Figure 1. Resonances representing the nonrepeating N and C-terminal sequences, which flank the 33 repeat-region, are not visible.

a conformational change to β -sheet structure when treated with methanol/water mixtures.⁵ The 2D ^1H - ^{15}N HSQC spectrum of the 49-kDa silk mimetic in nonaqueous HFIP and 25°C is shown in Figure 2. The spectrum exhibits 33 prominent cross peaks attributed to backbone resonances, and seven additional cross peaks, which were assigned to side-chain residues. These resonances were assigned to the 33 amino acids of the repeat region (semicrystalline) shown in the primary sequence in Figure 1. Spectral assignment was obtained from a combination of standard 3D data sets acquired at room temperature in 100% HFIP, that is, 3D HNCA, 3D HNCO, 3D CBCA(CO)NH, and 3D HNCOCACB. The chemical shift assignments are available as supplemental information. The semicrystalline region is composed of 15 identical units of 33 amino acids; therefore, the observed resonances indicate chemical shift degeneracy between the 15 repeats. Furthermore, the degenerate cross peaks suggest similar local environments, and is consistent with a single structural unit repeated 15 times. Signals from the putative amorphous N-terminal and C-terminal amino-acid sequences (84 amino acids combined) were either very weak or not visible in the 2D ^1H - ^{15}N HSQC spectrum. The signal intensity from this region of the sequence is expected to be reduced by a factor of 15

relative to the 33-amino acid repeat regions. However, these signals were broadened and exhibited greater reduction in intensity than anticipated. The discrepancy observed for the N and C terminal sequences can be attributed to line-broadening resulting from conformational exchange in the slow regime. The N-terminus of the CRGD-silk mimetic also includes an S-tag sequence (residues 14–38) used in purification protocols, which was reported to be unstructured in a corticotrophin-releasing factor receptor 2 β construct by NMR spectroscopic studies.³² Nevertheless, correlations to Asp54 (i.e., $i-1$), which is adjacent to the 33 repeat region can be made in the 3D experiments. The RGD sequence modification is not expected to influence the conformation of the protein but has functional implications, that is, as an epitope for cell-surface binding and is expected to exhibit little propensity to form structure. Earlier studies on authentic silk from *N. clavipes* in the solution state also observed a loss of spectral signals for residues in the amorphous regions and not in the structured sequences.³³ In that instance, loss of spectral signals was attributed to short ^{13}C T_1 times, which correlate with fast molecular motions. Figure 3 shows a typical 3D HNCA/CBCA(CO)NH pair used for making sequential assignments of backbone resonances.

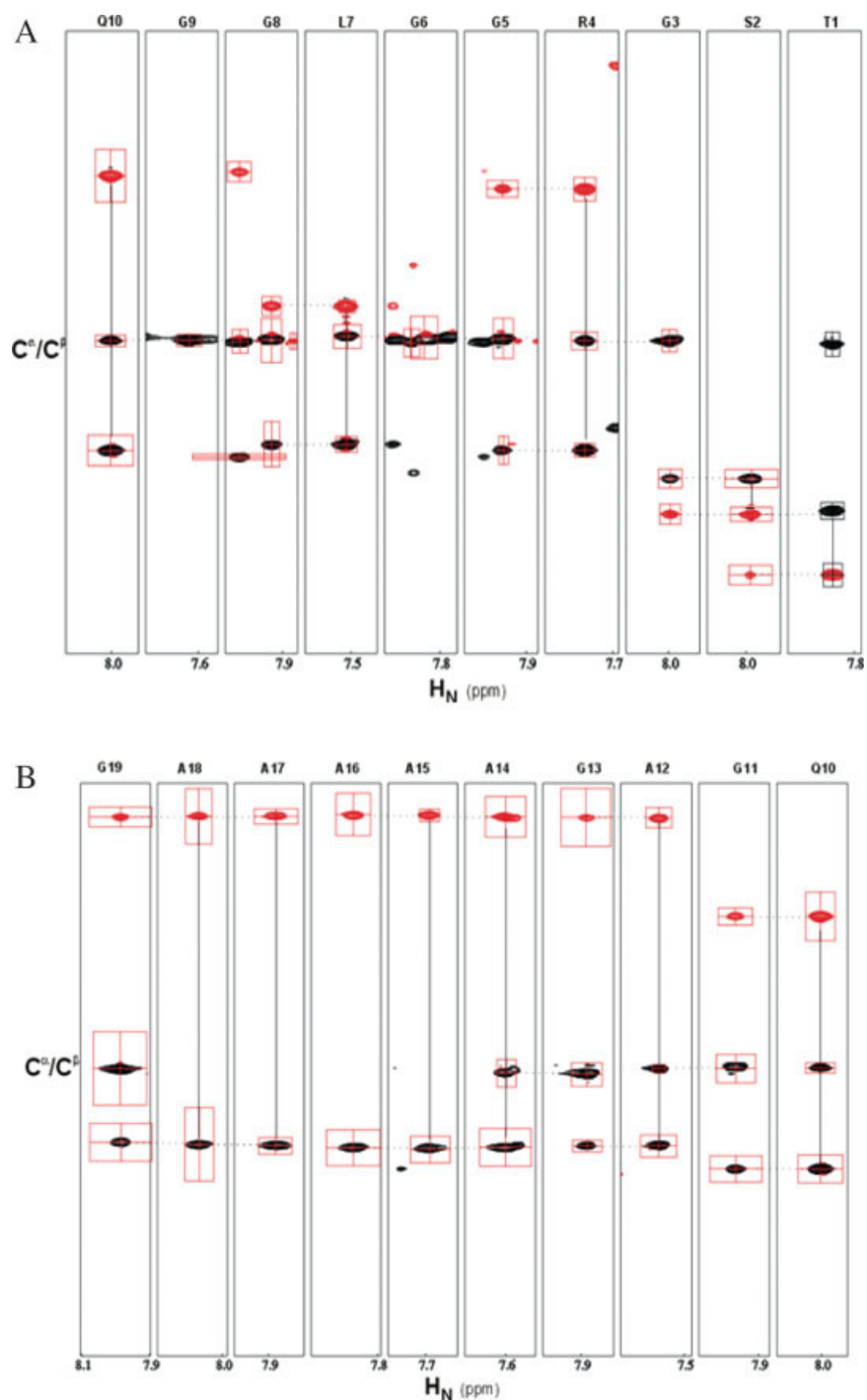


Figure 3. Strips of 3D HNCA/CBCA(CO)NH data sets utilized for sequential assignment of the backbone C_{α} resonances. Panels **A** and **B** displays the sequential connectivity for residues 10 to 1 and 19 to 10, respectively, of the 33-amino acid repeat. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Backbone secondary structure propensity

The backbone chemical shift and 3J coupling constant analyses are sensitive to protein secondary structure. In particular, deviation from random coil values provides information about the conformational propensities for the peptide backbone. Figure 4 shows the ^{13}C chemical shift deviation from random coil values for C_{α} , C_{β} , and CO of the CRGD protein.^{34,35} The C_{α} chemical shift of residues involved in α -helices are

shifted downfield of the random coil chemical shift and residues involved in β -sheet conformation are shifted upfield of the random coil value. The CO shifts are generally positive for α -helix and negative for β -sheet conformation. Taken together, it can be seen that the secondary structural content in the CRGD-silk mimetic, in 100% HFIP contains two α -helical segments. These are the polyalanine stretches spanning residues 12–19 and residues 26–32 of the repeat

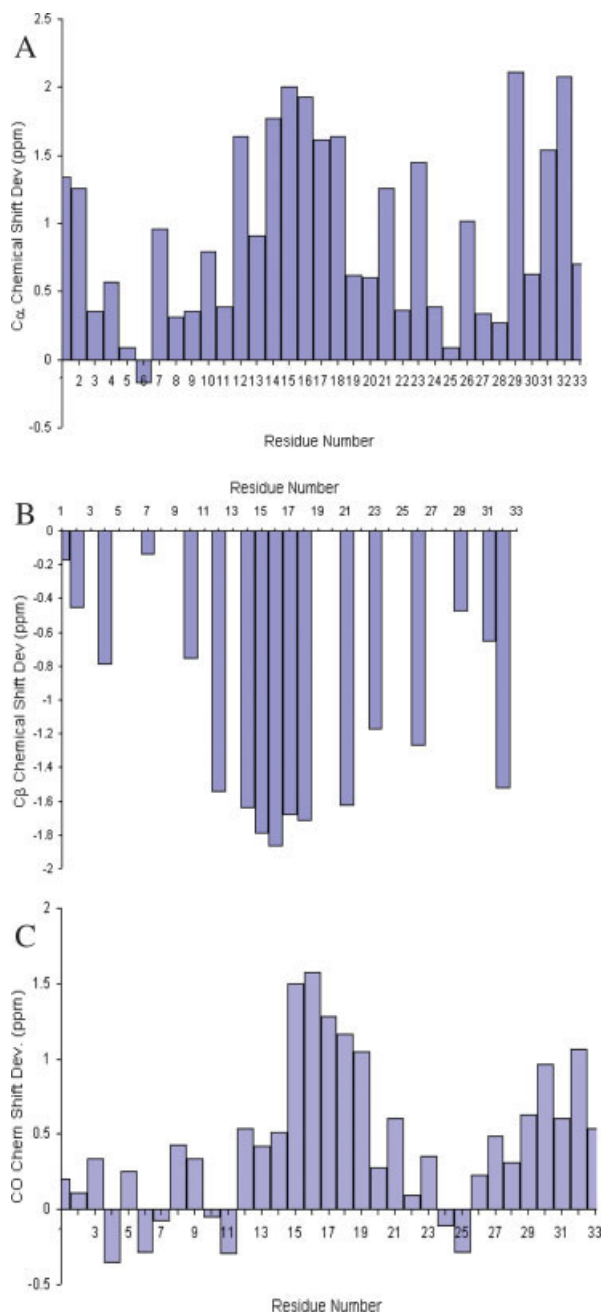


Figure 4. Deviations from random coil values of the (A) C α , (B) C β , and (C) CO chemical shifts. Well populated helical segments are positive with deviations greater than 1 ppm from coil values for C α and CO chemical shifts. The inverse relationship is observed for the C β chemical shifts. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

region, highlighted in Figure 1. Furthermore, the general convention for interpreting C α chemical shift deviation from random coil values is that values greater than 1 represents highly populated helical regions of a peptide chain, which indicates a high α -helical content for these amino acid sequences.³⁶ Unique interresidue interactions, which are characteristic of secondary structure is shown in Figure 5 for the 33 repeat

regions and were obtained from the 3D ^{15}N NOESY-HSQC and the 3D ^{13}C NOESY experiments. The H α HN ($i, i+3$) and H α HN ($i, i+4$) NOEs are clear indicators of helical stretches along the protein backbone, and are observed for residues in the previously identified helical segments. Whereas, the H α H β ($i, i+3$) NOEs are overlapped for the majority of the polyalanine sequence, these NOEs are observed for residues 26–29, 28–31 and 29–32.

Further evidence of the secondary structural propensity of the CRGD-silk mimetic in solution was determined from the 3J -coupling constants for the 33 residues observed in the 2D HSQC spectrum (Table I). The ϕ/φ dihedral angles provides residue specific evidence for the peptide backbone secondary structural propensity, which is related to the J -couplings by the Karplus equation. The 3J values for residues 12–21 and 25–30 are <5 Hz, which indicates helical propensity.³⁷ Residues 2 and 3 exhibits 3J -couplings consistent with a type 1 turn (4 and 7 Hz), and residues 22–23 (as well as, 30–31) are possibly in type 2 turns (4 and 5 Hz). The 3J -couplings associated with β -strand or half turns were not observed, that is consecutive 9.9 Hz and 4.9 Hz, respectively. These results are consistent with the chemical shift analysis that two highly helical stretches populate the polypeptide sequence. Although the N-terminus of the repeat motif displays no propensity to form an α -helix, the polyalanine stretches are highly α -helical. Generally, polyalanine sequences in folded peptides are populated by a mixture of α -helical or 3_{10} helical conformation in solution.^{38,39} It is expected that silk proteins sample a series of metastable backbone conformations during passage in the spinning ducts, before fiber formation. Therefore, it is likely that the secondary structural propensities observed for the repeat region in nonaqueous solvent represent an accessible conformational state that may be populated under appropriate conditions.

The molecular structure of natural silk fibers have been determined to be largely constructed of β -sheet secondary structure, arrayed in an ordered lattice.^{17,40}

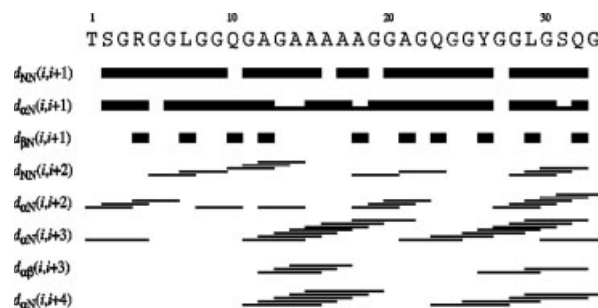


Figure 5. Summary of sequential and medium range backbone and sidechain NOE interactions. The sequential (residues i to $i+1$) and medium range (residues i to $i+3$ and i to $i+4$) NOEs are represented by the lines below the primary sequence and the line thickness is an indication of the relative intensities of the NOEs.

Table I. Phi-Value (ϕ) Analysis of the 33-Amino Acid Repeat

Residue number	3J -Coupling constants (Hz)
1	5.5
2	4.4
3	7.0
4	6.0
5	5.0
6	3.0
7	5.9
8	5.3
9	5.7
10	6.3
11	6.0
12	1.7
13	4.2
14	4.2
15	2.4
16	3.7
17	2.9
18	2.4
19	3.5
20	3.0
21	3.5
22	3.3
23	5.5
24	6.0
25	4.2
26	3.6
27	4.3
28	3.7
29	3.2
30	3.0
31	4.9
32	4.9
33	4.8

Regular secondary structure is associated with $^3J_{\text{HNH}\alpha}$ -couplings. $^3J_{\text{HNH}\alpha}$ -coupling constants of 3.9 Hz (α -helix), 4.2 Hz (3_{10} helix), 8.9 Hz (antiparallel β -sheet), and 9.7 Hz (parallel β -sheet).³⁷

In contrast, the molecular features of prespun silk have been reported to exhibit both ordered helical structures and disordered sequences, arrayed in either a cholesteric or nematic liquid crystalline state.^{2,41–43} The CRGD-silk mimetic is expected to exhibit similar molecular architecture as native silks, since the engineered primary sequence's design is based on the repetitive amino acid units found in *N. Clavipes* dragline silk, with a similar percentage of alanines and glycines.⁷ Recently, studies performed by Asakura and colleagues on the effects of solvent on the secondary structure of silk, reported that a 33-amino acid peptide, which encoded the 33-amino acid consensus sequence from spider dragline silk, adopted a helical conformation in nonaqueous solvents.⁴⁴ Additionally, a similar NMR spectroscopic analysis of the consensus sequence from the silkworm *Samia cynthia ricini*, revealed a highly α -helical structure in solution.⁴⁵ The observed secondary features of the 49,000 Da CRGD-silk mimetic are in agreement with the peptide-based studies and suggest that the determined secondary

structural elements predominates in at least the nonaqueous fractions reported for prespun silks.

Topology and fold of the CRGD-silk mimetic

Pulsed field gradient (PFG) translational diffusion measurements can provide quantitative and qualitative information about the relative size, shape and hydrodynamic radius of the molecular species in solution.^{46,47} To confirm that we were studying the full-length construct and not a 33-amino acid fragment, we performed PFG translational diffusion measurements with the CRGD 15-mer in 100% HFIP and 25°C. Figure 6 shows the signal attenuation of a selected peak from the NH region of the spectrum, which fits well with a two-component exponential. A fast component with D_s of 2.9×10^{-6} cm²/s and a slower component with D_s of 0.45×10^{-6} cm²/s were calculated. The fast component is comparable to measured values for a 2.5 kDa (19 amino acids) peptide in buffer. In contrast, the determined D_s for the slow component is disparate to the expected D_s of 0.75×10^{-6} cm²/s, obtained for the 66 kDa (570 amino acids) globular protein bovine serum albumin. Therefore, it is inferred that the molecular topology of the CRGD-silk mimetic in nonaqueous HFIP is nonspherical. This conclusion is also consistent with the 3D-NOESY HSQC data sets, which exhibited a paucity of long-range NOEs and is typical of an elongated or extended conformation. Figure 7 shows strip plots obtained from the ¹⁵N-NOESY HSQC spectrum of the silk mimetic, from residues 24–33 of the repeat region.

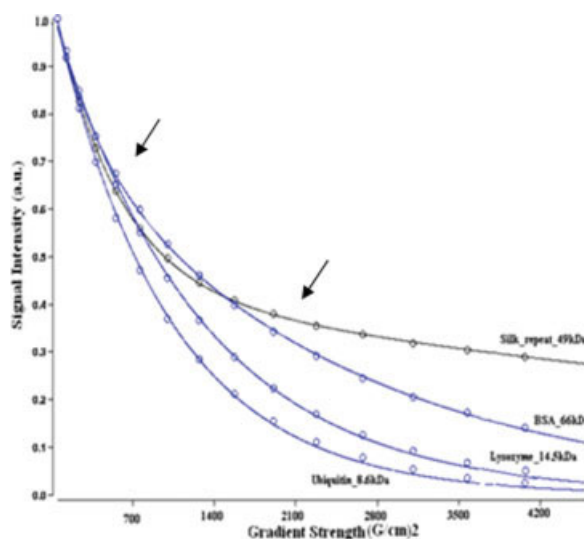


Figure 6. Pulse field gradient translational diffusion measurement of CRGD in 100% HFIP and 25°C. Shown is the attenuation of the signal intensity as a function of the square of the gradient strength for the 1.25 mM sample and known globular proteins. The arrows point to the fast and slow components of the decay curve. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

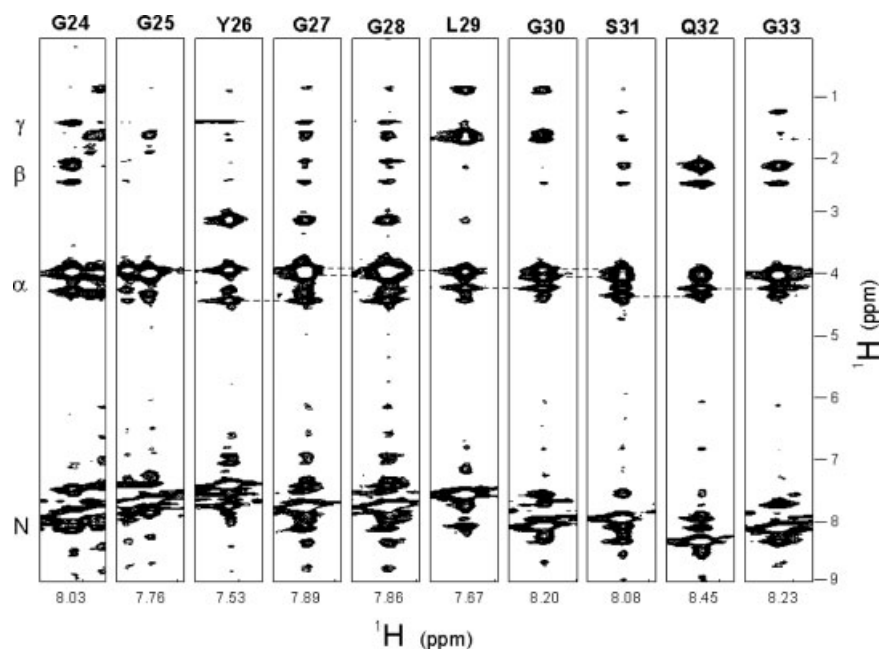


Figure 7. 3D ^{15}N -NOESY-HSQC strip-plot of the 33-amino acid repeat region from residue 24–33. The strips are obtained from the ^{15}N plane of the NOESY-HSQC and strong intraresidue and sequential αH_i to NH_{i+1} NOEs are shown with dotted lines. The weaker amide proton to side-chain NOEs, indicative of long-range interactions are labeled.

Interactions between adjacent repeat units are identified as weak NOEs, indicating that the distance between repeats are ≥ 5 Å NOE limits. Furthermore, the combination of degenerate NOEs and identical primary sequence within the repeat hampered the identification of unambiguous long-range interactions between repeat units. Nevertheless, this type of folding architecture, typically observed for Ankyrin repeat-motifs, has been characterized by both NMR spectroscopy and X-ray crystallography, and is suggested to function as scaffolds for macromolecule assembly.^{48,49} A recurring helix–turn–helix folding-motif, representing the 15 tandem repeats was implied from the overlapping resonances and sparsely populated 3D-NOESY spectrum. An elongated topology might be advantageous in facilitating the transition from α -helix/coil to β -sheet rich fibers.

The effects of water and peptide concentration on 2°-structural propensities

The authentic silk sequence is highly hydrophobic and exhibits limited solubility in water, yet is stored at high concentration in vivo.³³ Moreover, native spider silk undergoes supercontraction when exposed to water, suggesting restricted exposure to water in vivo.^{50,51} Supercontraction leads to a reduction of the silk fiber length by 50% and is accompanied by changes in protein's conformation and in the mechanical properties.^{51–53} To examine the interaction of H_2O with the silk-like mimetic, we used far UV-CD spectroscopy to monitor changes in secondary structure, since titration with H_2O resulted in protein precipita-

tion at NMR concentrations (require millimolar sample concentration). Figure 8 displays the superposition of CD spectra of the CRGD-silk mimetic dissolved in mixtures of HFIP/ H_2O (v/v) from 80% HFIP/20% H_2O to 10% HFIP/90% H_2O . The signal intensity of the 4 mg/mL sample appears to decrease with increasing H_2O content; however, the ratio of the signals at 222 nm and 208 nm, which is the hallmark of α -helicity, remains constant at a value of 0.7 for all measurements. Similarly, the signal at 220 nm, which corresponds to β -turn, undergoes minimal changes with increasing H_2O content. The spectra suggest that the average helical content is constant at the % H_2O used, and the loss in signal intensity with increasing H_2O content might be due to mild protein precipitation and not due to some other phenomenon such as supercontraction. Interestingly, Jelinski and coworkers determined that supercontraction of major ampullate dragline-silk of *N. clavipes*, corresponded to the consensus sequence, YGGLGS(N)QGAGR.⁵⁴ Additionally, it was suggested that substitution to an analogous, but less hydrophilic sequence might attenuate the contraction of silk in water. The engineered CRGD-silk sequence incorporates the YGGLGSQG repeat sequence, which corresponds to the indicated modification, and suggests a possible explanation for our observation. Furthermore, it is expected that the hydrophobic sequence will make minimal contact with H_2O , which suggest that the helical stretches resides within the hydrophobic 33-amino acid repeats.

Interestingly, in the spider, water is removed from the spinning solution before fiber formation.⁵⁵ It

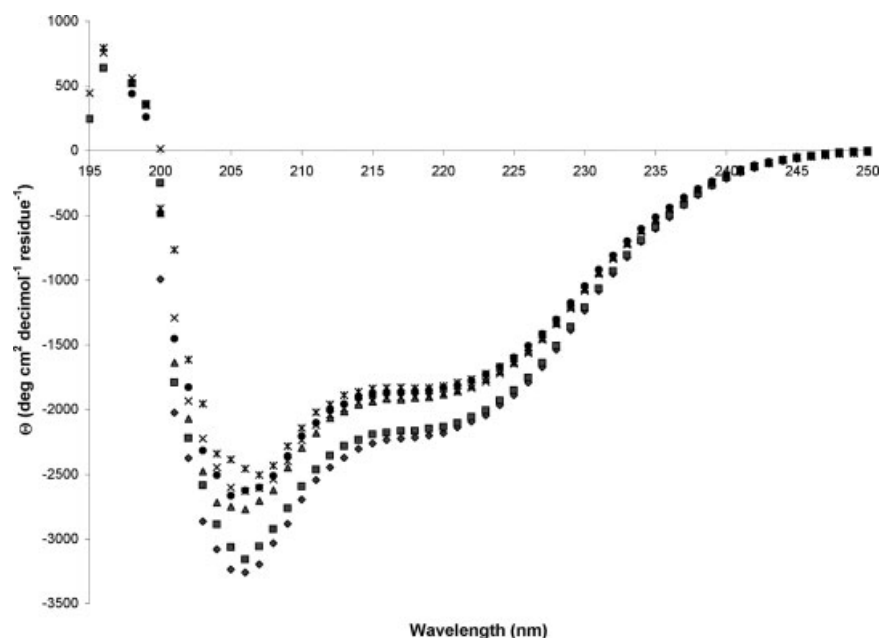


Figure 8. Circular dichroism spectra of CRGD in HFIP at 25°C. The effect of increasing H₂O content on the secondary structure of the CRGD peptide at high concentration is shown. ◆ 80% HFIP/20% H₂O, ■ 50% HFIP/50% H₂O, ▲ 40% HFIP/60% H₂O, ● 30% HFIP/70% H₂O, * 20% HFIP/80% H₂O, x 10% HFIP/90% H₂O. Each spectrum was collected at a constant protein concentration of 4 mg/mL.

might be that water functions in some other capacity than as a solubilizing agent, occupying a separate phase and not penetrating the core of the hydrophobic semicrystalline blocks. Alternatively, water may have a dual functionality in silk processing, serving as a coacervate at high protein concentration and plasticizer at lower concentrations.^{56,57} In this way, the highly concentrated liquid crystalline spinning solution, which has been suggested to be as much as 50% w/v might be streamlined as it traverse the spinning duct and then softened before draw-down, where a plastic conformation could facilitate the conversion from α -helix/coil to β -sheet formation. This model might be advantageous in facilitating facile mechanical and chemical control of the silk microenvironment in the prespun state, that is, regulation of temperature, pressure transduction, ionic uptake, and so forth.

Conclusions

We have utilized a spider-silk mimetic, which encodes similar percentages of glycines and alanines (42% and 21%, respectively), as expressed in dragline silks, to examine the secondary structural propensities and the conformation of highly concentrated silk-like proteins in a nonaqueous solution. We find a predominance of α -helix encompassing the polyalanine sequences and the GGLGSQ sequences within the semicrystalline repeats. In addition, the silk mimetic polypeptide chain folds into an elongated conformation. The helical population was unaltered with increasing water/solvent mixtures, indicating that this structural motif was sequestered from water.

Methods

Chemical and reagents

Hexafluoro-isopropanol (HFIP) and deuterated HFIP were purchased from Cambridge isotopes. Double-distilled deionized water was used in the titration experiments.

Expression and purification

The 48,558 Da CRGD 15-mer was subcloned in a pET30a vector and ¹⁵N/¹³C isotope-labeled protein was obtained from M9 minimal media growth.^{5,7} Purified, isotope-enriched, His-tagged protein was obtained by a one-step process from a Ni⁺-column.

Circular dichroism

Far-UV CD studies were carried out on an Aviv spectropolarimeter, model 202, equipped with an automated temperature controller (Aviv associates, Lakewood, NJ). The peptide concentration was 4 mg/mL, and each data set was collected in 400 μ L of a clear sample, containing the CRGD silk mimetic solubilized in 100% to 10% HFIP/H₂O mixtures (v/v) at 25°C. Spectral measurements were taken in a 1-mm path-length quartz cell from 250 to 190 nm wavelength; with a 1 nm bandwidth and three scans per spectrum. Each data point represents the average of three experiments.

NMR spectroscopy

The NMR sample contained 1.25 mM CRGD protein solubilized in 100% deuterated HFIP. NMR measurements were performed on a Bruker DRX 600 or

Varian INOVA 600, each equipped with three RF channels, at 25°C. The Bruker 600 is equipped with a three-axis PFG triple resonance probe, and the Inova 600 is equipped with a z-axis PFG triple resonance cryoprobe. Initially, a high-resolution 2D ¹H-¹⁵N HSQC spectrum was collected with 2048 data points in F1 and 8192 data points in the F2 dimension. The assignment of all ¹H, ¹³C, and ¹⁵N chemical shifts were made using through-bond correlations along the backbone and side chains using the following pulse sequences: 3D HNCO, 3D HN(CA)CO, 3D HNCA, 3D HN(CO)CA, 3D HNCACB, 3D CBCA(CO)NH, 3D HCCH-TOCSY, 3D HCCH-COSY, 3D ¹⁵N TOCSY-HSQC, 3D ¹⁵N NOESY-HSQC (50–100 ms mixing times), and 3D ¹³C-edited NOESY-HSQC (50 ms mixing time).^{58–67} Backbone phi dihedral angle constraints were obtained from ³J (HN, H α) scalar couplings measured using a 3D HNCA-E.COSY experiment.⁶⁸ A reference 2D ¹H-¹⁵N HSQC spectrum was collected before and after each set of multidimensional NMR experiments. All chemical shifts were referenced to the tetramethylsilane proton signal at 0 ppm. All NMR spectra were processed using the software program NMRPipe and analyzed using the program NMRVIEW.^{69,70} PFG translational diffusion measures were carried out at 25°C in 100% HFIP.⁴⁶ The equation mentioned later correlates the diffusion coefficient with the signal attenuation as a function of increasing gradient strength:

$$I = I_0 \exp -[(\gamma_H)^2 \delta^2 (\Delta - \delta/3) D_s G^2]$$

where G is the gradient strength, D_s is the diffusion coefficient, Δ is the time between pulses, δ is the length of the pulse, and γ_H is the proton gyromagnetic ratio. The diffusion coefficient is calibrated relative to the published value of 1.08×10^{-6} cm²/s, which was obtained for the globular protein lysozyme in phosphate buffer, at 25°C.⁷¹

Acknowledgments

The authors are grateful to Dr. Yujia Xu for use of the circular dichroism spectropolarimeter in the department of chemistry at Hunter College CUNY. The work at Hunter College was supported by a grant from the Air Force Office of Scientific Research and an infrastructure grant from the National Institutes of Health (NCRR - RR03037).

References

- Gosline JM, Guerette PA, Ortlepp CS, Savage KN (1999) The mechanical design of spider silks: from fibroin sequence to mechanical function. *J Exp Biol* 202: 3295–3303.
- Vollrath F, Knight DP (2001) Liquid crystalline spinning of spider silk. *Nature* 410:541–548.
- Sofia S, McCarthy MB, Gronowicz G, Kaplan DL (2001) Functionalized silk-based biomaterials for bone formation. *J Biomed Mater Res* 54:139–148.
- Wong Po Foo C, Kaplan DL (2002) Genetic engineering of fibrous proteins: spider dragline silk and collagen. *Adv Drug Deliv Rev* 54:1131–1143.
- Bini E, Foo CW, Huang J, Karageorgiou V, Kitchel B, Kaplan DL (2006) RGD-functionalized bioengineered spider dragline silk biomaterial. *Biomacromolecules* 7:3139–3145.
- Wang Y, Kim HJ, Vunjak-Novakovic G, Kaplan DL (2006) Stem cell-based tissue engineering with silk biomaterials. *Biomaterials* 27:6064–6082.
- Wong Po Foo C, Patwardhan SV, Belton DJ, Kitchel B, Anastasiades D, Huang J, Naik RR, Perry CC, Kaplan DL (2006) Novel nanocomposites from spider silk-silica fusion (chimeric) proteins. *Proc Natl Acad Sci USA* 103: 9428–9433.
- Zhou CZ, Confalonieri F, Medina N, Zivanovic Y, Esnault C, Yang T, Jacquet M, Janin J, Duguet M, Perasso R, Li Z-G (2000) Fine organization of *Bombyx mori* fibroin heavy chain gene. *Nucleic Acids Res* 28:2413–2419.
- Valluzzi R, Winkler S, Wilson D, Kaplan DL (2002) Silk: molecular organization and control of assembly. *Philos Trans R Soc Lond B Biol Sci* 357:165–167.
- Bini E, Knight DP, Kaplan DL (2004) Mapping domain structures in silks from insects and spiders related to protein assembly. *J Mol Biol* 335:27–40.
- Rising A, Nimmervoll H, Grip S, Fernandez-Arias A, Storckenfeldt E, Knight DP, Vollrath F, Engstrom W (2005) Spider silk proteins—mechanical property and gene sequence. *Zool J Linn Soc* 22:273–281.
- Vollrath F, Knight DP (1999) Structure and function of the silk production pathway in the spider *Nephila edulis*. *Int J Biol Macromol* 24:243–249.
- Meinel L, Hofmann S, Karageorgiou V, Zichner L, Langer R, Kaplan D, Vunjak-Novakovic G (2004) Engineering cartilage-like tissue using human mesenchymal stem cells and silk protein scaffolds. *Biotechnol Bioeng* 88: 379–391.
- Kardestuncer T, McCarthy MB, Karageorgiou V, Kaplan D, Gronowicz G (2006) RGD-tethered silk substrate stimulates the differentiation of human tendon cells. *Clin Orthop Relat Res* 448:234–239.
- Knight DP, Vollrath F (2001) Changes in element composition along the spinning duct in a *Nephila* spider. *Naturwissenschaften* 88:179–182.
- Dicko C, Vollrath F, Kenney JM (2004) Spider silk protein refolding is controlled by changing pH. *Biomacromolecules* 5:704–710.
- Simmons AH, Michal CA, Jelinski LW (1996) Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science* 271:84–87.
- Kenney JM, Knight D, wire MJ, Vollrath F (2002) Amyloidogenic nature of spider silk. *Eur J Biochem* 269: 4159–4163.
- Azuma M, Ohta Y (1998) Changes in H⁺-translocating vacuolar-type ATPase in the anterior silk gland cell of *Bombyx mori* during metamorphosis. *J Exp Biol* 201:479–486.
- Zong XH, Zhou P, Shao ZZ, Chen SM, Chen X, Hu BW, Deng F, Yao WH (2004) Effect of pH and copper(II) on the conformation transitions of silk fibroin based on EPR, NMR, and Raman spectroscopy. *Biochemistry* 43: 11932–11941.
- Holland C, Terry AE, Porter D, Vollrath F (2006) Comparing the rheology of native spider and silkworm spinning dope. *Nat Mater* 5:870–874.
- Seidel A, Liivak O, Jelinski LW (1998) Artificial spinning of spider silk. *Macromolecules* 31:6733–6736.
- Trabacchi KA, Yager P (1998) Comparative structural characterization of naturally- and synthetically-spun fibers of *Bombyx mori* fibroin. *Macromolecules* 31:462–471.
- Yao J, Masuda H, Zhao C, Asakura T (2002) Artificial spinning and characterization of silk fiber from *Bombyx*

- mori* silk fibroin in hexafluoroacetone hydrate. *Macromolecules* 35:6–9.
25. Zhao C, Yao J, Masuda H, Kishore R, Asakura T (2003) Structural characterization and artificial fiber formation of *Bombyx mori* silk fibroin in hexafluoro-iso-propanol solvent system. *Biopolymers* 69:253–259.
 26. Ha SW, Asakura T, Kishore R (2006) Distinctive influence of two hexafluoro solvents on the structural stabilization of *Bombyx mori* silk fibroin protein and its derived peptides: ¹³C NMR and CD studies. *Biomacromolecules* 7:18–23.
 27. Luo P, Baldwin RL (1997) Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* 36:8413–8421.
 28. Walgers R, Lee TC, Cammers-Goodwin A (1998) An indirect chaotropic mechanism for the stabilization of helix conformation of peptides in aqueous trifluoroethanol and hexafluoro-2-propanol. *J Am Chem Soc* 120:5073–5079.
 29. Andersen NH, Dyer RB, Fesinmeyer RM, Gai F, Liu Z, Neidigh JW, Tong H (1999) Effect of hexafluoroisopropanol on the thermodynamics of peptide secondary structure formation. *J Am Chem Soc* 121:9879–9880.
 30. Hong DP, Hoshino M, Kuboi R, Goto Y (1999) Clustering of fluorine-substituted alcohols as a factor responsible for their marked effects on proteins and peptides. *J Am Chem Soc* 121:8427–8433.
 31. Ha SW, Gracz HS, Tonelli AE, Hudson SM (2005) Structural study of irregular amino acid sequences in the heavy chain of *Bombyx mori* silk fibroin. *Biomacromolecules* 6:2563–2569.
 32. Grace CRR, Perrin MH, DiGruccio MR, Miller CL, Rivier JE, Vale WW, Riek R (2004) NMR structure and peptide hormone binding site of the first extracellular domain of a type B1 G protein-coupled receptor. *PNAS* 101:12836–12841.
 33. Hijirida DH, Do KG, Michal C, Wong S, Zax D, Jelinski LW (1996) ¹³C NMR of *Nephila clavipes* major ampullate silk gland. *Biophys J* 71:3442–3447.
 34. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD (1995) ¹H, ¹³C and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR* 5:67–81.
 35. Schwarzingler S, Kroon GJ, Foss TR, Wright PE, Dyson HJ (2000) Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *J Biomol NMR* 18:43–48.
 36. Spera S, Bax A (1991) Empirical correlation between protein backbone conformation and C.alpha and C.beta ¹³C nuclear magnetic resonance chemical shifts. *J Am Chem Soc* 113:5490–5492.
 37. Wuthrich K (1986) *NMR of proteins and nucleic acids*. New York: John Wiley & Sons.
 38. Marqusee S, Robbins VH, Baldwin RL (1989) Unusually stable helix formation in short alanine-based peptides. *PNAS* 86:5286–5290.
 39. Fiori WR, Miick SM, Millhauser GL (1993). Increasing sequence length favors alpha-helix over 310-helix in alanine-based peptides: evidence for a length-dependent structural transition. *Biochemistry* 32:11957–11962.
 40. Marsh RE, Corey RB, Pauling L (1955) An investigation of the structure of silk fibroin. *Biochim Biophys Acta* 16:1–34.
 41. Knight DP, Vollrath F (1999) Liquid crystals and flow elongation in a spider's silk production line. *Proc R Soc B Biol Sci* 266:519–523.
 42. Knight DP, Vollrath F (2002) Biological liquid crystal elastomers. *Philos Trans R Soc Lond B Biol Sci* 357:155–163.
 43. Dicko C, Kenney JM, Knight D, Vollrath F (2004) Transition to a beta-sheet-rich structure in spidroin in vitro: the effects of pH and cations. *Biochemistry* 43:14080–14087.
 44. Yang M, Nakazawa Y, Yamauchi K, Knight D, Asakura T (2005) Structure of model peptides based on *Nephila clavipes* dragline silk spidroin (MaSp1) studied by ¹³C cross polarization/magic angle spinning NMR. *Biomacromolecules* 6:3220–3226.
 45. Yao J, Nakazawa Y, Asakura T (2004) Structures of *Bombyx mori* and *Samia cynthia* ricini silk fibroins studied with solid-state NMR. *Biomacromolecules* 5:680–688.
 46. Altieri AS, Hinton DP, Byrd RA (1995) Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements. *J Am Chem Soc* 7566–7567.
 47. Wilkins DK, Grimshaw SB, Recevour V, Dobson CM, Jones JA, Smith LJ (1999). Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR Techniques. *Biochemistry* 38:Biochemistry. 16424–16431.
 48. Main ERG, Xiong Y, Cocco MJ, D'Andrea L, Regan L (2003) Design of stable α -helical arrays from an idealized TPR motif. *Structure* 11:497–508.
 49. Saito S, Aizawa T, Kawaguchi K, Yamaki T, Matsumoto D, Kamiya M, Kumaki Y, Mizuguchi M, Takiya S, Demura M, Kawano K (2007) Structural approach to a novel tandem repeat DNA-binding domain, STPR, by CD and NMR. *Biochemistry* 46:1703–1713.
 50. Gosline JM, Denny MW, DeMont ME (1984) Spider silk as rubber. *Nature* 309:551–552.
 51. Work RW (1985) Viscoelastic behaviour and wet supercontraction of major ampullate silk fibres of certain orb-weaving spiders (Araneae). *J Exp Biol* 118:379–404.
 52. Shao Z, Young RJ, Vollrath F (1999) The effect of solvents on spider silk studied by mechanical testing and single-fibre Raman spectroscopy. *Int J Biol Macromol* 24:295–300.
 53. Holland GP, Lewis RV, Yarger JL (2004) WISE NMR characterization of nanoscale heterogeneity and mobility in supercontracted *Nephila clavipes* spider dragline silk. *J Am Chem Soc* 126:5867–5872.
 54. Yang Z, Liivak O, Seidel A, LaVerde G, Zax DB, Jelinski LW (2000) Supercontraction and backbone dynamics in spider silk: ¹³C and ²H NMR studies. *J Am Chem Soc* 122:9019–9025.
 55. Tillinghast EK, Chase SF, Townley MA (1984) Water extraction by the major ampullate duct during silk formation in the spider, *Argiope aurantia* Lucas. *J Insect Physiol* 30:591–596.
 56. Vollrath F, Edmonds DT (1989) Modulation of the mechanical properties of spider silk by coating with water. *Nature* 340:305–307.
 57. Jin HJ, Kaplan DL (2003) Mechanism of silk processing in insects and spiders. *Nature* 424:1057–1061.
 58. Kay LE, Ikura M, Tschudin R, Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson* 89:496–514.
 59. Ikura M, Kay LE, Bax A (1991) Improved three-dimensional ¹H-¹³C-¹H correlation spectroscopy of a ¹³C-labeled protein using constant-time evolution. *J Biomol NMR* 1:299–304.
 60. Grzesiek S, Bax A (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc* 114:6291–6293.

61. Grzesiek S, Bax A (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson* 96:432–440.
62. Grzesiek S, Anglister J, Bax A (1993) Correlation of backbone amide and aliphatic side-chain resonances in $^{13}\text{C}/^{15}\text{N}$ -enriched proteins by isotropic mixing of ^{13}C magnetization. *J Magn Reson B* 101:114–119.
63. Kay LE, Xu GY, Singer AU, Muhandiram DR, Formankay JD (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^1H and ^{13}C correlations in H_2O samples of proteins. *J Magn Reson B* 101:333–337.
64. Muhandiram DR, Farrow NA, Xu GY, Smallcombe SH, Kay LE (1993) A gradient ^{13}C NOESY-HSQC experiment for recording NOESY spectra of ^{13}C -labeled proteins dissolved in H_2O . *J Magn Reson B* 102:317–321.
65. Wittekind M, Mueller L (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J Magn Reson B* 101:201–205.
66. Zhang O, Kay LE, Olivier JP, Forman-Kay JD (1994) Backbone ^1H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J Biomol NMR* 4:845–858.
67. Engelke J, Rüterjans H (1995) Sequential protein backbone resonance assignments using an improved 3D-HN(CA)CO pulse scheme. *J Magn Reson B* 109:318–322.
68. Weisemann R, Rüterjans H, Schwalbe H, Schleucher J, Bermel W, Griesinger C (1994) Determination of $\text{HN}, \text{H}\alpha$ and HN, C' coupling constants in ^{13}C , ^{15}N -labeled proteins. *J Biomol NMR* 4:231–240.
69. Johnson BA, Blevins RA (1994) NMR view—a computer-program for the visualization and analysis of NMR data. *J Biomol NMR* 4:603–614.
70. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
71. Dubin SB, Clark NA, Benedek GB (1971) Measurement of the rotational diffusion coefficient of lysozyme by depolarized light scattering: configuration of lysozyme in solution. *J Chem Phys* 54:5158–5164.