Characterizing the Secondary Protein Structure of Black Widow Dragline Silk Using Solid-State NMR & X-ray Diffraction

Janelle E. Jenkins[a], **Sujatha Sampath**[a,b], **Emily Butler**[a], **Jihyun Kim**[a], **Robert W. Henning**[c], **Gregory P. Holland**[a,*], and **Jeffery L. Yarger**[a,*]

[a] Department of Chemistry and Biochemistry, Magnetic Resonance Research Center, Arizona State University, Tempe, Arizona 85287-1604, USA

[b] Department of Physics, University of Wisconsin, Milwaukee, WI 53211

[c] Center for Advanced Radiation Sources, The University of Chicago, Chicago, IL 60637.

Abstract

This study provides a detailed secondary structural characterization of major ampullate dragline silk from *Latrodectus hesperus* (black widow) spiders. X-ray diffraction results show that the structure of black widow major ampullate silk fibers is comprised of stacked β-sheet nanocrystallites oriented parallel to the fiber axis and an amorphous region with oriented (anisotropic) and isotropic components. The combination of two-dimensional (2D) ^{13}C-^{13}C through-space and through-bond solid-state NMR experiments provide chemical shifts that are used to determine detailed information about amino acid motif secondary structure in black widow spider dragline silk. Individual amino acids are incorporated into different repetitive motifs that make up the majority of this protein-based biopolymer. From the solid-state NMR measurements, we assign distinct secondary conformations to each repetitive amino acid motif and hence to the amino acids that make up the motifs. Specifically, alanine is incorporated in β-sheet (poly(Ala$_n$) and poly(Gly-Ala)), 3_1-helix (poly(Gly-Gly-X$_{aa}$), and α-helix (poly(Gln-Gln-Ala-Tyr)) components. Glycine is determined to be in β-sheet (poly(Gly-Ala)) and 3_1-helical (poly(Gly-Gly-X$_{aa}$)) regions, while serine is present in β-sheet (poly(Gly-Ala-Ser)), 3_1-helix (poly(Gly-Gly-Ser)), and β-turn (poly(Gly-Pro-Ser)) structures. These various motif-specific secondary structural elements are quantitatively correlated to the primary amino acid sequence of major ampullate spidroin 1 and 2 (MaSp1 and MaSp2) and are shown to form a self-consistent model for black widow dragline silk.

INTRODUCTION

Due to the strength of dragline silk, researchers are attempting to use it as a model for synthetic fibers.[1,2] However, the production of synthetic fibers has proven to be more difficult than initially predicted. Although large portions of the proteins that make up dragline silk have been sequenced[3-6] and considerable amounts of these proteins can be expressed in different media,[7-9] to date, fibers formed from expressed proteins do not have mechanical properties that compare to natural dragline silk.[8,10-12] Before progress can be made towards the production of synthetic silk, research must take a step back and develop a fundamental understanding of the molecular protein structures that are formed in the natural

fibers. Once molecular level structure is elucidated, then it can be correlated to mechanical function and aid in developing a foundation for the structural engineering of synthetic fibers. Black widow dragline silk provides a unique opportunity to better quantitatively correlate the primary amino acid sequence with secondary structure, as currently it is the first and only dragline spider silk to be fully sequenced.[13]

Major silk (often called dragline silk) from the black widow spider, *Latrodectus hesperus,* is a biopolymer with mechanical properties that exceed many manmade materials in terms of strength and elongation.[14,15] Black widow (BW) major ampullate (Ma) silk is primarily composed of two large proteins, major ampullate spidroin 1 (~250 kDa) and major ampullate spidroin 2 (~300 kDa) (MaSp1 and MaSp2).[13,16] Each of the proteins is characterized by a highly repetitive primary amino acid sequence, where 90% of the proteins are made up of only five amino acids, glycine (Gly), alanine (Ala), glutamine (Gln), tyrosine (Tyr), and serine (Ser).[17] Individual amino acids are incorporated into simple motifs that are repeated throughout the primary amino acid sequence (Figure 1). Many of these amino acid motifs are conserved in dragline silk between species.[18] For example, Ala is found in poly(Ala_n), poly(Gly-Ala), and poly(Gly-Gly-X_{aa}), where X_{aa} = Ala, Gln, and Tyr in multiple species.[13,19,20] Each of the repeated motifs is thought to adopt and pseudo regular secondary structure and linked to some specific aspect of the mechanical properties of spider silk (i.e, toughness, elasticity).[16,21] Studies have shown that Ala in spider's silk can be incorporated into a β-sheet, as well as a helical secondary structure.[21] It was shown that Ala in poly(Ala) and poly(Gly-Ala) were responsible for the β-sheet and the helical component was from poly(Gly-Gly-Xaa), where Xaa = Ala.

Different techniques have been utilized to study the physical and structural characteristics of BW Ma silk. Stress-Strain mechanical testing have been performed on black widow silk and shows BW dragline silk to be a very tough biopolymer fiber.[15,22] Atomic force microscopy (AFM) has been used to study the surface of black widow silk in the stretched and unstretched form.[23] Transmission electron microscopy (TEM) and XRD have been utilized the characterize crystallite size in BW Ma silk.[24] To date, however, there has not been a detailed molecular secondary structural study of BW dragline silk fibers.

Solid-state NMR has been used extensively to characterize the molecular structure of major silk from various *Nephila* species.[21,25-29] The major silk from these species contain similar amino acid motifs that are found in BW Ma silk.[13,19,20,30] The poly-Ala regions of all spider silks studied by NMR have been determined to be in a β-sheet structure.[3,31,32] Poly(Gly-Ala) in *N. clavipes* has also been determined to be incorporated in a β-sheet structure utilizing two-dimensional (2D) ^{13}C-^{13}C correlation experiments with dipolar-assisted rotational resonance (DARR) solid-state NMR.[21] The same study showed that the Ala in poly(Gly-Gly-Xaa) motif not found next to β-sheet can be assigned to a disordered 3_1-helix.[21] Ala was also assigned to both β-sheet structure and 3_1-helix for *N. edulis* utilizing 2D ^{13}C-^{13}C proton-driven spin-diffusion (PDSD) correlation and total through-bond correlation spectroscopy (TOBSY) with solid-state NMR.[29] Although solid-state NMR has been used to investigate the structure of many silks, such as silkworm silk, spider silk, and synthetic mimics of silk,[6,21,25,28,33] black widow major ampullate (BW Ma) silk has essentially been uncharacterized utilizing solid-state NMR.[22]

Both 2D through-space ^{13}C-^{13}C correlation spectrum with a DARR[34] mixing period and 2D through-bond ^{13}C-^{13}C double quantum/single quantum (DQ/SQ) refocused Incredible Natural Abundance DoublE QUAntum Transfer Experiment (INADEQUATE),[35] solid-state NMR experiments are utilized to extract chemical shifts for individual amino acids that would not otherwise be assessable from the inhomogeneously broadened one-dimensional ^{13}C spectra of BW Ma silk. The chemical shifts are then used to determine the

secondary protein structures adopted by the amino acids found in the various repetitive amino acid motifs that make up the proteins in BW Ma silk. In this way, the average secondary structures of the primary amino acid motifs that make up the BW dragline silk proteins were characterized. The secondary structures are then quantitatively correlated to the primary amino acid sequence of MaSp1 and MaSp2. Establishing a quantitative correlation between the secondary structure in spider silk and the primary amino acid sequence is an important step towards better understanding the structure and function relationship in silk. This information should assist in directed molecular engineering of synthetic spider silk fibers.

MATERIALS AND METHODS

Sample Preparation

Adult female *Latrodectus hesperus* (black widow, BW) spiders were forcibly silked at a rate of 2 cm/s under a dissection microscope to ensure that only major silk was collected, similar to previous methods.[36] Spiders were silked every other day while being fed isotopically enriched amino acids dissolved in water. On average, spiders drank 20-30 μL each silking. Three groups of spiders (4-6 spiders per group) were each fed a uniformly labeled amino acid solution; (1) U-^{13}C/^{15}N-Ala (15% w/v), (2) U-^{13}C/^{15}N-Pro (15% w/v), and (3) U-^{13}C/^{15}N-Ser (4% w/v). The three amino acids metabolize differently in the spider and can be used to observe different portions of the silk. U-^{13}C/^{15}N-Ala predominately enriches the ^{13}C residues for Ala, Gly, and Gln. U-^{13}C/^{15}N-Pro is used to label the Pro that is only found in MaSp2. U-^{13}C/^{15}N-Ser enriches Ser, but also metabolized into Gly, so a large Gly peak is observed. All uniformly labeled ^{13}C and ^{15}N amino acids were purchased from Cambridge Isotopes Laboratories, Inc. In addition to the amino acid solutions, spiders were fed one cricket per week. For the natural abundance NMR silk samples and the XRD silk samples, spiders were silked on average every other day and fed 1-2 crickets per week. For NMR studies of wet spider silk, the major silk was soaked in 99.9 % D_2O for 30 minutes to ensure that the silk was water-saturated.

Solid-State NMR

All NMR experiments were collected on a Varian VNMRS 400 MHz wide-bore spectrometer with a 3.2 mm triple resonance magic angle spinning (MAS) probe. All ^{13}C experiments were referenced externally to the downfield (methylene) adamantane peak at 38.56 ppm. The $^1H \rightarrow {}^{13}C$ cross-polarization (CP) MAS NMR spectrum of natural abundance *Latrodectus Hesperus,* black widow*,* major silk was collected at a MAS speed (ω_r) of 10 kHz. The $^1H\rightarrow{}^{13}C$ CP condition was matched to the -1 spinning sideband (ssb) in the Hartmann-Hahn (HH) profile. CP conditions consisted of a 3.6 μs ^1H pulse, followed by a 1 ms ^1H spin-lock pulse of 70 kHz rf field strength. Free induction decays were collected using a recycle delay of 3 s, 12,000 transients and 1 k complex points, and two-pulse phase-modulated (TPPM) ^1H decoupling with a rf field strength of 100 kHz was used during acquisition.[37]

All 2D solid-state NMR experiments were collected at ω_r = 20 kHz with the $^1H\rightarrow{}^{13}C$ CP condition matched to the -1 ssb in the HH profile and TPPM ^1H decoupling with a rf field strength of 100 kHz used during acquisition. The DARR experiments were collected using a sweep width of 25 kHz in both dimensions and 512 complex points collected in the direct dimension and 128 points (t_1) in the indirect dimension. CP conditions consisted of a ^1H pulse, followed by a 1 ms ramped (10%) ^1H spin-lock pulse of 75 kHz rf field strength. During DARR mixing (τ_{mix}) periods, continuous wave (CW) irradiation was applied on the ^1H channel at a rotary-resonance condition, $\omega_r = \omega_{rf}$ = 20 kHz. A τ_{mix} = 150 ms was utilized for the U-^{13}C/^{15}N-Pro labeled silk and a τ_{mix} = 1 s was used for the U-^{13}C/^{15}N-Ser

labeled silk. A recycle delay of 2 s was used and 32 transients were collected, for an average experimental time of 5 hours.

The refocused double quantum/single quantum (DQ/SQ) INADEQUATE NMR experiment was collected on U-^{13}C/^{15}N-Ser labeled Ma silk with a sweep width of 25 kHz in the direct dimension and 50 kHz in the indirect dimension with 128 points collected. CP conditions consisted of a 3 μs ^{1}H pulse, followed by a 1 ms ramped (11%) ^{1}H spin-lock pulse of 83 kHz rf field strength. A τ delay of 3 ms was found to be optimal for the sample. The recycle delay of 2 s was used and 128 transients collected for a total experimental time of 18 hours.

^{13}C directly detected NMR spectra with ^{1}H decoupling (^{13}C [^{1}H] DD-MAS) of dry and wet U-^{13}C/^{15}N-Ser labeled BW Ma silk were collected at ω_r = 20 kHz. The spectra were collected with a 50 kHz sweep width, a recycle delay of 1 s, and 12 k transients, and ^{1}H TPPM decoupling of 100 kHz was used during acquisition. Fully relaxed ^{13}C [^{1}H] DD-MAS NMR spectrum was collected with a recycle delay of 60 s for fitting the Ser C_β. Fully relaxed ^{13}C[^{1}H] DD-MAS NMR spectrum was collected on wet U-^{13}C/^{15}N-Ala labeled major silk at ω_r = 20 kHz using recycle delay of 50 s, with 100 kHz ^{1}H TPPM decoupling during acquisition. Dmfit was used for deconvolution of the spectra.[38] The fits were purely Lorentzian distributions. To fit the carbonyl peak, the linewidth and chemical shift for individual carbonyl resonances were extracted from a 2D refocused INADEQUATE (Supplemental). Initially, the linewidths and chemical shifts were held constant and the amplitude of each peak was varied. These initial parameter did not provide a good fit. Therefore the linewidths were incremented by 5% until the linewidths were 25% larger, providing a good fit of the carbonyl peak. Although the linewidths were increased by 25%, the overall percentage of the fits only changed by ± 3%. Because of the short recycle delay in the INADEQUATE experiment, it appears that this experiment enhances the most mobile components of the sample and therefore underestimates the linewidths. In this case, it is appropriate to use larger linewidths for fitting. Three distinct chemical shifts for the Ser C_β are observed in the INADEQUATE experiment, however, only two distinct sites are observed in the ^{13}C [^{1}H] DD-MAS NMR spectrum. Two sites were fit for the Ser C_β, one for the β-sheet and one that contains the helical and turn component. Chemical shifts were extracted from the INADEQUATE experiment (Figure 6). The chemical shift for the β-sheet component was held constant and the chemical shift for the helical component was allowed to vary. The chemical shift for the non-β-sheet component was determined to be between the turn and helical chemical shifts that were extracted from the INADEQUATE experiment.

Primary amino acid motif model and quantification method

The complete protein sequence for MaSp1 and MaSp2 of *Latroductus hesperus* are from GenBank, accession no. EF595246 and EF595245, respectively. The molar ratio of MaSp1 to MaSp2 was calculated to be 3:1 by using the percent Pro content found in the silk.[30] The following rules were used to establish an amino acid motif model and quantify the amount of β-sheet in the proteins: (I) Poly-Ala has previously been established to be in a β-sheet secondary structure in spider silk,[3,21,39] therefore any poly(Ala) motifs of 4 or more amino acids were counted as β-sheet. (II) Every Gly, Ser, or Gly-Ala on either side of the poly-Ala motifs were also included in a β-sheet. In almost every case in MaSp1, the β-sheet runs are terminated by Gln. In the case of MaSp2, the β-sheet runs are generally terminated by Pro on either side. All other amino acids not assigned to a β-sheet are glycine-rich and make up the helical, turn, and random coil regions within silk protein biopolymer. This is consistent with previous counting methods used to make a quantitative model of the structural motifs from *N. clavipes.*[27]

X-Ray Diffraction

X-ray fiber diffraction was performed on the native BW Ma silk (original length L_0), supercontracted silk (shrunk to 0.7 L_0 upon wetting with water) and the restretched silk from the supercontracted state (0.9 L_0), at the BioCARS 14BM-C beamline in the Advanced Photon Source at Argonne National Laboratory, Argonne IL, USA. The wavelength of the x-ray beam was 0.9 Å with a flux of 6×10^{11} photons/sec and the beam size on the sample was 150×200 μm. Data were recorded using an ADSC Quantum-315 detector. A bundle of 20 individual fibers were held taut (but unstretched) in a metal frame made out of a thick wire, with the fiber axis normal to the x-ray beam. The sample to detector distance was 200 mm. Data collection time was 60 seconds for one image. Background measurements were performed with the sample displaced from the beam and the image was recorded under the same conditions as used with the sample in the beam. Multiple images were taken to get better statistics and improve on the signal/background ratio. CeO_2 powder was used for instrument calibration. The 2D wide angle x-ray diffraction (WAXD) patterns were analyzed using the software package FIT2D. Radial and azimuthal 1D profiles were obtained from the deconvolution of 2D WAXD images using FIT2D. Radial profiles are intensity as a function of radius integrated azimuthally over a sector typically 20-30 degrees wide, while azimuthal profiles are integrated intensity as a function of all azimuthal angles over a very thin annular ring centered at the peak maximum of the desired reflection.[40,41] Microcal Origin was used for the deconvolution of the 1D x-ray data and peak fitting. The peaks were fitted with Gaussian functions using nonlinear least squares fitting. Like previous x-ray diffraction analyses on spider silks, sequential analysis instead of whole pattern fitting has been applied.[40,41]

Before wetting, the metal frame was bent so that the ends were brought close together such that the silk, which was previously held taut, was now hanging slackly. Distilled water was added to the silk by a pipette and the silk contracted freely and reduced in length. Excess water was removed with a paper towel and the silk was let to dry in room temperature for 5 minutes. The ends of the frame were separated and straightened such that the silk was held taut again. WAXD patterns were taken under the same conditions described above for the native silk. To restretch the silk, the supercontracted silk was wetted with distilled water and cautiously stretched to as close as was possible to its original length (0.9 L_0) and dried before performing WAXD.

RESULTS

The full black widow MaSp1 and MaSp2 sequences are color coded in Figure 2 to illustrate the amino acid motif model that is used to predict the percentage of β-sheet in the silk. All poly(Ala_n), as well as Gly and Ser terminating poly(Ala_n) are predicted to be in a β-sheet secondary structure and are colored red, blue, and green, respectively. The rest of the sequences are colored black indicating that these amino acids are incorporated into non-β-sheet structures, including helices, turns, and random coil.

The ^{13}C CP-MAS NMR spectrum of natural abundant BW Ma silk is shown in Figure 3. The peaks are significantly broadened primarily due to the heterogeneity of spider silk protein structure, causing a distribution of chemical shifts for each carbon resonance. Although the peaks are broad, amino acids that make up > 3 % of the silk, alanine (Ala), glycine (Gly), glutamine (Gln), tyrosine (Tyr), and serine (Ser), are observed in the spectrum of the natural abundant silk. These resonances are assigned based on chemical shift. Distinct peaks are observed for Ala C_α, Ala C_β, Gly C_α, Ser C_β, Tyr C_ε, and the Tyr C_ξ. The C_α of Leu, Gln, Ser, and Tyr overlap around 52 ppm. Tyr C_γ and Tyr C_ε overlap at 132 ppm. Gln C_β and Gln C_γ are indistinguishable from each other at 32 ppm. The Ala C_β exhibits a broad

peak that spans from approximately 16 ppm to 22 ppm indicating that Ala is incorporated into multiple secondary structures.

The 2D ^{13}C-^{13}C correlation spectra collected of BW Ma silk enriched with U-^{13}C/^{15}N Pro is shown in Figure 4. When fed a solution of U-^{13}C/^{15}N Pro, the Pro and Ala resonances become ^{13}C enrichment. Proline is only found in MaSp2, therefore this labeling method provides a way to observe this protein in BW dragline silk. The 2D correlation was collected with a short τ_{mix} = 150 ms to provide intramolecular correlations for individual amino acids. These intramolecular correlations permit the chemical shift assignment for each carbon site within a specific amino acid. The chemical shifts for Pro C_α, Pro C_β, Pro C_γ, Pro C_δ, and Pro CO are 60.4, 29.6, 24.9, 47.3 and 173.8 ppm, respectively. Three distinct correlations with differing chemical shifts are observed in the silk between the Ala C_α and the Ala C_β, indicating that Ala is found in three different secondary structures. The main correlation is between the Ala C_α at 48.6 ppm and the Ala C_β at 20.4 ppm. The two less intense Ala correlations are between Ala C_α at 49.4 ppm and the Ala C_β at 17.0 ppm and between Ala C_α at 52.4 ppm and the Ala C_β at 16.4 ppm. Intramolecular Gln, Gly, and Pro resonances are also observed. Intramolecular Gln correlations are observed between Gln C_α (52.4 ppm), Gln C_β (27.4 ppm), Gln C_γ (30.9 ppm), and Gln CO (172.8 ppm).

The 2D ^{13}C-^{13}C correlation NMR spectrum of ^{13}C-enriched BW Ma silk by feeding a solution of U-^{13}C/^{15}N Ser is shown in Figure 5A. When fed U-^{13}C/^{15}N Ser, the Ser resonances of the silk are ^{13}C-enriched. Serine metabolizes to Gly, therefore Gly resonances are also ^{13}C-enriched. Both intra- and intermolecular amino acid correlations are observed using τ_{mix} = 1 s. Strong intramolecular correlations are observed for Ala C_α-C_β, Ala C_α-CO, and Ala C_β-CO. Three intermolecular correlations are observed between the Ala C_β and the Gly C_α. A projection in the direct dimension (*F2*) taken at the Ala C_β chemical shift of 19.3 ppm in the indirect dimension (*F1*) shows that Gly C_α has two distinct chemical shifts of 41.3 and 43.0 ppm as shown in Figure 5B. The Gly C_α peak at 41.3 ppm exhibits a correlation to the Ala C_β at 20.4 ppm, as well as to the lower ppm shoulder of the Ala C_β peak, while the Gly C_α at 43.0 ppm only exhibits a strong correlation to the Ala C_β at 20.4 ppm.

The through-bond DQ/SQ refocused INADEQUATE NMR spectrum of ^{13}C-enriched BW Ma silk from spiders fed U-^{13}C/^{15}N Ser is shown in Figure 6. The spectrum reveals three double quantum correlations for Ser C_α-C_β, indicating that Ser is found in three discrete secondary structures. Double quantum correlations are seen for Ser C_α (54.6 ppm)-C_β (64.0 ppm), Ser C_α (55.6 ppm)-C_β (62.5 ppm), and Ser C_α (57.6 ppm)-C_β (60.8 ppm). To confirm the Ser chemical shift assignments in the refocused INADEQUATE, a direct ^{13}C MAS directly-detected (DD-MAS) NMR spectra was collected of the ^{13}C-enriched BW Ma silk with a short recycle delay of 1 s, both in the natural silk form and wetted state (Figure 7). The upfield (low ppm) side of the Ser C_β peak increases in intensity when wetted with water. Two resonances are also enhanced in the Ser C_α region, a peak at 53.4 ppm and 55.8 ppm. Additionally, there are two resolved resonances at 27.7 ppm and 31.9 ppm, which are consistent with Gln C_β and Gln C_γ, respectively.

The carbonyl peak of a fully relaxed ^{13}C[^{1}H] DD-MAS NMR spectrum was deconvoluted to extract the percentage of Ala and Gly incorporated into a β-sheet and helical secondary structure, as shown in figure 8. The chemical shift and linewidths for individual amino acids were extracted from a directly polarized INADEQUATE experiment (data not included) in the same manner as Holland, et al.[39] During deconvolution, the known chemical shift and linewidth of each peak are held constant and only the amplitude is varied. The fraction of Ala from black widow major silk determined to be in a β-sheet is 88 % and 12 % incorporated into a helical secondary structure. Glycine was determined to be 40 % β-sheet

and 60% helical from the carbonyl fit. For the deconvolution of the Ser C_β peak, chemical shifts were extracted from the directly polarized INADEQUATE experiment and held constant for fitting. However, due to the overlap of the Ser C_β peaks, linewidths could not be extracted from the INADEQUATE experiment, so the linewidths were varied during fitting. From the fit, Ser was determined to be 42 % β-sheet and 58 % non-β-sheet. The non-β-sheet components could include contributions from helical, turn, and random coil structures.

X-Ray Diffraction

The WAXD pattern of BW Ma native silk is shown in Figure 9a. The diffraction pattern depicts a semi-crystalline morphology characterized by Bragg reflections and an amorphous halo. The predominant structural model for spider silk fiber is that it contains a nanocrystalline component as well as an oriented (anisotropic) and an isotropic amorphous component.[40,42,43] The crystalline reflections are broad due to the small size of the crystallites (nanocrystalline).[40] Intense reflections are observed along the equator (perpendicular to the fiber axis). Along the meridian (parallel to the fiber axis) the main (002) reflection is strong, while subsequent higher order reflections are weaker. The reflections comprising the crystalline fraction correspond to the pattern of β-poly(L-alanine) structure.[40,42,43] The reflections were indexed based on an orthogonal unit cell with the *c*-axis as the fiber axis.[43,44] According to the prevalent structural model for spider silk, it has a hierarchical structure with each silk fiber being composed of many fibrils and each fibril in turn composed of anisotropic nano-crystalline domains connected by an amorphous matrix, which has both an oriented and isotropic components.[40,41] Crystallite size, orientation distribution of the crystallites about the fiber axis, and the percentage crystallinity in the silk fiber can be determined from WAXD.

Radial integration along the equator and meridian in the WAXD pattern gives the radial peak width of the (200), (120) and (002) reflections. Figure 10a shows the 1D radial intensity profile along the equator, which has been integrated azimuthally over a sector to include the two main equatorial Bragg reflections (200)/(120). The inner weaker (200) peak on the equator corresponds to the distance between β-sheets and is denoted as the *a*-axis of the crystallite while outer stronger (120) peak corresponds to the interchain distance and is denoted here as the *b*-axis of the crystallite.[40-42,44] The (002) peak on the meridian corresponds to the size of the crystallite in the *c* direction, along the fiber axis. The 1D equatorial radial-profile has been deconvoluted into two crystalline peaks and an amorphous peak, by fitting with Gaussian peaks for the Bragg reflections, a broader Gaussian peak for the amorphous component (includes both oriented and unoriented). The broad peak at higher Q is due to superposition of multiple weak reflections, included so as to have a higher angular range of data for fitting and has not been used for crystallite size calculation. The intensity at 20° from the equator is almost entirely from the amorphous halo, which has been used as a constraint for the amorphous peak in the fitting process. The radial width of the inner (200) peak is wider than that of the outer (120) peak, implying that the crystallites are thinner perpendicular to the hydrogen-bonded planes (200), as has been seen with other silks.[40,41,45] Figure 10b shows the 1D radial intensity profile of the (002) meridional peak. The shape of the (002) peak as a function of diffraction angle is asymmetric at higher angles, due to overlap with other layer-line reflections.[40] This peak has been fitted using two Gaussians one for the crystalline component and a broader Gaussian for the amorphous halo, the small sharp peak at higher angles has been treated as a separate peak which could be an overlap of other weaker reflections or due to the orientational broadening of layer line reflections. Average size of the crystallites along the *a*, *b* and *c* axes have been determined by using Scherrer's formula.[40] The average crystallite size was found to be 3.02, 4.15, and 6.71 *nm* along the *a*, *b* and *c* directions. The degree of crystallinity x_c was estimated from the radial intensity profile of the strongest equatorial reflections (200)/(120) (Figure 10a) using

the relation $x_c = (I_c/I_c + I_a)$, I_c is the sum of the integrated intensities of the (200) and (120) crystalline peaks and I_a is the integrated intensity of the amorphous halo. This results in a value of 40% for the crystalline fraction in BW Ma silk.

The axial orientation distribution of the crystallites about the fiber axis was determined from azimuthal broadening (FWHM) of the (200)/(120) equatorial reflections. Orientational order can be expressed quantitatively in terms of Herman's orientation function $f_c = (3 \langle \cos 2\phi \rangle - 1)/2$, where ϕ is the angle between the c axis and the fiber axis.[46] The parameter f_c is 0 for no preferred orientation in fibers and 1 if all crystallites are perfectly aligned with respect to each other. $\langle \cos^2 \phi \rangle$ can be obtained from azimuthal width of the two strongest equatorial reflections - (200)/(120) based on the equation $\langle \cos^2 \phi \rangle = 1 - A \langle \cos^2 \phi_1 \rangle - B \langle \cos^2 \phi_2 \rangle$, where A= 0.8, B = 1.2. Figure 11a is a plot of the intensity as a function of angle measured from the equator at the radial position of the strong (120) peak, and Figure 11b is the corresponding plot at the position of the (200) peak. Both (120) and (200) peaks have been fitted with two Gaussian functions, one wider than the other, according to the structural model used in previous studies on spider silks.[40,42] In the case of the (120) peak profile, small side-peaks due to residual intensity from the (201) reflections were treated as individual Gaussian functions. The degree of orientation of both the crystalline and the oriented amorphous material have been determined from their FWHM as described above. The calculated value of the orientation function for the crystalline material is $f_c \approx 0.979$ and ≈ 0.479 for the oriented amorphous material. As a comparison, for Ma silk from *Nephila clavipes*, the degree of orientation for crystallites and amorphous component are $f_c \approx 0.981$ and ≈ 0.808, respectively.[40] Figure 9b shows the WAXD pattern for the supercontracted fiber (0.6 times its original length). The crystalline reflections are smeared out considerably in the supercontracted fiber. Azimuthal scans at the positions of the equatorial reflections (120) and (200) are shown in Figures 12a and 12b, respectively. The subsidiary peaks (201) overlap considerably with the (120) peak. The position of the subsidiary peaks is fixed to the value of the well-oriented fibers as a constraint in the fitting process. The fitting procedure and the calculation of the orientation parameters have been performed in the similar manner as for the native silk (Figures 10 a and b). The values for orientation function for the crystallites is $f_c \approx 0.926$ and ≈ 0.703 for the oriented amorphous material in the supercontracted silk. The WAXD pattern of the fiber stretched from the supercontracted state to 0.9 times its original length, is shown in Figure 9c. The pattern looks similar to that of the native fiber (Figure 9a). The radial positions of the crystalline peaks, the crystallite sizes and crystallinity in the restretched fiber are about the same for the supercontracted fiber. Azimuthal scans of the equatorial (120) and (200) are shown in Figures 12c and 12d. The values for orientation function for the crystallites is $f_c \approx 0.978$ and ≈ 0.740 for the oriented amorphous material in the restretched silk.

DISCUSSION

The natural abundance BW Ma silk $^1H \rightarrow ^{13}C$ CP-MAS NMR spectrum (Figure 3) exhibits resonances similar to those previously observed in the most studied spider dragline silk, *Nephila* species.[3,21,28] The Ala C$_\beta$ peak spans the chemical shift range from helical to β-sheet conformation, indicating that Ala is in both structural conformations, as previously observed in *Nephila* species silk.[3,28,47] Although $^1H \rightarrow ^{13}C$ CP-MAS NMR spectra of BW MA silk is similar to that of *N. clavipes* major silk, distinct detailed differences in the secondary structure of BW MA silk are expected when comparing the consensus primary amino acid sequences as shown in Figure 1. Both MaSp1 and MaSp2 from BW contain different amino acid motifs when compared to MaSp1 and MaSp2 from *N. clavipes*.[13,19,20] Specifically, BW contains a Gly-Pro-Ser motif, as well as a Gln-Gln-Ala-Tyr motif[13] that are not found in *N. clavipes* silk.

The three Ala C_α-C_β correlations observed in the 2D ^{13}C-^{13}C NMR correlation spectrum provides evidence that Ala is in three different secondary structures in BW Ma silk (Figure 4). The strongest chemical shift correlation between 20.4 ppm and 48.6 ppm is consistent with a β-sheet structure. Like *N. clavipes* Ma silk, large portions of the proteins MaSp1 and 2 in BW silk contain poly(Ala$_n$) motifs. These poly(Ala$_n$) motifs have been shown to form β-sheets in numerous other spider silks and are assigned to a β-sheet in the case of BW dragline silk. The other two Ala C_α-C_β correlation shifts are indicative of two distinct non-β-sheet components. The first correlation with chemical shift of 17.0 ppm and 49.4 ppm corresponds to Ala incorporated in a disordered 3_1-helix and is attributed to the Gly-Gly-X motif.[47] Ala in a 3_1-helix has also been seen in *N. clavipes* major silk.[21] Both the poly-Ala and Gly-Gly-Ala motifs are found in BW and *N. clavipes* major silk, therefore it is not surprising to observe similar secondary structures in the two silk. The third distinct Ala component observed with chemical shift 16.4 ppm and 52.4 ppm more closely correlates to an α-helix than random coil or a β-turn structure.[48,49] This secondary structure has not previously been observed in *N. clavipes* major silk. One Ala containing motif found in black widow silk that is not found in *N. clavipes* silk is Gln-Gln-Ala-Tyr motif,[13] which is found in the MaSp2 protein. There is a Gln-Gln motif in *N. clavipes* silk, however it does not contain Ala.[20] Therefore, the Gln-Gln-Ala-Tyr motif is assigned to a α-helix secondary structure.

Proline chemical shifts were extracted from the 2D ^{13}C-^{13}C correlation spectra with τ_{mix} = 150 ms by taking a projection in *F2* at the Pro C_α. The difference in chemical shift between the Pro C_β and the Pro C_γ (Δ β/γ) are know to be an indicator of secondary structure in proteins.[50] The Δ β/γ for BW Ma silk is 4.7 ppm. This is indicative of a β-turn structure similar to the one found in elastin. This structure has most recently been observed in the Ma silk from *A. aurantia* and is believed to be correlated to the elongation properties of silk fibers.[51]

As seen in Figure 5B, two distinct Gly C_α peaks observed in the projection taken at 19.3 ppm in *F2* from the 2D ^{13}C-^{13}C correlation spectrum with a τ_{mix} = 1 s show that Gly assumes two different secondary protein structures. The downfield peak with chemical shift of 43.0 ppm is indicative of Gly in a β-sheet structure.[52] As shown in Figure 5A, this 43.0 ppm resonance correlates to the Ala C_β peak at 20.5 ppm, which is also assigned to a β-sheet structure, indicting that Gly terminating the poly-Ala runs are incorporated into a β-sheet. Gly incorporated in a β-sheet has been attributed to the poly(Gly-Ala) motif found predominately in MaSp1 of dragline silk, minor spider silk, and silkworm silk.[19,21,53-55] BW MaSp1 and 2 both contain poly(Gly-Ala) that borders poly-Ala.[13] The Gly component in a β-sheet conformation in BW Ma silk can then be ascribed to these poly(Gly-Ala) regions flanking poly-Ala.

The Gly C_α peak with chemical shift of 41.3 ppm, seen in Figure 5, is consistent with a 3_1-helix secondary structure. A 3_1-helix structure has also been observed in *N. clavipes* major silk and is attributed to the Gly-Gly-Ala motif that does not flank poly-Ala.[21] In BW Ma silk, Gly is found in Gly-Gly-X (X is Ala, Gln, or Tyr) in MaSp1 and in MaSp2, X is primarily Ala. Like the *N. clavipes* silk, these non-β-sheet Gly in BW Ma silk are ascribed to the Gly-Gly-X regions. In the 2D correlation, Gly C_α of the 3_1-helix exhibits two correlations to the Ala C_β, one to the β-sheet peak and one to the peak shoulder that is in a helical structure. This correlation indicates that Gly is spatially close to both the Ala β-sheet components, as well as the Ala found in non-β-sheet structures, suggesting that these Gly are incorporated in the interface between the crystalline and non-crystalline portions of the silk. This is the first evidence observed utilizing solid-state NMR that shows Gly as the intermediate between the helical and β-sheet regions. This is consistent with our primary amino acid motif model of silk.

A refocused INADEQUATE NMR experiment of ^{13}C-enriched BW Ma silk is shown in Figure 6. This is the first example in spider dragline silk where Ser has been observed in three different secondary structures (Figure 6). The first Ser DQ chemical shift pair of Ser C_α (54.6 ppm)-C_β (64.0 ppm) is indicative of a β-sheet secondary structure. These are assigned to Ser in BW Ma silk found within or terminating the poly-Ala or poly(Gly-Ala) regions that make up the β-sheets found in spider silk. The Ser C_α (55.6 ppm)-C_β (62.5 ppm) correlations most closely match a helical structure. The Gly-Ser-Gly motif found in MaSp2 of BW Ma silk most likely accounts for the Ser that is incorporated into a helical structure, with the exception of Gly-Ser-Gly terminating poly-Ala. The last correlation of Ser C_α (57.6 ppm)-C_β (60.8 ppm) corresponds to Ser in a turn-like structure.[52] The final Ser motif of Gly-Pro-Ser almost certainly explains the turn structure found in silk. Pro is known to form β-turn secondary structures,[50,56] therefore it is likely that the Ser following the Pro will be incorporated into a β-turn structure.

Figure 7 shows the $^{13}C[^1H]$ DD-MAS NMR spectrum collected with a 1 s recycle delay of ^{13}C-enriched BW Ma silk in both the native dry state, as well as the wetted state. Utilizing a 1 s delay enhances any resonance that becomes mobile when wet, indicating a non-β-sheet structure. This spectrum can be used to confirm the Ser chemical shift assignments established in the refocused INADEQUATE experiment. When wet, the upfield (high ppm) shoulder of the Ser C_β peak exhibits an increase in intensity, while no change is observed on the downfield (low ppm) side. The increase in intensity has been observed in other silks and is attributed to water penetrating non-β-sheet regions of the silk and plasticizing these regions.[57] Any resonance that becomes mobile will exhibit an intensity enhancement and therefore be discounted from being in rigid, β-sheet secondary structures. Therefore, this indicates that the upfield side of the Ser C_β is in a non-β-sheet conformation and the downfield side is in a β-sheet. Two peaks are enhanced in chemical shift range of Ser C_α. The chemical shift of the downfield resonance of 55.8 ppm is too high to be any other amino acid but Ser C_α and is therefore ascribed to Ser in a non-β-sheet structure. The chemical shift of the more upfield peak at 52.7 ppm overlaps the range of Gln C_α. Gln is one of the amino acids in silk that has been shown to become plasticized by water.[28,57] Therefore, the more upfield peak in the Ser C_α region that becomes enhanced is most likely due to the Gln C_α becoming plasticized. This peak overlaps the Ser C_α chemical shift range for a β-sheet structure and accounts for the change that is observed in the DD-MAS spectrum of the wet silk that is not expected if the peak was only Ser C_α in a β-sheet structure, confirming the chemical shift assignment in the refocused INADEQUATE.

The assigned secondary structures of Gly, Ala, and Ser in BW dragline silk were quantitatively correlated to the primary acid sequence. All Ala in poly-Ala and all Gly or Ser terminating either side of poly-Ala are counted as β-sheet. Using this scheme, the primary amino acid sequences predict that 88% of Ala, 42% of Gly, and 41% of Ser are incorporated into a β-sheet structure. When comparing the predicted % β-sheet to the percentages determined by solid-state NMR, very good agreement is observed. The percentages for Ala and Gly incorporated into a β-sheet were extracted by fitting the carbonyl peak of a fully relaxed DD-MAS spectrum. From this data, 88% of Ala and 40% of Gly is incorporated into a β-sheet. From this data, it is clear that BW Ma silk has a larger amount of Ala and Gly in β-sheet secondary structure compared to *N. clavipes*, which has 82±4% of Ala and 28±5% of Gly in a β-sheet.[27] The total amount of β-sheet in BW Ma silk, which includes all Ala, Gly, and Ser incorporated into a β-sheet is 47%. This is over 10% more than *N. clavipes* major silk (34% β-sheet), but very similar to *N. clavipes* minor silk (45% β-sheet).[27] BW Ma silk has a higher toughness then *N. clavipes* major silk.[15] This mechanical property could be related to the total percent of β-sheet in the silk.

X-ray diffraction results show that structure of BW Ma silk is similar to that of Ma silks from other species, with crystallites comprising of stacked β-sheets and an amorphous region comprising of oriented and isotropic components.[40] The degree of crystallinity is higher in BW Ma silk compared to other species like *Nephila clavipes* and*Argiope auruntia.* While the crystallites are well oriented with respect to the fiber axis with $f_c \approx 0.979$, the oriented amorphous component is very poorly oriented ($f_{am} \approx 0.479$). In other silks studied by the authors, the degree of orientation of the amorphous component decreased with increase in crystallinity and BW Ma silk seems to follow this trend.[40]

In the supercontracted state, there is a loss of orientation of the nano-crystallites, similar to the behavior seen in Ma silk from *Nephila clavipes* (NC).[58] However, the degree of orientation of the oriented amorphous component increases in the supercontracted state for BW silk (from 0.479 to 0.703), a trend not seen in earlier studies on other Ma silks.[42,58]

In the restretched fiber, while the crystallites become well oriented again with a f_c value of 0.978, the oriented amorphous component also becomes better oriented with a value of 0.740. These results show that there are structural changes associated with supercontraction and that the restretched fiber does not completely return to the structure of the native fiber. However, in the restretched fiber, the β-sheet crystallites retain their structure close to that of the native fiber, but there is a permanent change in the structure of the amorphous component caused by supercontraction. Hysteresis in mechanical properties have been observed in dragline silk, when the silk fiber is restretched to its original length after supercontraction and the results have been attributed to change in nanostructure of the fiber.[59] The results from this work have demonstrated that the oriented amorphous component has a significant structural change in the supercontracted state and that the structure of the restretched fiber does not completely return to that of the native fiber.

CONCLUSION

Various solid-state NMR and XRD experiments were utilized to characterize the secondary structure of the only fully sequenced spider dragline silk proteins, MaSp1 and MaSp2 from the black widow spider, *Latroductus hesperus*. Utilizing chemical shift assignments and correlations from 2D through-space ^{13}C -^{13}C correlation spectrum and 2D through-bond ^{13}C-^{13}C refocused INADEQUATE experiments, significant progress has been made in determining the secondary structures of various amino acid motifs that together make up the black widow major silk. Alanine was determined to be in three distinct secondary conformations, β-sheet, 3_1-helix, and an α-helix. Proline was ascribed to a β-turn structure. Glycine was found to be in two discrete secondary structures, a β-sheet and a 3_1-helix. Additionally, the Gly in a helical conformation was determined to be spatially close to both the Ala in a β-sheet and Ala in a helical structure indicating that the helical portion of Gly is an intermediate between the two. Finally, Ser was found to also be in three distinct structures, a β-sheet, helix, and turn.

These secondary structural assignments were then quantitatively correlated and related to the primary amino acid sequences for both MaSp1 and MaSp2. Utilizing solid-state NMR, this study determined that 88% of Ala, 40% of Gly, and 42% of Ser in BW Ma silk is incorporated into a β-sheet secondary structure. These values are in good agreement with the predicted values from the primary amino acid sequences of 88%, 42%, and 41% for Ala, Gly, and Ser, respectively. In total, solid-state NMR shows that 47% of all BW Ma silk is incorporated into a β-sheet secondary structure. This is further confirmed from XRD analysis to also give a crystalline fraction of 47%. This study has characterized and quantified the secondary structures found in BW Ma silk. The data was then correlated to

the amino acid sequence for spider silk, providing a molecular structure template for the production of synthetic fibers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCE

1. Teulé F, Cooper AR, Furin WA, Bittencourt D, Rech EL, Brooks A, Lewis RV. Nat. Protoc. 2009; 4:341. [PubMed: 19229199]

2. Vollrath F, Porter D. Polymer. 2009; 50:5623.

3. Simmons A, Ray E, Jelinski L. Macromolecules. 1994; 27:5235.

4. Michal C, Jelinski L. J. Biomol. NMR. 1998; 12:231. [PubMed: 9751996]

5. van Beek JD, Hess S, Vollrath F, Meier BH. PNAS. 2002; 99:10266. [PubMed: 12149440]

6. van Beek JD, Meier BH. J. Magn. Reson. 2006; 178:106. [PubMed: 16243550]

7. Scheller J, Gührs K, Grosse F, Conrad U. Nat. Biotechnol. 2001; 19:573. [PubMed: 11385464]

8. Lazaris A, Arcidiacono S, Huang Y, Zhou JF, Duguay F, Chretien N, Welsh EA, Soares JW, Karatzas CN. Science. 2002; 295:472. [PubMed: 11799236]

9. Zhang Y, Hu J, Miao Y, Zhao A, Zhao T, Wu D, Liang L, Miikura A, Shiomi K, Kajiura Z, Nakagaki M. Mol. Biol. Rep. 2008; 35:329. [PubMed: 17525867]

10. O'Brien JP, Fahnestock SR, Termonia Y, Gardner KH. Adv. Mater. 1998; 10:1185.

11. Seidel A, Liivak O, Jelinksi L. Macromolecules. 1998; 31:6733.

12. Brooks AE, Stricker SM, Joshi SB, Kamerzell TJ, Middaugh CR, Lewis RV. Biomacromolecules. 2008; 9:1506. [PubMed: 18457450]

13. Ayoub NA, Garb JE, Tinghitella RM, Collin MA, Hayashi CY. Plos One. 2007:1.

14. Gosline J, DeMont E, Denny M. Endeavour. 1986; 10:37.

15. Swanson BO, Blackledge TA, Beltran J, Hayashi CY. Appl. Phys. A. 2006; 82:213.

16. Hayashi C, Shipley N, Lewis R. Int. J. Biol. Macromol. 1999; 24:271. [PubMed: 10342774]

17. Casem ML, Turner D, Houchin K. Int. J. Biol. Macromol. 1999; 24:103. [PubMed: 10342753]

18. Gatesy J, Hayashi C, Motriuk D, Woods J, Lewis R. Science. 2001; 291:2603. [PubMed: 11283372]

19. Xu M, Lewis R. PNAS. 1990; 87:7120. [PubMed: 2402494]

20. Hinman M, Lewis R. J. Biol. Chem. 1992; 27:19320. [PubMed: 1527052]

21. Holland GP, Creager MS, Jenkins JE, Lewis RV, Yarger JL. J. Am. Chem. Soc. 2008; 130:9871. [PubMed: 18593157]

22. Lawrence BA, Vierra CA, Moore AF. Biomacromolecules. 2004; 5:689. [PubMed: 15132648]

23. Gould SAC, Tran KT, Spagna JC, Moore AMF, Shulman JB. Int. J. Biol. Macromol. 1999; 24:151. [PubMed: 10342759]

24. Trancik JE, Czernuszka JT, Bell FI, Viney C. Polymer. 2006; 47:5633.

25. Holland GP, Lewis RV, Yarger JL. J. Am. Chem. Soc. 2004; 126:5867. [PubMed: 15125679]

26. Izdebski T, Akhenblit P, Jenkins JE, Yarger JL, Holland GP. Biomacromolecules. 2010; 11:168. [PubMed: 19894709]

27. Jenkins JE, Creager MS, Lewis RV, Holland GP, Yarger JL. Biomacromolecules. 2010; 11:192. [PubMed: 20000730]

28. Holland GP, Jenkins JE, Creager MS, Lewis RV, Yarger JL. Biomacromolecules. 2008; 9:651. [PubMed: 18171016]

29. Marcotte I, van Beek B, Meier B. Macromolecules. 2007; 40:1995.

30. Brooks AE, Steinkraus HB, Nelson SR, Lewis RV. Biomacromolecules. 2005; 6:3095. [PubMed: 16283732]

31. Liivak O, Flores A, Lewis R, Jelinski L. Macromolecules. 1997:7127.

32. Parkhe A, Seeley S, Gardner K, Thompon L, Lewis R. J. Mol. Rec. 1997; 10:1.

33. Ashida J, Ohgo K, Komatsu K, Kubota A, Asakura T. J. Biomol. NMR. 2003; 25:91. [PubMed: 12652118]

34. Takegoshi K, Nakamura S, Terao T. J. Chem. Phys. 2003; 118:2325.

35. Cadars S, Sein J, Duma L, Lesage A, Pham TN, Baltisberger JH, Brown SP, Emsley L. J. Magn. Reson. 2007; 188:24. [PubMed: 17588789]

36. Work RW, Emerson PD. J. Arachnol. 1982; 10:1.

37. Bennett A, Rienstra CM, Auger M, Lakshmi KV, Griffin RG. J. Chem. Phys. 1995; 103:6951.

38. Massiot D, Fayon F, Capron M, King I, Le CalvÈ L, Alonso B, Durand JO, Bujoli B, Gan Z, Hoatson G. Magn. Reson. Chem. 2002; 40:70.

39. Holland GP, Jenkins JE, Creager MS, Lewis RV, Yarger JL. Chem. Commun. 2008:5568.

40. Sampath S, Isdebski T, Jenkins JE, Ayon JV, Henning RW, Orgel J, Antipoa O, Yarger JL. Soft Matter. 2012; 8:6713. [PubMed: 23569461]

41. Yang Z, Grubb DT, Jelinski LW. Macromolecules. 1997; 30:8254.

42. Riekel C, Muller M, Vollrath F. Macromolecules. 1999; 32:4464.

43. Bram A, Branden CI, Craig C, Snigireva I, Riekel C. Journal of Applied Crystallography. 1997; 30:390.

44. Warwicker JO. Journal of Molecular Biology. 1960; 2:350. [PubMed: 13783274]

45. Du N, Liu XY, Narayanan J, Li LA, Lim MLM, Li DQ. Biophysical Journal. 2006; 91:4528. [PubMed: 16950851]

46. Stein, RS. Structure and Properties of Oriented Polymers. Wiley; New York: 1975.

47. Shoji A, Ozaki T, Saito H, Tabeta R, Ando I. Macromolecules. 1984; 17:1472.

48. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. J. Biomol. NMR. 1995; 5:67. [PubMed: 7881273]

49. Asakura T, Nakazawa Y, Ohnishi E, Moro F. Protein Sci. 2005; 14:2654. [PubMed: 16195552]

50. Kricheldorf HR, Muller E. Macromolecules. 1983; 16:615.

51. Liu Y, Sponner A, Porter D, Vollrath F. Biomacromolecules. 2008; 9:116. [PubMed: 18052126]

52. Ishida M, Asakura T, Yokio M, Saito H. Macromolecules. 1990; 23:88.

53. Colgin M, Lewis R. Protein Sci. 1998; 7:667. [PubMed: 9541398]

54. van Beek JD, Beaulieu L, Schäfer H, Demura M, Asakura T, Meier BH. Nature. 2000; 405:1077. [PubMed: 10890452]

55. Asakura T, Ito T, Okudaira M, Kameda T. Macromolecules. 1999; 32:4940.

56. Kricheldorf H, Müller D. Int. J. Biol. Macromol. 1984; 6:145.

57. Yang Z, Liivag O, Seidel A, LaVerde G, Zax D, Jelinski L. J. Am. Chem. Soc. 2000; 122:9019.

58. Grubb DT, Ji GD. International Journal of Biological Macromolecules. 1999; 24:203. [PubMed: 10342766]

59. Ene R, Papadopoulos P, Kremer F. Soft Matter. 2009; 5:4568.

Latrodectus hesperus

    Major Silk

        MaSp1: GGAGQGGQGGYGQGGYGQGGAGQGGGAGAAAAAA

        MaSp2: GGAGPGRQQAYGPGGSGAAAAAAAA

Nephila clavipes
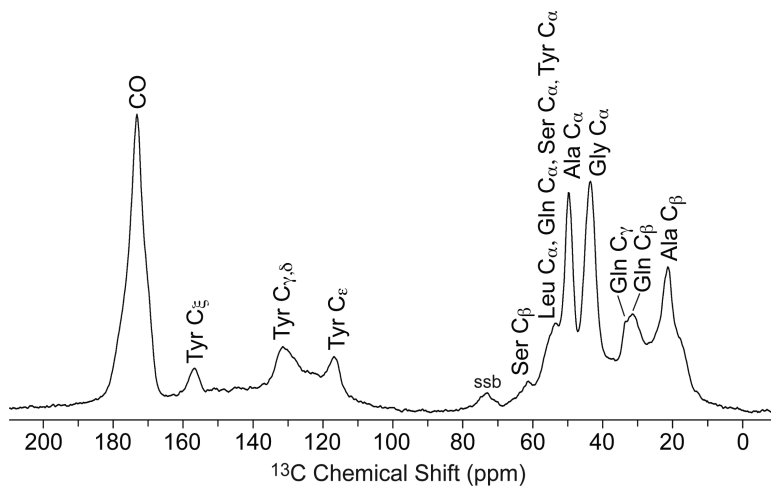
    Major Silk

        MaSp1: GGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAAA

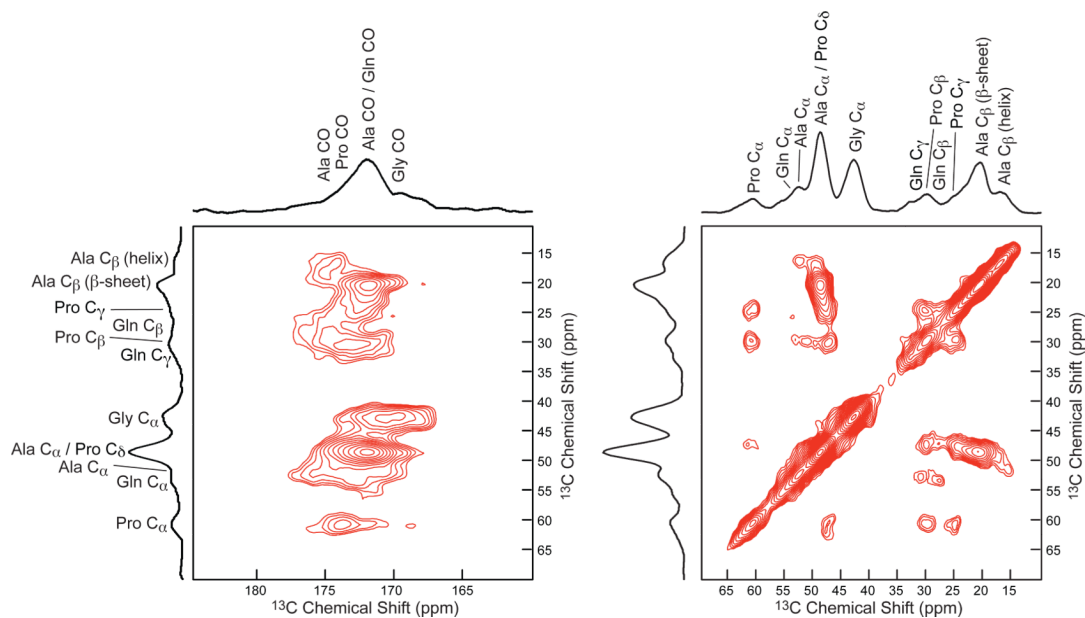   MaSp2: GPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAAAA

Figure 1.

Consensus amino acid sequences for major ampullate spidroin 1 (MaSp1) and major ampullate spidroin 2 (MaSp2) of *Latrodectus hesperus* and *N. clavipes* dragline silk.

MaSp1

MTWSTRLALSFLFVLCTQSLYALAQANTPWSSKANADAFINSFISAASNTGSFSQDQMEDMSLIGNTLMAAMDNMGGRITPSKLQALDMAFASSVAEIAASEGGDLGVTTNAIADALTS
AFYQTTGVVNSRFISEIRSLIGMFAQASANDVYASAGSSGGGGYGASSASAASASAAAPSGVAYQAPAQAQISFTLRGQQPVYGQGGAGPGGAGAAAAAAAAAGGAGQGGQGGGYGQGGY
GQGGAGQGGSGAAAAAAAAAGGTGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGTGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGGGSGAAAAAAAAGGAGQ
GGGQGGGYGQGGYGQGGAGQGGAGAAAAAAAAGGAGQGGYGRGGAGQGGAAAAAAAGAGQGGYGGQGAGQGGSGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGSGAAAAAAAAGGAGQG
GQGGGYGQGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGG
AGQGGAAAAAAAAAGGAGQGGYGRGGAGQGGAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAG
QGGAAAAAAAAGGAGQGGYGRGGAGQGGAAAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAAGGAGQGGQGDYGRGGYGQGGAGQGGAGAAAAAAAAAGGAGQGGQGGGYGQGGYGQGGA
GQGGAAAASAAAAGGAGQGGYGRGGAGQGGAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGAAAAAATAAGGAGQGGQGGGYGQGGYGQGGAG
QGGAAAAAAAAGGAGQGGYGRGGAGQGGAAAAAAAAGAGQGGYGGQGAGQGGAGAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAG
QGGAAAAAAAAAGGAGQGGYGRGGAGQGGAAAAGAGQGGYGGQGAGQGGAGAAAAASRGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAGQGGGA
AAAAAAGGAGQGGYGRGGAGQGGAAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAGQGGAAAA
AAAAAAAGGAGQGGYGRGGAGQGGAAAAAAAAGSGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAGQGGAA
AAAAAAAGGAGQGGYGRGGAGQGGAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGTAAAAAAGGAGQGGQGGGYGQGGYGQGGAGQGGAAA
AAAAAAGGAGQGGYGRGGAGQGGAAAAAAAAGGAGQGGYGGQGAGQGGAGAAAAAAAAGGASQGGQGGGYGQGDYGQGGAGQGGA
AAAAAAAGGAGQGGYGRGGAGQGGAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGARGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGTGQGGAAAA
AAAAAGGAGQGGYGRGGAGQGGAAAAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAGQGG
AAAAAAAAGGAGQGGYGRGGAGQGGAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAGQGG
AAAAAAAAGGAGQGGYGRGGAGQGGAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGNGQG
GAGQGGAAAAAAAAGGAGQGGYGRGGAGQGGAAAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGASQGGQGGGYGQGDYGQ
GGAGQGGAAAAAAAAGGAGQGGYGRGGAGQGGAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGARGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGA
GQGGAAAAAAAAGGAGQGGYGRGGAGQGGAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGARGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGGAGQGGA
GGAAAAAAAAAVGGAGQGGYGRGGAGQGGAAAAAAAAAGSGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAAAAGGAGQGGQGGGYGQGGYGQGG
AGQGGAAAAAAAAGGAGQGGYGRGGAGQGGAAAAAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGGYGRGGYGQGGAGQGGAGAAAAAGGAGQGGQGGGYGQGGYGQGGAGQ
GGAAAAAAAAAAGGAGQGGYGGYGQGQGGAGAAAAAASGPGQIYYGPQSVAAPAAAAASALAAPATSARISSHASALLSNGPTNPASISNVISNAVSQISSSNPGASACDVLVQALLELV
TALLTIIGSSNIGSVNYDSSGQYAQVVTQSVQNAFA

MaSp2

MTTMNWSTRLVLSILVVLCTQSLCALGQANTPWSSKENADAFIGAFMNAASQSGAFSSDQIDDMSVISNTLMAAMDNMGGRITQSKLQALDMAFASSVAEIAVADGQNVGAATNAISDA
LRSAFYQTTGVVNNQFITGISSLIGMFAQVSGNEVSYSSAGSSSAAASEAVSAGQGPAAQPVYAPSGASAAAAAASGAAPAIQQAYERGGSGSAAAAAGSGPSGYGQGAGGPGGAGAAA
GAAAAGGSGPGGYGQGPAAYGPSGPSGQQGYGPGGSGAAAAAAAAAGSGPSGYGPGAGGPGGAGAAAAAAAAAGGSGPGGYGQGQASYGPSGPSGQQGYGPGGSGAAAAAAAAAGSGPSG
YGPGAAAAAAAGSAGPGTQQGYGPGGSGAAAAAGSGPRGYGPRGPGGAGAAATAARGSGPGGYGQGPAGYGTSGPSRQQGYGPGGSGAAAAAAAAAGGAGPGRQQGYGPGGSGAAAATA
AGGPGYVGQQRYGPGGAGAAAAAAAAGSAGPSRQQAYGPGGSGPAAATAAGSGPSGYGPGASGPVGADAAAAAATGSAGPGRQQAYGPGEGAAAAAAAGSAGPGRQLGYGPGGSGAAAA
AAAGGPGYGGQQGYGPGGAGAAAAAAAAGGAGPGRQQTYGPGGSGAAATAAGSGPGGYGQGPSGYGPSGPGGQQGYGPGGSGAAAAAAGEAGPGRQQGYGPRGSGAAAAAAAAGGPGYG
GQSGYGPGGAGAAAAAAAAGGAGPGRQQEYGPGGSGAAAAAAAAAGSGPSGYGPGAAGPIGPGGAGAAAAGGSGPVGYGQGPSGYGASGTGGEQDYGPGGSGAAAAAAAAAGGAGPGRQQ
GYGPGGSGAAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAAGGAGPGRQQPYGPGGAGAAAAAAAGGSGPGGYGQGPSGYSASGPGGQQGYGPGGSGLAAA
AAAAGGTGYGGQQGYGPGGAGAAAAAAAAGAGPGRQQEYGPGGTGAAAAAAAAAGSGPSGYGQGAAGPSGPGGEGTAAAAAAAGGSGPGGYGQGPSGYSASGPGGQQGYGPGGSGLAAA
AAAAAGGAGTGRQQGYGPGGSGAAAAAAAAVGPGYGGQQGYGPGGAGAAAAAAAAGGAGPGRQQAYGPGGSGATAAAAVAGSGPSGYGPGGAGAAAAAAAAGGAGPGRQQAYGPGGSGAAAA
AASGAGPGRQQVYGPGGSGAAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAAGGAGQGTRQAYGPGGSGAAAAAAGPGPSGYGPGAAGPSGPGLAGAAAAAAAAGGSGPGGNGQRPSGYGQS
GTGGQQGYGPGGSGAAAAAAAAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAAGGAGPGTQQAYGPGGSGAAAAAAAAAAAGGAGPGRQQGYGPGSSGAAAAA
AAGGPGYGGQQGYGPGGAGAAAAAAAAGGAGRQQAYGPGGSGAAAAGSGPSGYEPGAAGPGGAGAAAAAAAVGAGPGRQQAYGGGGSGAVAAAAAGGPGYGGQQGYEQGGAGAASAAA
AGGEGPARQQAYGPGGSGAAAAAAAGGAGPGRQQGYGPGGSGAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAGGAGPGRQQAYGPGGSGAAAAAAAGTGPSGYGPGAAGPGGAGAAAAA
AAGSAGPGRQQAYGPGGSGAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAGGAGPGTQQAYGPGGSGAAAAAAAAAGGAGPDRQQGYGPGSSGAAAAAAAGGPGYGGQQGYGPGGAGAA
AAAAAAGPGPSGYGPGGAGAAAAAAAAGSGPGGYGQGPSGYSASGPGGQQGYGPGPGGSGLAAAAAAAAGGAGTGRQQGYGPGGSGAAAAAAAVGPGYGGQQGYGPGGAGAAAAAAAGG
AGPGRQQAYGPGGSGATAAAAAAGSGPSGYGPGGAGAAAAAAAGAGPGRQQAYGPGGSGAAAAAASGAGPGRQQVYGPVGSGAAAAAAAGGPGYGGQQGYEQGGAGAAAAAAAGGAGGQ
GTRQAYGPGGSGAAAAAAGPGPSGYGPGAAGPSGPGLAGAAAAAAAAGGSGPGGNGQRPSGYGQSGPGGQQGYGPGGSGAAAAAAAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGYGGQ
QGYGPGGAGAAAAAAAGGAGPGTQQAYGPGGSGAAAAAAAAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAGGAGAGRQQAYGPGGSGAAAAGSGPSGYE
SGAAGPGGAGAAAAAAVGAGPGRQQAYGQGGSGAVAAAAAGGPGYGGQQGYEQGGAGAAAAAAAGGEGPARQQAYGPGGSGAAAAAAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGYGG
QQGYGPGGAGAAAAAAAGGAGPGRQQAYGPGGSGAAAAAAAGTGPSGYGPGAAGPGGAGAAAAAAAGGAGPGRQQAYGPGGSGAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAGGAGP
GTQQAYGPGGSGAAAAAAAAAGGAGPDRQQGYGPGSSGAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAAAGPGPSGYGPGGAGAAAAAAAAAGGSGPGGYGQGPSGYGPSGPGGQQGNG
PGGSGAAAAAAAAAGGAGPGRQQGYGPGGAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAAGGAGPGRQQAYGPGGAGAAAAAAAAAAGPGPSGYGPGASGPSGTGGAGAAAAAAAAAGGSG
PGGYGQGASGYGPSGPGGQQGYGPGGSGAAAAAAAAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGYGGPQGYGPGGAGAAAAAAAGGAGPGRQQAYGPGGSGAAAAAGSGPSGYGPGA
AGPGGTGAAAVAAAGGAGPGRQQAYGPGGSGAAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAGGAGPGTQQLYGPGGSGAAAAAAAAAGGSGPSGYGPGAAGPSGPGAGAAAAAAASAGG
SGPGGYGQGPSGYGPTGPVGQQGYGPGGSGAAAAAAAAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGYGGQQGYGPGGAGAAAAVAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGYGGQ
QGYGPGGAGAAAAVAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGYGGQQGYGLGVAGAAAAVAAGGAGPGRQQAYGPGGSGAAAAAAAGSRSGYGPGAAGTGGAGAAAAAAAGGAGSG
RQQAYGPGGSGAAAASAAGGPGYGGQQGYGPGGAGAAAAAAAAGGAGPGTQQAYGPGGSGAAAAAAAAAGSGPGPSGYEPGAAGPSGPAGAGAAAAAAAAAGSGPGGYGQGPSGYGPSGPGG
QQGYGPGGSGAAAAAAAAGGAGPGRQQGYGQGGSGAAAAAAAAGGPGYGGQQVYGPGGAGAAAAVAAGGAGPGRQQAYGPGGSGAAAGSGPSGYGPGAAAAAAAAAGGAGPGRQQAYGPGG
SGAAAAAAAAGGPGYGGQQGYGPGGAGAAAAAAAGSGPGGYGQGPSGYGPSGSGGQGYGQGGSGAAAAAAAAGGAGPGRQQGYGPGSSGAAAAAAAGGPGFGGQQGYGPGGSGAAAAAAAA
GAGPGRQQAYGPGGSGAAAAAAAAAGSGPSGYGPSAAGPSGPGGSGAAGGSGPGGFGQGPAGYGPSGPGGQQGYGPGASGAAAAAAASGSGGYGPSQYVPSSVASSAASAASALSSPTT
HARISSHASTLLSSGPTNAAALSNVISNAVSQVSASNPGSSSCDVLVQALLEIITALISILDSSSVGQVNYGSSGQYAQIVGQSMQQAMG

Figure 2.

The full primary amino acid sequences for *Latrodectus hesperus* major ampullate spidroin 1 and 2 (MaSp1 and 2) proteins that make up the majority (>99 %) of major silk. The colored Ala, Gly, and Ser indicate β-sheet secondary structure. The black indicates non-β-sheet components, including helices, turns, and random coil. GenBank accession no. EF595246 and EF595245.[13]

Figure 3.
$^1H \rightarrow ^{13}C$ CP-MAS NMR spectrum of natural abundance *Latrodectus hesperus* major silk. All amino acids that make up at least 3% of the amino acid composition are observed in the natural silk.
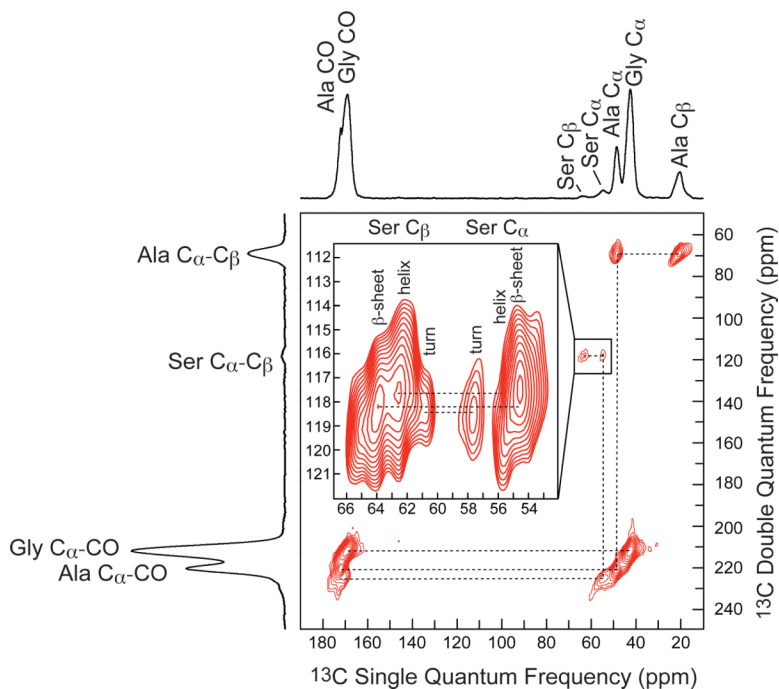
Figure 4.
2D ^{13}C -^{13}C correlation NMR spectrum with τ_{mix} = 150 ms of U-^{13}C/^{15}N-proline enriched *Latrodectus hesperus* major silk provides intramolecular correlations. The chemical shifts for each proline resonance is observed. Three correlations are observed between Ala C$_\beta$-Ala C$_\alpha$ (20.4 / 48.6 ppm, 17.0 / 49.4 ppm 16.4 / 52.4 ppm), indicating that Ala is incorporated into three distinct structural environments.
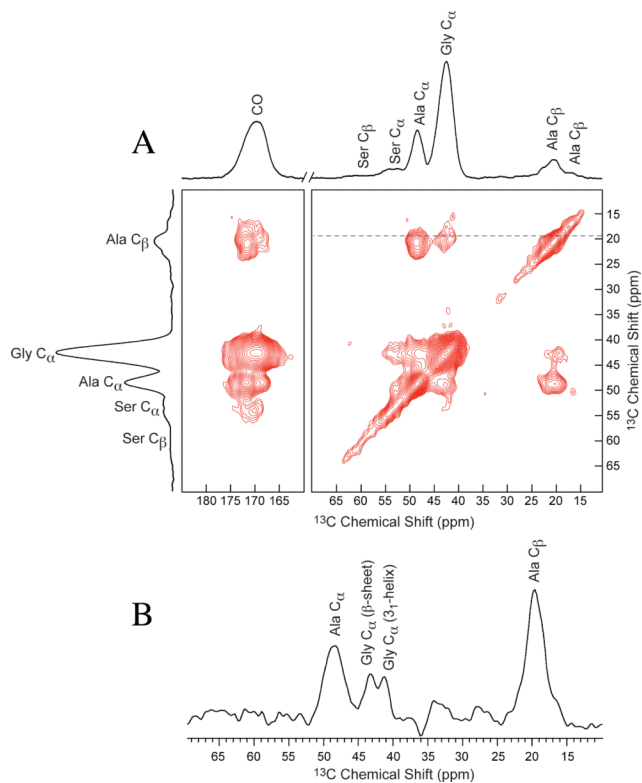
Figure 5.
(A) 2D ^{13}C -^{13}C correlation NMR spectrum with τ_{mix} = 1 s of U-^{13}C/^{15}N-serine enriched *Latrodectus hesperus* major silk provides intermolecular correlations between amino acids. (B) The slice taken at 19.3 ppm, is take between the Ala C_β in a helical structure and β-sheet structure to illustrate both Gly C_α resonances.

Figure 6.
2D ^{13}C-^{13}C through-bond DQ/SQ refocused INADEQUATE NMR spectrum of U-^{13}C/^{15}N-serine enriched *Latrodectus hesperus* major silk. The three distinct serine resonances are assigned to serine incorporated in three distinct secondary structures, a β-sheet, helix, and β-turn.
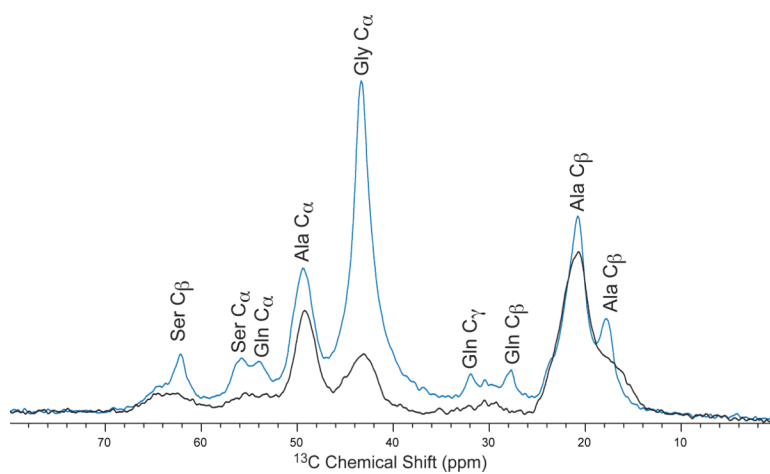
Figure 7.
Directly-detected (DD-MAS) ^{13}C[^{1}H] NMR spectrum with a 1 s recycle delay of U-^{13}C/^{15}N-serine enriched *Latrodectus hesperus* major silk in the native dry state (black) and water-wetted, supercontracted state (blue).
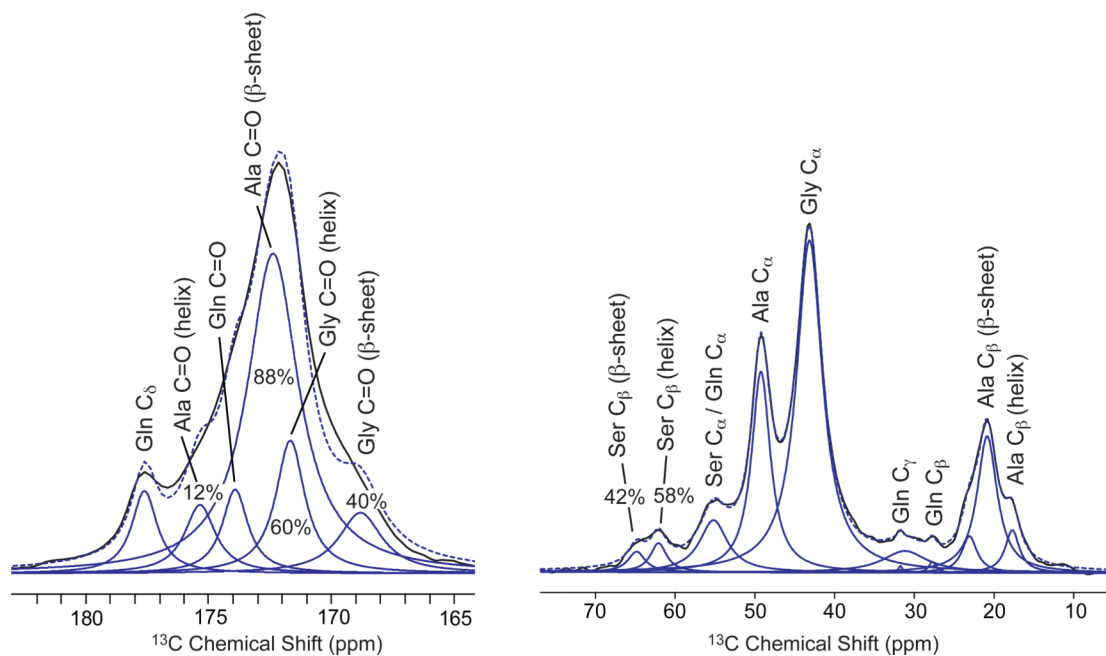
Figure 8.
The fitted carbonyl peaks from a fully relaxed, quantitative $^{13}C[^1H]$ DD-MAS NMR spectrum of *Latrodectus hesperus* major (BW Ma) wet (supercontracted) silk shows the percent of Ala, Gly, and Ser that are incorporated into a β-sheet secondary structure.
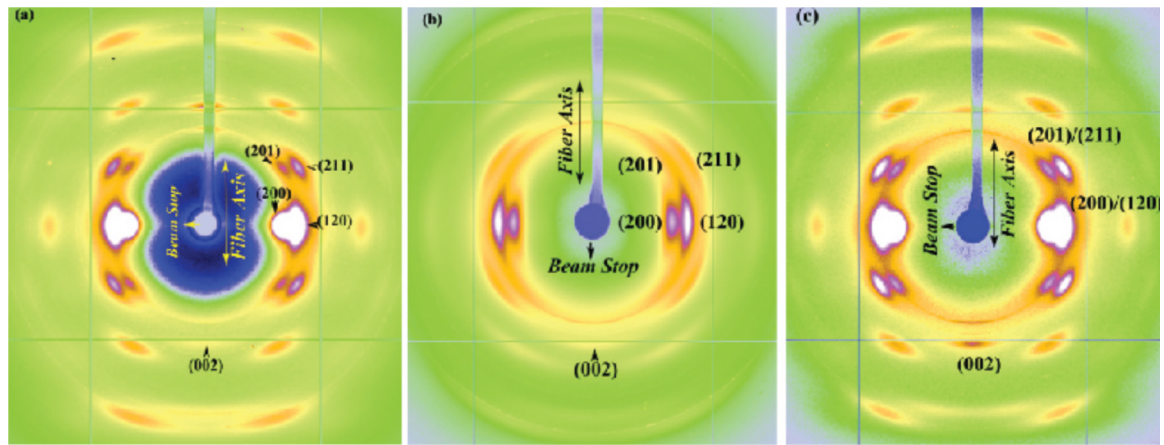
Figure 9.
WAXD pattern of *Latrodectus hesperus* Major Ampullate silk (a) - native fiber, original length L_0 (b) supercontracted fiber due to wetting with water, shrunk to 0.6 L_0, (c) restretched fiber from supercontracted state to 0.9 L_0
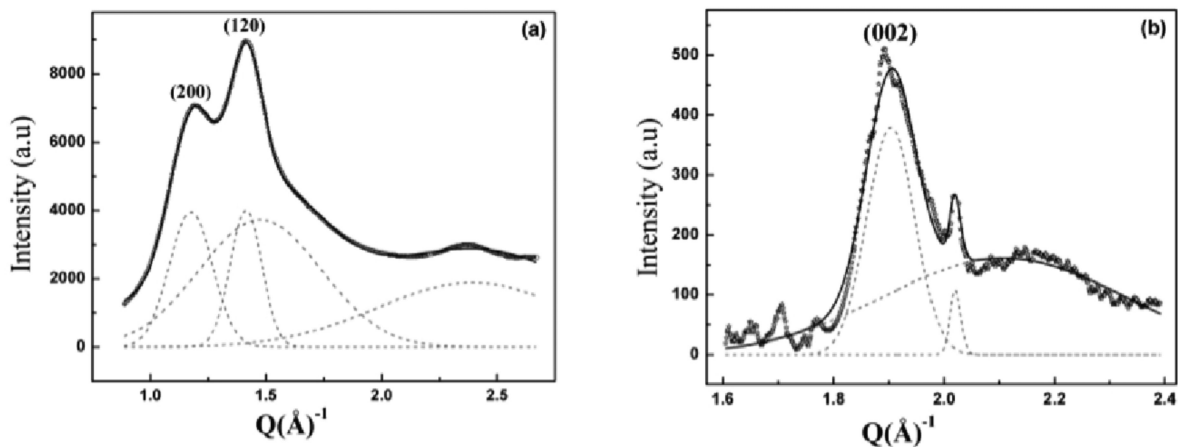
Figure 10.
1D XRD Radial Intensity profile for BW Ma silk fibers of azimuthally integrated (a) (120)/ (200) equatorial peaks (O) data and the (b) (002) meridian peak (★) data. Gaussian fits (--) are used for each component, including the broad amorphous component for *Latrodectus hesperus* MA native silk, (–) Total fit.
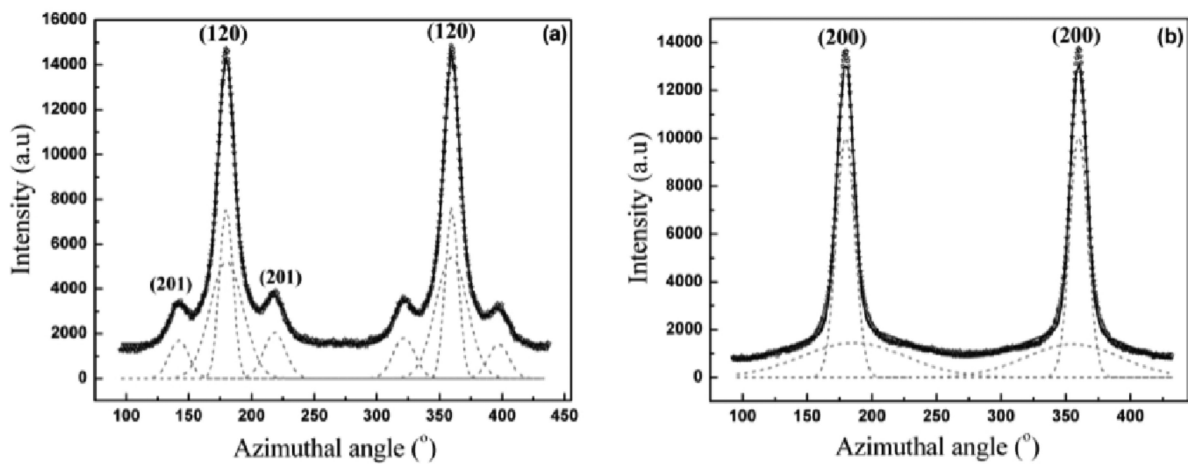
Figure 11.
1D Azimuthal Intensity profile for BW Ma silk fibers of the radially integrated (a) (120) and (b) (200) peak with Gaussian fits.
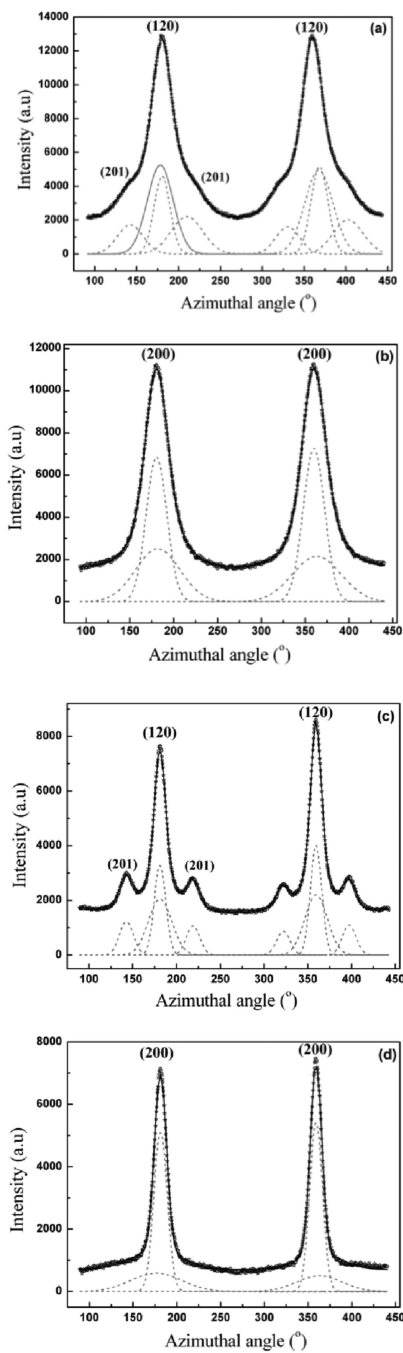
Figure 12.
(a) 1D Azimuthal Intensity profile for BW Ma silk fibers of the radially integrated (a) (120) peak with Gaussian fits for supercontracted silk fibers, (b) (200) peak with Gaussian fits for supercontracted silk fibers, (c) (120) peak with Gaussian fits for restretched silk fibers, and (d) (200) peak with Gaussian fits for restretched silk fibers.

Table 1

Amino acid motif model for quantification of amino acids in specific secondary structures. Percent β-sheet and helical structure from BW Ma silk is predicted from the primary amino acid sequence (AA%) and extracted from NMR data fits (NMR%) (Figure 8).

| | Black Widow, *Latrodectus hesperus, Major Silk* | | | |
| | β-sheet | | Helical | |
	AA %	NMR %	AA %	NMR %
Ala	88	88 ± 3	12	12 ± 3
Gly	42	40 ± 3	58	60 ± 3
Ser	41	42 ± 5	59	58 ± 5