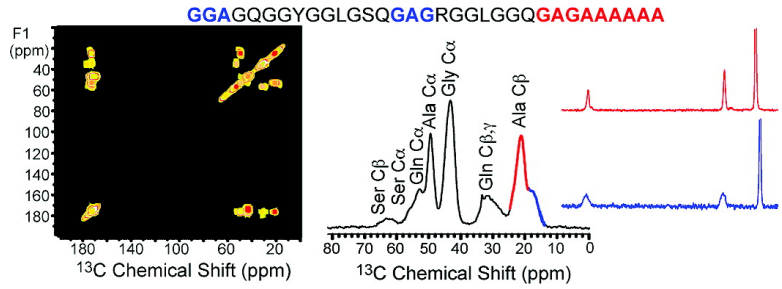
Article

Determining Secondary Structure in Spider Dragline Silk by Carbon#Carbon Correlation Solid-State NMR Spectroscopy

Gregory P. Holland, Melinda S. Creager, Janelle E. Jenkins, Randolph V. Lewis, and Jeffery L. Yarger

J. Am. Chem. Soc., **2008**, 130 (30), 9871-9877 • DOI: 10.1021/ja8021208 • Publication Date (Web): 02 July 2008



More About This Article

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

- Supporting Information
- 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 07/02/2008

Determining Secondary Structure in Spider Dragline Silk by Carbon−Carbon Correlation Solid-State NMR Spectroscopy

Gregory P. Holland,* Melinda S. Creager, Janelle E. Jenkins, Randolph V. Lewis, and Jeffery L. Yarger*

Magnetic Resonance Research Center, Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604, and Department of Molecular Biology, University of Wyoming, Laramie, Wyoming 82071

Received March 21, 2008; E-mail: Greg.holland@asu.edu; jyarger@gmail.com

Abstract: Two-dimensional (2D) $^{13}C-^{13}C$ NMR correlation spectra were collected on ^{13}C-enriched dragline silk fibers produced from *Nephila clavipes* spiders. The 2D NMR spectra were acquired under fast magic-angle spinning (MAS) and dipolar-assisted rotational resonance (DARR) recoupling to enhance magnetization transfer between ^{13}C spins. Spectra obtained with short (150 ms) recoupling periods were utilized to extract distinct chemical shifts for all carbon resonances of each labeled amino acid in the silk spectra, resulting in a complete resonance assignment. The NMR results presented here permit extraction of the precise chemical shift of the carbonyl environment for each ^{13}C-labeled amino acid in spider silk for the first time. Spectra collected with longer recoupling periods (1 s) were implemented to detect intermolecular magnetization exchange between neighboring amino acids. This information is used to ascribe NMR resonances to the specific repetitive amino acid motifs prevalent in spider silk proteins. These results indicate that glycine and alanine are both present in two distinct structural environments: a disordered 3_1-helical conformation and an ordered β-sheet structure. The former can be ascribed to the Gly-Gly-Ala motif while the latter is assigned to the poly(Ala) and poly(Gly-Ala) domains.

Introduction

Spider silk spun from the major ampullate gland (dragline silk) possesses incredible mechanical properties, with a toughness that exceeds Kevlar and an extensibility that is double that of high-tenacity nylon.[1–4] Spiders are capable of producing this material from an aqueous spinning dope at ambient pressures and temperatures. The unique combination of mechanical properties coupled with the benign way in which spiders produce dragline silk makes mimicking this process in the laboratory very appealing, with countless applications imaginable. Significant progress has been made in determining the primary amino acid sequences of the proteins that comprise spider silk[3,5–7] and there has even been some success in producing large amounts of the protein with the tools of biotechnology.[8] However, producing a fiber synthetically that matches native spider silk remains elusive. A more thorough understanding of the secondary, tertiary, and hierarchical structure of spider silk is required to better understand the structure−function relationship and guide the production of a biomaterial that matches the properties of natural spider silk.

Spider dragline silk is composed of two proteins,[5,6,9] major ampullate spidroin 1 (MaSp1) and major ampullate spidroin 2 (MaSp2), and is approximately 15% crystalline.[10,11] The crystalline domains have been well characterized via X-ray diffraction[10–14] (XRD), transmission electron microscopy[15–20] (TEM), and solid-state NMR[21,22] methods. These crystalline regions are aligned parallel to the fiber axis[12] and are composed primarily of polyalanine [poly(Ala)] that forms antiparallel β-sheet structures similar to silkworm silk.[23–25] The size of the crystalline domains

(1) Gosline, J. M.; DeMont, M. E.; Denny, M. W. *Endeavour* **1986**, *10*, 37–43.
(2) Gosline, J. M.; Denny, M. W.; DeMont, M. E. *Nature* **1984**, *309*, 551–552.
(3) Lewis, R. V. *Chem. Rev.* **2006**, *106*, 3762–3774.
(4) Lucas, F. *Discovery London* **1964**, *25*, 1–7.
(5) Xu, M.; Lewis, R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7120–7124.
(6) Hinman, M.; Lewis, R. V. *J. Biol. Chem.* **1992**, *267*, 19320–19324.
(7) Ayoub, N. A.; Garb, J. E.; Tinghitella, R. M.; Collin, M. A.; Hayashi, C. Y. *PLoS ONE* **2007**, *2*, e514.
(8) Lazaris, A.; Arcidiacono, S.; Huang, Y.; Zhou, J.; Duguay, F.; Chretien, N.; Welsh, E.; Soares, J.; Karatzas, C. *Science* **2002**, *295*, 472–476.
(9) Lewis, R. V. *Acc. Chem. Res.* **1992**, *25*, 392–298.
(10) Yang, Z.; Grubb, D.; Jelinski, L. W. *Macromolecules* **1997**, *30*, 8254–8261.
(11) Grubb, D. T.; Jelinski, L. W. *Macromolecules* **1997**, *30*, 2860–2867.
(12) Parkhe, A. D.; Seeley, S. K.; Gardner, K.; Thompson, L.; Lewis, R. V. *J. Mol. Recognit.* **1997**, *10*, 1–6.
(13) Riekel, C.; Vollrath, F. *Int. J. Biol. Macromol.* **2001**, *29*, 203–210.
(14) Riekel, C.; Bränden, C.; Craig, C.; Ferrero, C.; Heidelbach, F.; Müller, M. *Int. J. Biol. Macromol.* **1999**, *24*, 179–186.
(15) Thiel, B.; Kunkel, D.; Viney, C. *Biopolymers* **1994**, *34*, 1089–1097.
(16) Thiel, B. L.; Viney, C. *J. Microsc.* **1997**, *185*, 179–187.
(17) Thiel, B.; Viney, C. *MRS Bull.* **1995**, *20*, 52–56.
(18) Thiel, B. L.; Guess, K.; Viney, C. *Biopolymers* **1997**, *41*, 703–719.
(19) Frische, S.; Maunsbach, A.; Vollrath, F. *J. Microsc.* **1998**, *189*, 64–70.
(20) Barghout, J. Y. J.; Thiel, B. L.; Viney, C. *Int. J. Biol. Macromol.* **1999**, *24*, 211–217.
(21) Simmons, A.; Ray, E.; Jelinski, L. W. *Macromolecules* **1994**, *27*, 5235–5237.
(22) Simmons, A. H.; Michal, C. A.; Jelinski, L. W. *Science* **1996**, *271*, 84–87.
(23) Marsh, R. E.; Corey, R. B.; Pauling, L. *Biochim. Biophys. Acta* **1955**, *16*, 1–34.
(24) Fraser, R. D. B.; MacRae, T. P.; Stewart, F. H. C.; Suzuki, E. *J. Mol. Biol.* **1965**, *11*, 706–712.
(25) Lotz, B.; Cesari, F. C. *Biochimie* **1979**, *61*, 205–214.

was determined from wide-angle XRD[10,11] to be 2 × 5 × 7 nm while a ^{1}H NMR spin diffusion[26] analysis on wetted silk estimated average domains that were 6 ± 2 nm. Although the structure of the crystalline domains has been well characterized in spider silk, the structure of the remaining fraction of the material (~85%) is relatively unsolved due to the high degree of disorder and structural heterogeneity.

A few solid-state NMR studies have begun to address the molecular structure of the disordered regions of the silk.[27–30] Three of these studies rely on selectively ^{13}C-labeling the backbone carbonyl of glycine (that is believed to be present in high abundance in the disordered region) and performing a backbone torsion angle measurement with double-quantum single-quantum correlation experiments for static samples[28] (DOQSY) or direction exchange with correlation for orientation-distribution evaluation and reconstruction[28,29] (DECODER). These experiments have provided some convincing preliminary evidence that a fraction of glycine is present in an approximate 3-fold helical conformation similar to the type II 3_1-helical structure of poly(Gly).[31,32] It was also shown that this glycine helical domain has a high degree of orientation with respect to the fiber axis.[28] One of these studies determined that a fraction of Gly was present in the crystalline β-sheet region, although the percentage was not reported.[28] In a separate study, a specific ^{13}C/^{15}N isotopic labeling scheme was used to measure a distance with the rotational-echo double-resonance (REDOR) method that determined the Leu-Gly-Xaa-Gln motif was in a type-I β-turn structure.[30]

The ^{13}C CP-MAS NMR spectrum of spider silk is extensively broadened due to the heterogeneous nature of spider silk and the conformation dependence of the ^{13}C isotropic chemical shift.[33–35] This makes unambiguously assigning the NMR spectrum difficult, and extraction of distinct chemical shifts for each amino acid site is incomplete. The ^{13}C chemical shift can be used to determine secondary structure in fibrous proteins; however, in the case of silks the NMR spectrum is heterogeneously broadened where resonances from multiple amino acids and/or multiple conformations are overlapped. This is particularly obvious for the carbonyl region where one asymmetric broad resonance is observed that spans >10 ppm, making extraction of the carbonyl chemical shift for any of the amino acids intractable (see Figure 1). It should be mentioned that there could be a minor line broadening contribution from ^{13}C−^{13}C J coupling because the sample is ^{13}C-enriched; however, this component is negligible when one considers that the magnitude of the J coupling is 35−55 Hz and the observed line widths are in the kilohertz range.

In the present study, ^{13}C−^{13}C NMR correlation spectra are collected on a moderately (16−45%) ^{13}C-enriched spider silk

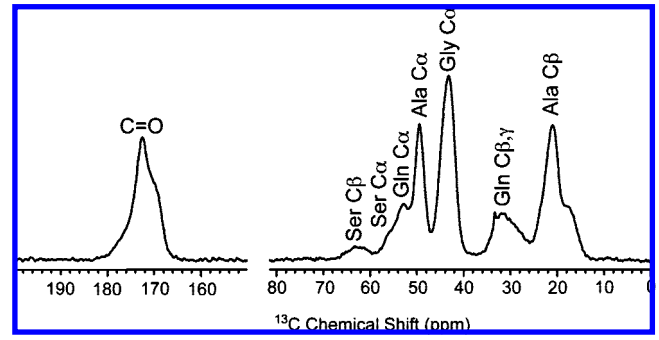


Figure 1. ^{13}C CP-MAS NMR spectrum of ^{13}C-enriched *N. clavipes* spider dragline silk. Spectrum was collected with 40 kHz MAS and 1 ms CP contact time.

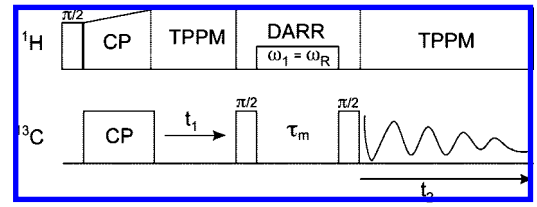


Figure 2. NMR pulse sequence for collecting ^{13}C−^{13}C correlation spectra with DARR recoupling. Experimental details are described under Materials and Methods.

sample. These experiments utilize ^{13}C−^{1}H dipolar-assisted rotational resonance[36] (DARR) recoupling to enhance magnetization exchange during a mixing period in the NMR pulse sequence (see Figure 2). This establishes spatial connectivity within each amino acid, permitting the chemical shift of each residue site to be obtained for the first time from the 2D spectrum. Not only do these experiments provide a complete spectral assignment, but at longer DARR mixing times, minor intermolecular contacts between amino acids are obtained that can be applied to determine the motif that is contributing to a specific resonance. This information coupled with the conformation dependence of the ^{13}C isotropic chemical shift permits a secondary structural characterization of a given repetitive motif in silk. This solid-state NMR approach in heterogeneous solids like silks has been relatively unexplored to date,[37] and the results presented herein provide new insight into the secondary structure of spider silk fibers.

Materials and Methods

Materials. Major ampullate silk was drawn from adult female *Nephila clavipes* spiders at a rate of 2 cm/s similar to the method described by Work and Emerson.[38] During forcible silking, the spiders were carefully monitored under a dissection microscope to ensure that only major ampullate silk was collected. The spiders were forcibly silked on average every other day. During a silking session, each spider was hand-fed one cricket in addition to a ^{13}C-enriched MEM5550 solution (Sigma−Aldrich, St. Louis, MO). This feeding and silk collection protocol was repeated every other day for 30 consecutive days. The fibers used in this study were gathered from silkings 11 and 12. The silk does not show uniform ^{13}C enrichment for each amino acid in spider silk fibers. Rather, the silk displays selective ^{13}C enrichment at Gly, Ala, Gln, and Ser.

(26) Holland, G. P.; Lewis, R. V.; Yarger, J. L. *J. Am. Chem. Soc.* **2004**, *126*, 5867–5872.
(27) Kummerlen, J.; van Beek, J. D.; Vollrath, F.; Meier, B. H. *Macromolecules* **1996**, *29*, 2920–2928.
(28) van Beek, J. D.; Hess, S.; Vollrath, F.; Meier, B. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10266–10271.
(29) Eles, P.; Michal, C. *Biomacromolecules* **2004**, *5*, 661–665.
(30) Michal, C. A.; Jelinski, L. W. *J. Biomol. NMR* **1998**, *12*, 231–241.
(31) Crick, F. H. C.; Rich, A. *Nature* **1955**, *176*, 780–781.
(32) Ramachandran, G. N.; Sasisekharan, V.; Ramakrishnan, C. *Biochim. Biophys. Acta* **1966**, *112*, 168–170.
(33) Saitô, H. *Magn. Reson. Chem.* **1986**, *24*, 835–852.
(34) Saitô, H.; Tabeta, R.; Asakura, T.; Iwanaga, Y.; Shoji, A.; Ozaki, T.; Ando, I. *Macromolecules* **1984**, *17*, 1405–1412.
(35) Shoji, A.; Ozaki, T.; Saitô, H.; Tabeta, R.; Ando, I. *Macromolecules* **1984**, *17*, 1472–1479.
(36) Takegoshi, K.; Nakamura, S.; Terao, T. *Chem. Phys. Lett.* **2001**, *344*, 631–637.
(37) Marcotte, I.; van Beek, J. D.; Meier, B. H. *Macromolecules* **2007**, *40*, 1995–2001.
(38) Work, R. W.; Emerson, P. D. *J. Arachnol.* **1982**, *10*, 1–10.

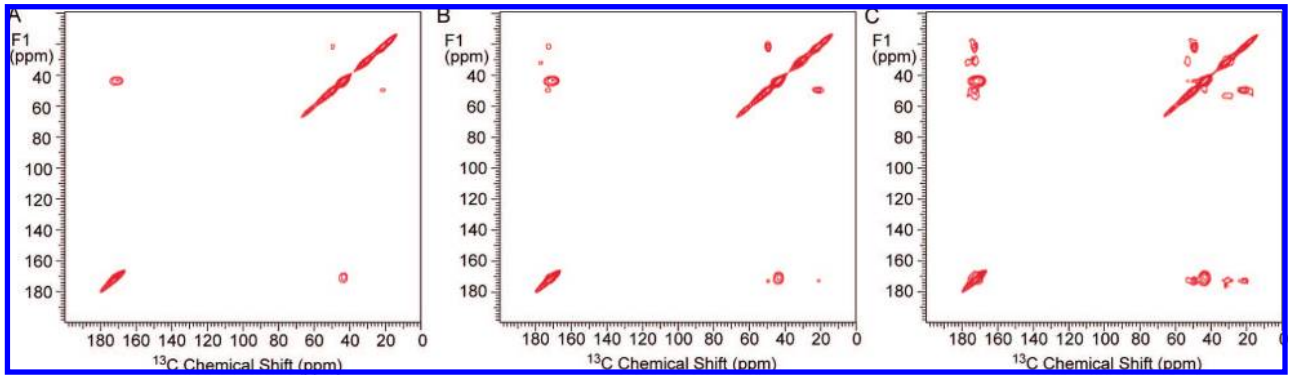


Figure 3. Two-dimensional $^{13}C-^{13}C$ NMR correlation spectra of *N. clavipes* spider dragline silk collected with different DARR mixing periods: (A) 50 ms, (B) 150 ms, and (C) 1 s. Spectra were obtained with 40 kHz MAS, 1 ms CP contact time, and a DARR condition of $n = 1$ ($\omega_1 = \omega_R$).

According to amino acid analysis conducted on major ampullate silk from *N. clavipes* spiders, these four amino acids constitute 85% of the fiber.[39] The ^{13}C enrichment was estimated by comparing peak intensities from the ^{13}C CP-MAS spectrum of the isotopically labeled sample to those observed for a natural-abundance silk sample. The enrichment is approximately 45%, 23%, 16%, 28%, 27%, and 18% for Ala C_β, Ala C_α, Gln $C_{\beta,\gamma}$, Gly C_α, Ser C_β, and the carbonyl resonance, respectively. It should be noted that there is a wealth of information in the selective differential uptake of ^{13}C by the various amino acids that is beyond the scope of the present paper.

Solid-State NMR. Two-dimensional (2D) solid-state $^{13}C-^{13}C$ correlation spectra were collected on a Varian VNMRS 800 MHz spectrometer equipped with a 1.6 mm triple-resonance MAS probe operating in double-resonance mode ($^1H/^{13}C$). The pulse sequence utilized to collect the $^{13}C-^{13}C$ correlation spectra is depicted in Figure 2. The experimental parameters for CP were a 2.2 μs ^1H $\pi/2$ pulse, a 1 ms ramped (~15%) ^1H spin-lock pulse with a radio frequency (rf) field strength of 125 kHz at the ramp maximum, and a ^{13}C square contact pulse. The MAS frequency (ω_R) was 40 kHz and the $^1H \rightarrow {}^{13}C$ CP condition was matched to the -1 spinning side band in the Hartmann−Hahn profile on the ^{13}C channel. Two-pulse phase-modulated[40] (TPPM) ^1H decoupling with a 150 kHz rf field strength was applied during acquisition with a 12° phase shift. Magnetization exchange between ^{13}C spins was enhanced by applying continuous wave (CW) irradiation on the ^1H channel at an $n = 1$ ($\omega_1 = \omega_R$) rotary-resonance (R^2) condition (see Figure 2 for pulse sequence). This 2D $^{13}C-^{13}C$ correlation experiment is analogous to the previously described DARR[36,41] and radio frequency-assisted diffusion[42] (RAD) experiments. Typical acquisition parameters were 1024 points in the direct dimension, 512 t_1 points for the indirect dimension, 32 scan averages, a 2 or 3 s recycle delay, and a 50 kHz sweep width in both dimensions. The hypercomplex (States) method was implemented to obtain phase-sensitive 2D spectra.[43] The ^{13}C chemical shift was referenced indirectly by setting the downfield (high ppm) resonance of adamantane to 38.56 ppm. Peak fitting routines were conducted with the DMFIT software package.[44]

Results and Discussion

A series of 2D $^{13}C-^{13}C$ NMR correlation spectra collected on ^{13}C-enriched *N. clavipes* spider dragline silk with increasing DARR mixing times is presented in Figure 3. At a mixing period of 50 ms (Figure 3A), off-diagonal correlations that originate from magnetization exchange between ^{13}C spins during the mixing period can be observed for Ala $C_\alpha-C_\beta$ and Gly $C_\alpha-CO$. Longer mixing periods of 150 ms (Figure 3B) are required to establish complete connectivity in Ala to include the Ala $C_\alpha-CO$ and $C_\beta-CO$ contacts. Weaker contacts are observed for Gln and Ser with the 150 ms mixing period; however, a complete connectivity between all the environments in these two amino acids can be established when a longer mixing time of 1 s is utilized (Figure 3C and expanded regions in Figure 4B). The longer mixing time requirement for Ser and Gln is due to the lower ^{13}C enrichment (see Materials and Methods) and signal-to-noise (S/N) for these two amino acids compared to Ala and Gly. The chemical shifts of all groups in each residue are readily determined from these 2D $^{13}C-^{13}C$ correlation spectra (see Table 1). It should be mentioned that the off-diagonal correlations in the 2D spectra presented in Figure 3A−C display some degree of asymmetry. This is particularly apparent when the contacts in Figure 3B are compared: more intense correlations are observed in the upper left corner of the spectrum compared to the lower right. The reason for this asymmetry in the 2D $^{13}C-^{13}C$ correlation spectra can be attributed to a combination of nonequilibrium initial magnetization resulting from CP and contributions from $^1H-^1H$ flip-flop exchange. These effects have been recently discussed in the literature and will not be addressed further here.[45]

Expansions of the high and low ppm region of the $^{13}C-^{13}C$ correlation spectrum collected with 1 s DARR mixing period are shown in Figure 4 to observe the amino acid contacts in more detail. In the carbonyl region of the spectrum (Figure 4A), the carbonyl resonance of each amino acid can be extracted by observing the $C_\alpha-CO$ and $C_\beta-CO$ contacts. For Gln, the two expected carbonyl groups are obtained: the backbone carbonyl

(39) Work, R. W.; Young, C. T. *J. Arachnol.* **1987**, *15*, 65–80.
(40) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951–6958.
(41) Takegoshi, K.; Nakamura, S.; Terao, T. *J. Chem. Phys.* **2003**, *118*, 2325–2341.
(42) Morcombe, C. R.; Gaponenko, V.; Byrd, R. A.; Zilm, K. W. *J. Am. Chem. Soc.* **2004**, *126*, 7196–7197.
(43) States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Reson.* **1982**, *48*, 286–292.
(44) Massiot, D.; Fayon, F.; Capron, M.; King, I.; Le Calvé, A.; Alonso, B.; Durand, J.-O.; Bujoli, B.; Gan, Z.; Hoatson, G. *Magn. Reson. Chem.* **2002**, *40*, 70–76.
(45) Ohashi, R.; Takegoshi, K. *J. Chem. Phys.* **2006**, *125*, 214503.
(46) Yang, Z.; Liivak, O.; Seidel, A.; LaVerde, G.; Zax, D. B.; Jelinski, L. W. *J. Am. Chem. Soc.* **2000**, *122*, 9019–9025.
(47) Holland, G. P.; Jenkins, J. E.; Creager, M.; Lewis, R. V.; Yarger, J. L. *Biomacromolecules* **2007**, *9*, 651–657.
(48) Asakura, T.; Yao, J.; Yamane, T.; Umemura, K.; Ulrich, A. S. *J. Am. Chem. Soc.* **2002**, *124*, 8794–8795.
(49) Asakura, T.; Yao, J. *Protein Sci.* **2002**, *11*, 2706–2713.
(50) Asakura, T.; Nakazawa, Y.; Ohnishi, E.; Moro, F. *Protein Sci.* **2005**, *14*, 2654–2657.
(51) Asakura, T.; Okonogi, M.; Nakazawa, Y.; Yamauchi, K. *J. Am. Chem. Soc.* **2006**, *128*, 6231–6238.

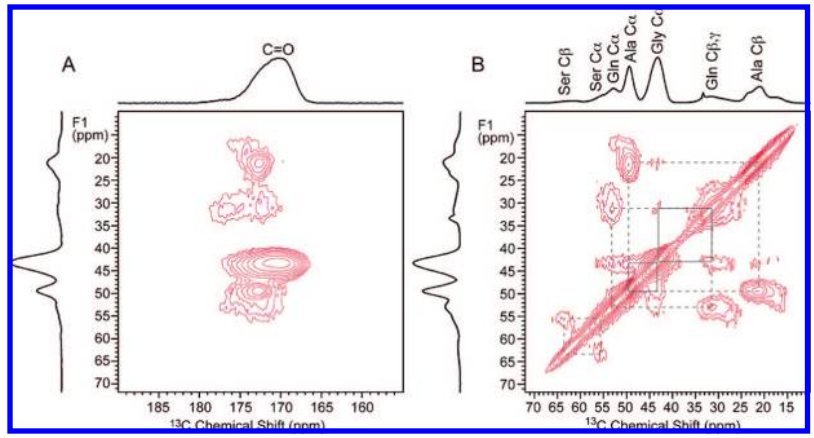


Figure 4. Two-dimensional $^{13}C-^{13}C$ NMR correlation spectra of *N. clavipes* spider dragline silk collected with 1 s DARR recoupling period. Expansions of the (A) downfield carbonyl and (B) upfield regions of the ^{13}C spectrum are shown. Strong intramolecular and weak intermolecular contacts are indicated with dashed and solid gray lines, respectively.

Table 1. ^{13}C Chemical Shifts of Major Ampullate Spider Silk and Model Polypeptides with Known Secondary Structures, Random Coil Conformation, and Silk I β-Turn Structure[a]

	^{13}C chemical shift (in ppm from TMS)					
residue	major silk	α-helix	β-sheet	random coil	3_1-helix	β-turn
Ala C_β	17.4	14.8−16.0	19.9−20.7	19.1	17.4	16.5−17.4
Ala C_β	20.9					
Ala C_β	23.3					
Ala C_α	50.0	52.3−52.8	48.2−49.3	52.5	48.9	50.8−51.7
Ala C_α	49.2					
Ala C_α	49.0					
Ala C=O	174.1	176.2−176.8	171.6−172.4	177.8	174.6	176.8−177.5
Ala C=O	172.6					
Ala C=O	172.5					
Gln C_β	32[b]	25.6−26.3	29.0−29.9	29.4		
Gln C_γ	33.2	29.7−29.8	29.7−29.9	33.7		
Gln C_α	52.9	56.4−57.0	51.0−51.4	56.2		
Gln C=O	172.1[c]	175.4−175.9	171.9−172.2	176.0		
Gln C_δ	176.5[c]			180.5		
Gly C_α	43.3		43.2−44.3	45.1	41.4−42.5	43.8−44.1
Gly C=O	171[b]		168.4−169.7	174.9	171.2−173.1	170.6−170.7
Ser C_α	55.5[c]	59.2[d]	54.4−55.0	58.3		58.0
Ser C_β	61.6[c]	60.7[d]	62.3−63.9	63.8		60.7
Ser C=O	172.2[c]		170.0−171.2	174.6		173.7

[a] Chemical shifts were extracted from $^{13}C-^{13}C$ correlation spectra collected with 150 ms DARR mixing time (Figure 2) except as noted. Model polypeptides with known secondary structures came from refs 33– 35, 49, 52, 53, 55, and 58–61; random coil conformation, from ref 62; and Silk I β-turn structure, from refs 48, 50, 55, 56, 63, and 64. [b] Broad asymmetric resonance was observed. [c] Chemical shifts were extracted from $^{13}C-^{13}C$ correlation spectra collected with 1 s DARR mixing time (Figure 2). [d] From *Antheraea pernyi* silk fibroin film, known to form α-helical structure.[65]

and C_δ. Observation of the low ppm region of the spectrum (Figure 4B) reveals all the expected contacts for Ala, Gln, Gly, and Ser. This allows the resonances to be assigned and distinct chemical shifts to be determined. This was particularly useful for the Ser C_α and Gln C_α. In the past, this region of the spectrum was assigned to the C_α of multiple amino acids, making extraction of the chemical shift of each C_α unobtainable.[46,47] The $^{13}C-^{13}C$ correlation spectrum makes it possible to assign the Ser C_α and Gln C_α by observing the Ser $C_\alpha-C_\beta$ and Gln $C_\alpha-C_{\beta,\gamma}$ contacts, respectively.

The 2D $^{13}C-^{13}C$ correlation spectrum presented in Figure 4 displays weak intermolecular contacts that can be used to characterize the repetitive amino acid motifs prevalent in the primary amino acid sequences of the proteins that comprise spider silk. For example, Gly C_α displays weak long-range contacts to Gln C_α, Gln $C_{\beta,\gamma}$, Ala C_α, and Ala C_β. This indicates that Gly is present adjacent to these other amino acids along the protein backbone, as expected from the primary amino acid sequence of MaSp1 and MaSp2 (see Supporting Information for consensus amino acid sequence).[3,5,6]

Weak intermolecular contacts can also be observed in the carbonyl region of the spectrum (Figure 4A). The distinct carbonyl resonance for Gly and Ala was obtained by taking a cross-sectional slice at the Gly C_α and Ala C_α resonances, respectively. These data are presented in Figure 5. There are a number of interesting features when the two carbonyl resonances are compared. The Gly carbonyl resonance has a full width at half-maximum (fwhm) of 1100 Hz, greater than twice that observed for the Ala carbonyl (fwhm = 480 Hz). This reflects the higher degree of structural heterogeneity for Gly compared to Ala. Closer inspection of the Gly carbonyl resonance reveals an asymmetric shape. This broad Gly carbonyl resonance actually covers the chemical shift range expected for helical and β-sheet conformations (see Figure 5A), indicating that Gly is present in at least two distinct structural environments. This is consistent with previous interpretations from solid-state NMR, where a β-sheet and helical environment was proposed for Gly.[27,28] The narrower line width and chemical shift (see Table 1) observed for the Ala carbonyl is consistent with an ordered β-sheet structure as proposed in previous work.[21] However, as

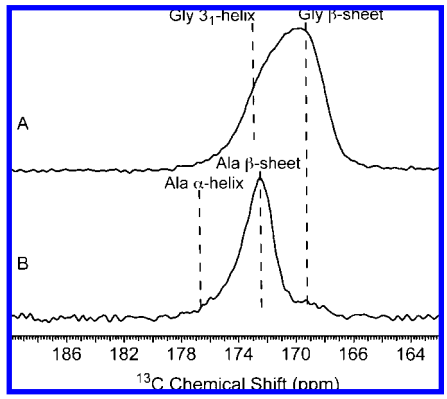


Figure 5. Slices extracted from the *F*1 dimension of the $^{13}C-^{13}C$ correlation spectrum collected with 1 s DARR mixing period. Slices were obtained at (A) Ala C_α and (B) Gly C_α. Typical carbonyl chemical shifts observed for helical and β-sheet conformations are indicated with dashed lines.

will be shown below, the Ala structural environment is more heterogeneous in spider dragline silk than previously thought.

One unique feature of the data presented in Figure 5B is a weak intermolecular contact that is observed at 169.1 ppm in the carbonyl range of the Ala C_α slice. This lies outside the carbonyl chemical shift range typically observed for Ala in differing structural conformations (see Table 1), although it is consistent with the carbonyl chemical shift expected for Gly in a β-sheet structure and lines up well with the upfield (low ppm) component of the asymmetric Gly carbonyl resonance (see Figure 5A). This allows us to conclude that this intermolecular contact results from Gly in a β-sheet structure neighboring Ala. There are only two motifs in the primary amino acid sequence to which this could correspond: poly(Gly-Ala) that flanks poly(Ala), and Gly-Gly-Ala that is located in the Gly-Gly-Xaa region of the primary amino acid sequence (see Supporting Information).

It has been recognized for over a decade that the poly(Ala) runs present in the primary amino acid sequence of the spider silk proteins form a β-structure within the fiber.[21] Close inspection of the Ala C_β resonance in the *F*2 projection from the $^{13}C-^{13}C$ correlation spectrum presented in Figure 4B reveals at minimum three components. The three components have chemical shifts at 17.4, 20.9, and 23.3 ppm. In order to better characterize these three environments, a cross-sectional slice was extracted at each position of the Ala C_β resonance (see Figure 6). The primary (20.9 ppm) and downfield components (23.3 ppm) of the Ala C_β resonance display similar C_α and carbonyl chemical shifts and are consistent with the β-sheet structural conformation (see Table 1). The difference in chemical shift for the Ala C_β component at 20.9 and 23.3 ppm is therefore attributed to packing differences of the Ala methyl groups in the β-sheet domains. This structural heterogeneity has never been detected in the ^{13}C CP-MAS spectrum of spider silk, but a structural heterogeneity in the Ala C_β environment was determined from ^2H NMR, where a highly oriented and less oriented domain was proposed.[22]

Ala heterogeneity similar to that observed here has been shown for the Ala C_β resonance in the ^{13}C CP-MAS spectra of *Bombyx mori* and *Samia cynthia ricini* silkworm silks.[48–51] In the silkworm silk studies, it was determined that the chemical shift positioned at 22−23 ppm originated from parallel methyl packing, while the component at 19−21 ppm resulted from antiparallel methyl packing in the β-sheet structure.[48–51] In one

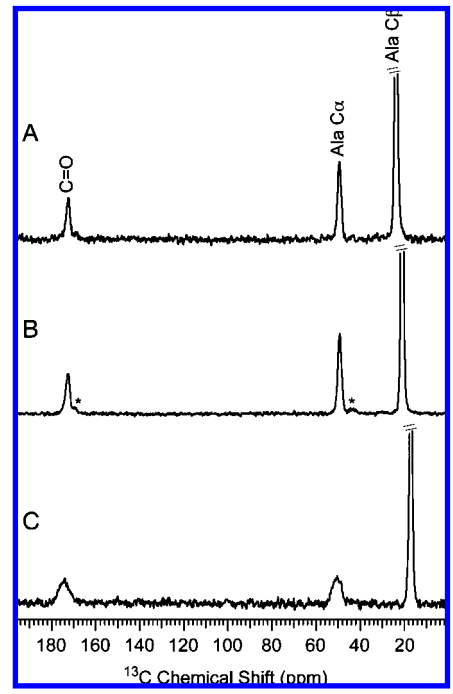


Figure 6. Ala C_β slices extracted from the *F*1 dimension of the $^{13}C-^{13}C$ correlation spectrum collected with 1 s DARR mixing period. Slices were obtained at different regions of the Ala C_β resonance: (A) 23.3, (B) 20.9, and (C) 17.4 ppm. Intermolecular contacts with Gly are denoted with asterisks.

study on model Ala_3 and Ala_4 peptides, it was found that the spin−lattice relaxation times (T_1) of parallel structures were 2−3 times longer compared to antiparallel structures, and this correlation could successfully be used to assign the components of the Ala C_β resonance in the ^{13}C CP-MAS spectrum of *S. cynthia ricini* silkworm silk.[51] This prompted a T_1 relaxation measurement with the inversion recovery method (data not shown). The T_1 relaxation time was determined to be 690 and 1040 ms for the resonances at 20.9 and 23.3 ppm, respectively. This is consistent with assigning the Ala component at 20.9 and 23.3 ppm to antiparallel and parallel β-sheet structures, respectively. The variation in T_1 values between the two structures is believed to result from methyl packing differences where shorter T_1 values are indicative of restricted mobility in the tightly packed antiparallel structure compared to the more loosely packed parallel structure.[48,51] This agrees with previous ^2H T_1 relaxation measurements on *N. clavipes* dragline silk isotopically labeled with deuterated Ala methyl groups where two distinct T_1 values were observed that were assigned to one highly oriented, densely packed and one poorly oriented, less densely packed β-sheet structure.[22]

(52) Asakura, T.; Yang, M.; Kawase, T.; Nakazawa, Y. *Macromolecules* **2005**, *38*, 3356–3363.
(53) Ashida, J.; Ohgo, K.; Komatsu, K.; Kubota, A.; Asakura, T. *J. Biomol. NMR* **2003**, *25*, 91–103.
(54) Lotz, B.; Keith, H. D. *J. Mol. Biol.* **1971**, *61*, 195–200.
(55) Ishida, M.; Asakura, T.; Yokoi, M.; Saito, H. *Macromolecules* **1990**, *23*, 88–94.
(56) Zhao, C.; Asakura, T. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 301–352.
(57) Asakura, T.; Sugino, R.; Okumura, T.; Nakazawa, Y. *Protein Sci.* **2002**, *11*, 1873–1877.
(58) Kricheldorf, H. R.; Müller, D. *Macromolecules* **1983**, *16*, 615–623.
(59) Murata, K.; Kuroki, S.; Ando, I. *Polymer* **2002**, *43*, 6871–6878.
(60) Henzler-Wildman, K. A.; Wilson, E. E.; Lee, D.-K.; Ramamoorthy, A. *Solid State Nucl. Magn. Reson.* **2003**, *24*, 94–109.

The ^{13}C cross-sectional slice of the primary Ala C_β component at 20.9 ppm (Figure 6B) displays intermolecular contacts to a component of the Gly carbonyl (169.1 ppm) similar to the slice taken at Ala C_α (Figure 5A). This component of the Ala C_β resonance also displays a contact to the Gly C_α. The conformation dependence of the Ala ^{13}C chemical shift indicated that this primary component of the Ala C_β resonance could be assigned to the poly(Ala) runs from the primary amino acid sequence that form a β-sheet structure.[21] The presence of these weak Gly contacts in the $^{13}C-^{13}C$ 2D spectrum indicates that this primary alanine component corresponds to Ala present not just in poly(Ala) but also in Gly-containing motifs like Gly-Gly-Ala and/or poly(Gly-Ala) as well. Additionally, the Gly carbonyl chemical shift extracted from these weak intermolecular contacts with Ala is consistent with Gly in a β-sheet structure (see Table 1). This indicates that Gly-Gly-Ala and/or poly(Gly-Ala) are also present in a β-sheet conformation.

The third component of the Ala C_β resonance in dragline silk fibers is positioned at 17.4 ppm. This component has been noted previously, although the only structural assignment that could be made was that it resulted from Ala present in conformations other than the β-sheet structure.[46] Recent work by our research group showed that this component sharpens considerably when dragline silk is wetted, and it was tentatively assigned to a 3_1-helical structure;[47] however, as can be seen in Table 1, this chemical shift also agrees reasonably well with the silk I β-turn structure, making the assignment somewhat ambiguous. There is one $^{13}C-^{13}C$ correlation spectrum that has been recently published on *Nephila edulis* spider silk that attempted to assign this Ala C_β resonance to a 3_1-helical structure but, an unambiguous conclusion could not be drawn.[37] The cross-sectional slice from the $^{13}C-^{13}C$ correlation spectrum reported here yields the chemical shifts of the C_α and carbonyl resonance that can be compared, along with the C_β, to known structures (Figure 6 C). Comparisons were made to known α-helix, β-sheet, random coil, silk I β-turn, and a polypeptide $(Ala\text{-}Gly\text{-}Gly)_{10}$ that is known to form a 3_1-helical structure[52,53] (see Table 1). When all three Ala chemical shifts are utilized, the best agreement is observed with the $(Ala\text{-}Gly\text{-}Gly)_{10}$ polypeptide that forms a 3_1-helical structure over the other structural motifs. The Ala C_α and carbonyl resonance is approximately twice as broad for this component (see Figure 6C) when compared with the resonances observed for the Ala components in the ordered β-sheet structure (see Figure 6A,B), reflecting a higher degree of structural disorder for the former. Thus, we assign this Ala component to a disordered 3_1-helical structure. This agrees with previous solid-state NMR torsion angle measurements made for Gly that showed that a fraction of Gly formed an approximate 3-fold helical conformation; however, this is the first time a 3_1-helix has been shown for Ala in spider silk fibers.[27,28] The formation of a stable 3_1-helix containing Ala in spider silk is further substantiated by the demonstration that poly(Ala-Gly-Gly) forms stable structures of this conformation via solid-state NMR[52,53] and XRD.[54]

The three Ala C_β components display nearly identical CP build-up kinetics (data not shown), making quantification of the different components possible. When this resonance is fit to three components, the fraction of Ala in the β-sheet structure is 80% ± 5% (combination of component at 20.9 and 23.3 ppm) and the 3_1-helical fraction (component at 17.4 ppm) is 20% ± 5% (see Supporting Information). If we return to the primary amino acid sequence of MaSp1 and MaSp2 and assume that the poly(Gly-Ala) motifs flanking the poly(Ala) runs all form β-sheet structures and the remaining fraction of Ala located in Gly-Gly-Ala motifs form 3_1-helices, the ratio is 85:15.[5,6] This is in excellent agreement with the ratio obtained from the fitted C_β resonance in the ^{13}C CP-MAS spectrum of ^{13}C-enriched spider dragline silk and also agrees well with fractions of Ala in β-sheet and helical domains extracted from wetted spider dragline silk.[47] It is therefore concluded that the poly(Ala) and poly(Gly-Ala) regions flanking poly(Ala) all form β-sheet structures, while the remaining Gly-Gly-Ala motif forms a disordered 3_1-helical structure. Although it was thought previously that poly(Gly-Ala) could be incorporated in the β-sheet domains, direct experimental evidence was lacking until now.[12,46] This result is of specific relevance since a number of studies have speculated that the Gly-Gly-Xaa motif may form a β-sheet structure.[5,15,17,28] This appears not to be the case, at least for the Gly-Gly-Xaa domains where Xaa = Ala that do not flank the poly(Ala) runs in the primary amino acid sequence of spider silk proteins.

The structural conformations of the other amino acids besides Gly and Ala that comprise spider silk fibers remain essentially uncharacterized. One REDOR NMR study determined that the Leu-Gly-Xaa-Gln motif (where Xaa = Ser, Gly, or Asn) formed compact turnlike structures.[30] The chemical shifts of all the Ser resonances have been obtained for the first time from the $^{13}C-^{13}C$ NMR correlation spectra (see Figure 7A for Ser C_α slice). This allows for an initial structural assessment to be made for Ser from the conformation-dependent ^{13}C chemical shift. One possibility for Ser is that it could be incorporated into the β-sheet structure, similar to silkworm silk fibers.[34,55–57] However, when the Ser chemical shifts observed for spider dragline silk were compared to those expected for a β-sheet structure, no agreement was observed (see Table 1). It should be noted, though, that the Ser resonances are broad and a minor component could represent the β-sheet structure, but the center of gravity does not appear to agree with the known chemical shifts for Ser in a β-sheet. Another possibility is that Ser could form a β-turn structure similar to Silk I. Again no agreement is observed for the Silk I form either. In fact, the Ser chemical shifts do not fall in the range of any of the known structures presented in Table 1. Unfortunately, there are no known Ser structures that form a 3_1-helix to our knowledge. But it should be pointed out that the observed chemical shifts for a 3_1-helical structure lie almost exactly between that observed for a β-sheet and an α-helix for Ala, and similar chemical shift behavior is observed for Ser C_α and C_β (see Table 1). This is speculative at this point and will require future work to confirm. It is also interesting to note that weak intermolecular contacts are observed to Gly C_α and possibly Gln $C_{\beta,\gamma}$ (see Figure 7A). However, the latter may just be the result of close overlap between the Ser C_α and Gln C_α resonances (see Figure 1).

The chemical shifts for Gln were determined from a Gln C_α and Gln C_γ slice (Figure 7B,D). The observed Gln chemical shifts were compared to those known for polyglutamates in different secondary structures (see Table 1).[35,58] The Gln C_α,

(61) Kishi, S.; Santos, A.; Ishii, O.; Ishikawa, K.; Kunieda, S.; Kimura, H.; Shoji, A. *J. Mol. Struct.* **2003**, *649*, 155–167.
(62) Wishart, D.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67–81.
(63) Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komatsu, K. *J. Mol. Biol.* **2001**, *306*, 291–305.
(64) Asakura, T.; Demura, M.; Date, T.; Miyashita, N.; Ogawa, K.; Williamson, M. P. *Biopolymers* **1997**, *41*, 193–203.
(65) Nakazawa, Y.; Asakura, T. *Macromolecules* **2002**, *35*, 2393–2400.

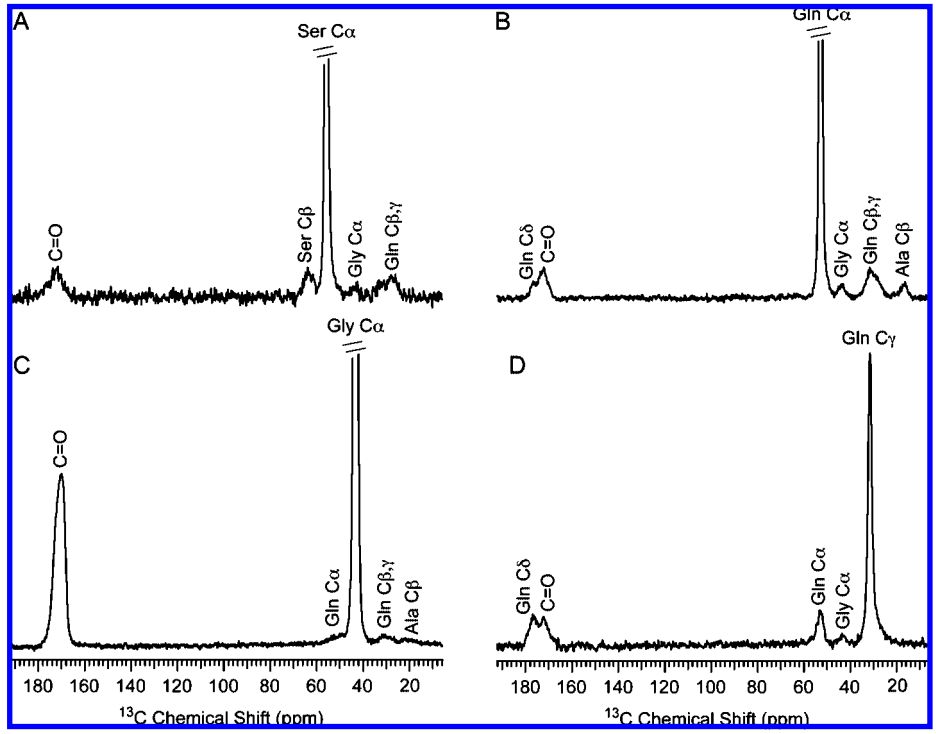


Figure 7. Slices extracted from the *F*1 dimension of the $^{13}C-^{13}C$ correlation spectrum collected with 1 s DARR mixing period. Slices were obtained at (A) Ser C_α, (B) Gln C_α, (C) Gly C_α, and (D) Gln C_γ.

C_β, and carbonyl ^{13}C chemical shifts observed for Gln in spider dragline silk fall outside the range observed for polyglutamates in a β-sheet or α-helical structure with the exception of the backbone carbonyl resonance. The backbone carbonyl resonance actually lies within the range observed for β-sheet structures. However, when the C_α and C_β shifts are considered, the β-sheet structure appears unlikely. This shows how knowing all the chemical shifts is useful when secondary structure is assessed from ^{13}C chemical shifts. Again, similar to Ser, the Gln resonances are broad and a minor component that agrees with the β-sheet structure could be present, but the center of gravity does not appear to agree with a β-sheet conformational assignment. It is concluded that the bulk of Gln residues are not present in an α-helical or β-sheet conformation. Clearly the Gln environment shows strong intermolecular contacts to Gly (see Figure 7B−D) consistent with these Gln resonances corresponding to the Gly-Gly-Gln motif. This motif may also form a 3_1-helix similar to Gly-Gly-Ala; however, this cannot be confirmed since there are no known model structures of Gln in a 3_1-helical conformation. Further study is needed to determine the secondary structure of Gln in this motif, which will be the subject of a future work. However, it is readily determined that the majority of Gln are not incorporated in the β-sheet structure as would be expected when the bulky side chain of Gln is considered.

Conclusions

The $^{13}C-^{13}C$ NMR correlation spectrum of *N. clavipes* spider dragline silk fibers has been presented. A complete NMR resonance assignment could be made, and the chemical shift from all residue sites of the labeled amino acids (Gly, Ala, Gln, and Ser) have been reported. These results indicate two structural environments from the conformation-dependent ^{13}C chemical shifts for both Gly and Ala. The structural environments are characterized by an ordered β-sheet structure and a disordered 3_1-helical structure. Weak intermolecular contacts from the $^{13}C-^{13}C$ correlation spectra and the quantitative nature of NMR have allowed us to assign these structures to three distinct amino acid motifs prevalent from the primary amino acid sequence of spider silk proteins. The ordered β-sheet structure was ascribed to both the poly(Ala) and poly(Gly-Ala) motifs flanking the poly(Ala) runs in the amino acid sequence of MaSp1, while the disordered 3_1-helical structure is ascribed to the Gly-Gly-Ala motif in the remaining region of the sequence. The Ala β-sheet structure was also shown to be heterogeneous, similar to silkworm silk fibers where antiparallel and parallel domains appear to be present. Gly is incorporated in the antiparallel domains, while Gln and Ser appear to be absent from any of the β-sheet structural regions. The NMR approach presented here has provided new insight into the structure of spider silk, is easy to implement, and should be widely applicable to the characterization of heterogeneous fibrous proteins like silks.

Acknowledgment. This research was supported by grants from the National Science Foundation (CHE-0612553 and DMR-0805197) and the National Institutes of Health (NIBIB-5R01EB000490-05). We thank Dr. Brian Cherry for help with NMR instrumentation, student training, and scientific discussions.

Supporting Information Available: Consensus primary amino acid sequence of spider silk proteins MaSp1 and MaSp2 and fit of the Ala methyl resonance in ^{13}C CP-MAS NMR spectrum. This material is available free of charge via the Internet at http://pubs.acs.org.

JA8021208