Published in final edited form as:
*Biomacromolecules*. 2010 January 11; 11(1): 192–200. doi:10.1021/bm9010672.

# Quantitative Correlation Between the Protein Primary Sequences and Secondary Structures in Spider Dragline Silks

**Janelle E. Jenkins**, **Melinda S. Creager**, **Randolph V. Lewis**, **Gregory P. Holland**, and **Jeffery L. Yarger**
Department of Chemistry and Biochemistry, Magnetic Resonance Research Center, Arizona State University, Tempe, Arizona 85287-1604, USA, and Department of Molecular Biology, University of Wyoming, Laramie, Wyoming 82071

## Abstract

Synthetic spider silk holds great potential for use in various applications spanning medical uses to ultra lightweight armor, however producing synthetic fibers with mechanical properties comparable to natural spider silk has eluded the scientific community. Natural dragline spider silks are commonly made from proteins that contain highly repetitive amino acid motifs, adopting an array of secondary structures. Before further advances can be made in the production of synthetic fibers based on spider silk proteins, it is imperative to know the percentage of each amino acid in the protein that forms a specific secondary structure. Linking these percentages to the primary amino acid sequence of the protein will establish a structural foundation for synthetic silk. In this study, Nuclear Magnetic Resonance (NMR) techniques are used to quantify the percentage of Ala, Gly, and Ser that form both β-sheet and helical secondary structures. The fraction of these three amino acids and their secondary structure are quantitatively correlated to the primary amino acid sequence for the proteins that comprise major and minor ampullate silk from the *Nephila clavipes* spider providing a blueprint for synthetic spider silks.

### Keywords

Solid-state NMR; Biopolymer; Spider silk; protein structure

## Introduction

Over nearly 400 million years [1], spiders have evolved the ability to produce up to six different silks and one glue-like substance that all possess amazing, yet unique mechanical properties [2, 3]. Major (Ma) ampullate silk is by far the most studied because it has an exceptional combination of mechanical properties including an impressive extensibility and a tensile strength that exceeds steel [2-4]. Much less is known about minor (Mi) ampullate silk, for the primary reason that silk collection is more difficult. Mechanical testing has been performed on multiple species of Ma silk, while fewer mechanical tests have been performed in Mi silk[3, 5-8]. One study comparing Ma and Mi silk from *Argiope argentata* showed that Ma silk is stronger than Mi silk, with an ultimate strength of 1.5 GPa, compared to 0.9 GPa for Mi. However, Mi silk has a larger extensibility[6]. With these varying mechanical properties, both silks could be used for multiple applications, such as synthetic tendon implants or proactive

greg.holland@asu.edu, jyarger@gmail.com.

**Supporting Information Available**. $^{1}H \rightarrow ^{13}C$ CP and $^{13}C$ DD-MAS spectra with a rapid recycle delay collected of Ma and Mi silk in their wet and dry state. This material is available free of charge via the Internet at http://pubs.acs.org.

body armor, similarly to Kevlar®[4]. Recent advances in bioengineering are making these possibilities closer to reality. However, to date synthetic silk fibers do not exhibit the same physical or mechanical properties of natural spider silk [9-12]. The viable production and use of these proteins as synthetic materials requires greater knowledge of the secondary, tertiary, and quaternary structure of the proteins that comprise spider silk.

Ma and Mi silk of the *N. clavipes* spider both consist of two repetitive proteins, major ampullate spidroin 1 and 2 (MaSp1/MaSp2) [13, 14] and minor ampullate spidroin 1 and 2 (MiSp1/ MiSp2), respectively [15]. MaSp1 and MaSp2 contain poly-Alanine (poly-A) runs up to ten residues long and Glycine-Glycine-X (GGX) runs (where X = Glutamine (Q), Leucine (L), or Tyrosine(Y)), with MaSp2 also having Glycine-Proline-Glycine-X-X (GPGXX) motifs (where XX = G-Y or Q-Q)[13, 14]. MiSp1 and 2 are made up of (i) poly-A repeats that are shorter than Ma silk, (ii) GGX motifs (where X = A, Q, or Y), (iii) long poly(Glycine-Alanine) (poly-(GA)) regions similar in length to the poly-A repeats in Ma silk, and (iv) a Ser-rich nonrepetitive spacer region ("spacer")[15]. The poly-A regions in Ma and Mi silk have been shown to be in a β-sheet structure utilizing solid-state NMR [16, 17]. X-ray diffraction established that these crystalline β-sheets are aligned with the fiber axis [18]. Solid-state $^{2}$H NMR also points to highly ordered β-sheet, as well as disordered Ala rich regions in Ma silk [19].

The poly-(GA) and GGX repeats in Ma and Mi silk have not been definitively characterized, but are believed to be in a β-sheet and helical conformation, respectively [20]. The two-dimensional (2D) solid-state NMR techniques, double-quantum NMR spectroscopy (DOQSY) and direction exchange with correlation for orientation distribution evaluation and reconstruction NMR (DECODER), were used to investigate backbone torsion angles, as well as the orientation of the backbone with respect to the fiber axis in Ma silk [21]. This data confirmed alanine is mostly in a β-sheet structure and indicated that glycine is incorporated both in β-sheet and $3_1$-helical conformations [21]. 2D NMR $^{13}$C-$^{13}$C correlation experiments with proton-driven spin-diffusion (PDSD) under magic angle spinning (MAS) were used to investigate Ma silk from *Nephila edulis* and show that Ala was in two conformations [22]. The predominant structure was assigned to a β-sheet and the other was tentatively assigned to a $3_1$-helix. More recently, 2D $^{13}$C-$^{13}$C NMR correlation experiments with fast MAS and dipolar-assisted rotational resonance (DARR) on *Nephila clavipes* Ma silk provided evidence that Ala and Gly are both present in the ordered β-sheet and disordered $3_1$-helical conformations [23]. Rotational-echo double-resonance (REDOR) of Ma silk explored the Leu-Gly-X-Gln motif (where X = Ser, Gly, or Asn), illustrating that this region could form compact turn-like structures [24].

All these studies have provided valuable pieces to the molecular puzzle that is spider silk. However, no study has been able to quantify the fraction of a specific amino acid in a given secondary structure. This is particularly important as 95%+ of spider dragline silk is comprised of only six different amino acids [25], each of which can be present in multiple structural environments [22, 23]. It is thus, imperative to be able to determine the fraction of a specific amino acid in a specific conformation and correlate these fractions to different regions of the amino acid sequence. It was recently shown in Ma silk from the *N. clavipes* spider, that by utilizing the known plasticizing effect water has on silk to improve NMR resolution [26, 27], two distinct carbonyl environments for both Gly and Ala [28] were observed with a through-bond $^{13}$C-$^{13}$C double-quantum (DQ)/single-quantum (SQ) refocused incredible natural abundance double quantum transfer experiment (INADEQUATE). The carbonyl peaks for Gly and Ala were resolved enough to extract both linewidths and chemical shifts from these resonances. This information made it possible to fit the carbonyl region of a $^{13}$C MAS spectrum of plasticized silk and extract the percent of Gly and Ala that is incorporated in the β-sheet structure for Ma silk. In this study, we apply this methodology to quantify the percent of Gly, Ala, and previously uncharacterized Ser found in β-sheet and helical regions in Mi silk from

the same species and comparisons are made with Ma silks. To the authors' knowledge, this is the first structural characterization of the Ser rich spacer region that is unique to Mi silk. More significantly, this is the first time that very large portions of Ma and Mi silk, Gly, Ala, and Ser, have been quantitatively assigned to specific secondary structures and then correlated to the primary amino acid sequence. This information is important in understanding the differences in mechanical properties that natural spider silk possesses and will lead to improved production of silk as a synthetic biopolymer.

## Materials and Methods

### Spider Silk Isotope Labeling

Nephila clavipes spiders were forcibly silked at a rate of 2 cm/s to collect both Ma and Mi silk [29]. Silking was performed under a dissection microscope to ensure that the Ma and Mi silks were separated during the entire process. Spiders were silked every other day on average. Spiders were fed a cricket in addition to a $^{13}$C-enriched MEM5550 solution (Sigma-Aldrich, St. Louis, MO) during silking. Studies have shown that depriving *N. clavipes* of specific amino acids can change the mechanical properties of the silk.[30] Additionally, studies have shown changing the diet of a spider from a fly to a cricket can change the amino acid composition slightly.[31] However, no reports have been made that the addition of isotopic labeled amino acids to a spiders diet changes the silk in any way. The Ma sample used in this study is from the 11$^{th}$ and 12$^{th}$ silking. Comparing the intensity from a $^{13}$C CP-MAS experiment of the labeled silk to a natural abundance spectrum, a rough $^{13}$C enrichment is estimated at Ala $C_\beta$ (45%), Ala $C_\alpha$ (23%), Glu C $_{\beta,\gamma}$ (16%), Gly $C_\alpha$ (28%), Ser (27%), and the carbonyl (18%). Due to smaller sample size for each silking, the Mi sample includes silkings 5, 6, 8, 9, 10, 13, and 14. This sample is roughly enriched at Ala $C_\beta$ (26%), Ala $C_\alpha$ (17%), Glu C $_{\beta,\gamma}$ (6%), Gly $C_\alpha$ (16%), Ser (20%), and the carbonyl (10%). No appreciable enrichment of Pro, Leu, or Tyr were detected in either the Ma or Mi sample. All water-wetted samples were soaked in $D_2O$ for at least 30 minutes to ensure saturation prior to experiment. Wet samples were checked following experiments to ensure that the sample stayed hydrated throughout acquisition.

### Nuclear Magnetic Resonance (NMR) Measurments

The DARR experiments were collected on a Varian VNMRS 800 MHz equipped with a 1.6 mm triple resonance probe operating in double resonance ($^1$H/$^{13}$C) mode. DARR [32] experiments were performed at 40 kHz MAS. The $^1$H→$^{13}$C CP condition was matched to the -1 spinning sideband (ssb) in the Hartmann-Hahn (HH) profile. CP conditions consisted of a 2.2 $\mu$s $^1$H pulse, followed by a 1 ms ramped (17%) $^1$H spin-lock pulse of 115 kHz rf field strength. The sweep width in both dimensions was 50 kHz with 512 $t_1$ points collected in the second dimension. The recycle delay was 3 s and two-pulse phase-modulated (TPPM) [33] $^1$H decoupling with a rf field strength of 150 kHz was used during acquisition. During DARR mixing periods, continuous wave (CW) irradiation was applied on the $^1$H channel at a rotary-resonance condition ($\omega_r=\omega_{rf}$) of 40 kHz. Experiments with mixing times of 150 ms and 1 s were collected. Zero filling was applied in both dimensions to 2048 complex points. Both data sets were processed with 100 Hz exponential line-broadening in the direct dimension. A Gaussian apodization of the form $\exp(-(t/gf)^2)$, where t is time, was applied in the indirect dimension with the gaussian apodization constant, gf, equal to 0.006 s.

Refocused INADEQUATE [34, 35] NMR experiments were performed on a Varian VNMRS 400 MHz wide-bore spectrometer with a 3.2 mm triple resonance MAS probe operating in double resonance mode ($^1$H/$^{13}$C). Samples were packed in zirconia rotors. Data was collected at 20 kHz MAS and TPPM $^1$H decoupling with a rf field strength of ~100 kHz was used during acquisition. The $^1$H→$^{13}$C CP condition was matched to the -1 ssb of the HH profile. $^1$H→$^{13}$C CP parameters were a 3 $\mu$s $^1$H pulse with a 1 ms ramped (13%) $^1$H spin-lock

pulse of 85 kHz rf field strength. The experiment was collected with a 50 kHz sweep width in both dimensions, a recycle delay of 2 s, with 256 $t_1$ points, and a τ delay of 3 ms. For the water wetted samples, the refocused INADEQUATE was modified to utilize $^{13}C$ direct excitation with a 3 $\mu$s $^{13}C$ pulse [28]. The spectral width was 50 kHz in both dimensions with 288 $t_1$ points in the indirect dimension. The τ delay was set to 3.5 ms and two spectra with recycle delays of 1 s and 3 s were collected. TPPM $^{1}H$ decoupling 100 kHz was applied throughout the pulse sequence of all INADEQUATE experiments. For all data sets zero filling was applied in both dimensions to 2048 complex points. For all the INADEQUATE data, exponential line-broadening of 100 Hz was applied in the direct dimension. A Gaussian apodization was applied in the indirect dimension with the constant, gf, equal to 0.003 s.

$^{13}C$ direct spectra with dipolar decoupling (DD-MAS) of wet Ma and Mi silk were collected with a recycle delay of 10 s and 20 s, respectively, producing fully relaxed $^{13}C$ spectra. Deconvolution of the spectra was performed using Dmfit [36]. For the carbonyl fit, the chemical shift and line width were extracted from the 2D NMR INADEQUATE experiment. These were held constant and only the amplitude of the peak were varied during the fit. The total ratio of Ala to Gly in the silk is known from the amino acid composition of silk. This information was used to ensure that the ratio of Ala to Gly in the fit was correct. The chemical shifts of the aliphatic region were held constant and the linewidth and the amplitude of each peak were varied during the fit. For all fits, the shape of the fit was varied and a Lorentzian lineshape was found to be the best fit. For the percent Ser fit of Mi silk, (Figure 4) three peaks are assigned to Ser $C_\beta$. Although the Pro $C_\alpha$ chemical shift overlaps with Ser $C_\beta$, Pro is discounted because of the nearly nonexistent percentage of Pro in Mi silk. The 2D NMR DARR and INADEQUATE data further emphasizes this point as no Pro connections are observed.

## Amino Acid Sequence and “Counting”

The complete primary amino acid sequences for *Nephila Clavipes* major (MaSp1 and MaSp2) and minor (MiSp1 and MiSp2) ampullate silk proteins are unknown. However, sections of the sequences have been published. It has been shown in major silk from the Black Widow spider that these proteins maintain the repetitive sequence for the entire length of the repetitive region [37]. The protein sequences for *N. Clavipes* MaSp1 and MaSp2 are from the GenBank data base, accession no. AAA29380 and AAA29381, respectively. MiSp1 and MiSp2 are taken from Colgin and Lewis [15]. The protein sequences from this paper are found in the GenBank data base as accession no. AAC14589, AAC14590, and AAC14591. A summary of the amino acid “consensus” sequences can be found in figures 5 and 6. For Ma silk, all poly-A and poly-(GA) preceding and terminating poly-A are counted as β-sheet. Any Gly or Ser preceding or terminating poly-A or poly-(GA) in Mi silk are also considered incorporated in these β-sheet runs and counted as β-sheet. Rarely there is a Val found within the β-sheet region. In this case either side of the Val is counted as β-sheet. The β-sheet regions in Ma silk are generally terminated by either Pro or Gln. Poly-A and poly-(GA) in Mi silk are counted as β-sheet. Gly and Ala that are found in GGX in Mi silk are not counted as β-sheet as they are considered part of a motif which is believed to form a $3_1$-helix. In the case of Mi, there are long runs of poly-(GA) and poly-(A) that are broken by amino acids with small side chains. The Ala, Gly, and Ser on either side of these amino acids are counted as β-sheet. The “spacer” regions in Mi silk differ within the sequences, but retain greater than 90% identity. Therefore, the “spacer” region used for counting is considered representative of all “spacer” regions in the silk. For the “spacer” and “nonrep” regions, runs of four amino acids or more that include only Ala, Gly, or Ser were counted as β-sheet.

## Results And Discussion

### Two-dimensional $^{13}C$-$^{13}C$ Correlation NMR Experiments

Two-dimensional $^{13}C$-$^{13}C$ correlation experiments with DARR mixing periods of 150 ms and 1 s were collected on $^{13}C$-enriched *N. clavipes* spider Mi ampullate silk (Figure 1). Data collected with a mixing time of 150 ms primarily exhibits intramolecular correlations. Strong correlations are observed for Ala $C_\alpha$-$C_\beta$, Ala $C_\alpha$-CO, and Ala $C_\beta$-CO, as well as Gly $C_\alpha$-CO. These correlations confirm chemical shift assignments for Mi silk, are similar to Ma silk chemical shifts, and are consistent with previously reported Ma and Mi silk data (see Table 1) [23, 27]. With longer DARR mixing periods of 1 s, weak intermolecular correlations are observed between Ala $C_\alpha$ and Gly $C_\alpha$. This is not surprising as the primary amino acid sequences of both MiSp1 and 2 contain large runs of poly-(GA) and the GGX motif, where X can be Ala [15]. Spectral projections in the *F2* dimension are extracted from the 1 s DARR data to display the details for Ala and Gly (Figure 1c). There are two components of the Ala $C_\beta$. The projection taken at 17.4 ppm corresponds to Ala in a $3_1$ helix (Figure 1c, top) and exhibits broader linewidths for the three Ala peaks than the Ala peaks for the projection taken at 21.7 ppm (Figure 1c, middle). This is consistent with Ala at 17.4 ppm being in a more disordered environment, while Ala at 21.7 ppm agrees with an ordered β-sheet structure. From the projection taken at Gly $C_\alpha$, 43.4 ppm (Figure 1c, bottom), the Gly CO is an asymmetric peak that spans the chemical shift range for both β-sheet and helical conformations indicating that Gly is in two distinct environments. This data shows that Gly in Mi silk adopts an ordered β-sheet structure, as well as a helical structure, similarly to Ma silk [23]. However, unlike Ma silk [23], the upfield (lower ppm) side of the Gly CO in Mi is more pronounced indicating that more Gly in Mi is incorporated in a β-sheet structure than in Ma silk.

To complement the through-space DARR experiments, two versions of the through-bond two-dimensional $^{13}C$-$^{13}C$ refocused INADEQUATE NMR experiments were used to better characterize the Mi silk. The refocused INADEQUATE is a DQ/SQ experiment that shows correlations between directly bonded carbons [34, 35, 38]. A cross polarization (CP) version of the INADEQUATE NMR pulse sequence was used for dry Mi silk (Figure 2a). This experiment shows correlations between Ala $C_\alpha$-$C_\beta$, Ala $C_\alpha$-CO, Gly $C_\alpha$-CO, as well as Ser $C_\alpha$-$C_\beta$. In addition to confirming the chemical shift assignments for the Ala and Gly resonances that were established by the through-space DARR experiments, this experiment permits further investigation into the structure of the Ser rich “spacer” region by providing the assignments of Ser $C_\alpha$ and $C_\beta$.

The impact that water has on spider dragline silk can be exploited to increase resolution in the NMR spectra [27, 28]. Water plasticizes both Ma and Mi silk, although the silks behave differently in water. When wet, Ma silk from the *N. clavipes* contracts to ~65% the original length of the fiber, a phenomenon known as supercontraction[39, 40]. Mi silk, on the other hand, does not supercontract to any appreciable extent [41]. Several NMR techniques have been used to investigate the interesting effect that water has on silk [26, 27, 41]. $^{13}C$ dipolar decoupled magic angle spinning (DD-MAS) with a short recycle delay enhances regions that are more mobile with shorter $T_1$ relaxation times. $^{13}C$ DD-MAS of wet Ma and Mi silk exhibit similar enhancement at the Gly, Gln, Ser, and Ala resonances associated with non β-sheet regions indicating that water penetrates these regions and causes an increase in molecular mobility [26, 27, 42]. $^{13}C$ CP-MAS of both water wetted Ma and Mi silk, shows that the resonances from the crystalline poly-A β-sheets remain unaffected while other resonances decrease intensity, demonstrating that water does not penetrate the β-sheet structures [26, 27]. $^{1}H/^{13}C$ wide-line separation (WISE) NMR for both Ma and Mi silk illustrates that water penetrates the non β-sheet regions, leaving the poly-A β-sheet crystalline structure unchanged [27, 43]. Although the interactions between water and the silks are not completely understood, the plasticizing effect water has on both silks can be used to enhance resolution of $^{13}C$ NMR spectra.

When the silk is wet, the increase in mobility improves NMR resolution, but causes $^{1}H \rightarrow ^{13}C$ cross polarization to be inefficient and signal to be lost for those amino acids that interact with water. To take advantage of the plasticizing effect water has on silk, a directly polarized version of the refocused INADEQUATE experiment was used to assign the carbonyl resonances for both Ala and Gly. This experiment was recently used on Ma silk to assign carbonyl chemical shifts and extract linewidths [28]. The direct INADEQUATE experiment benefits from the fact that Ma and Mi silk become mobile when wet, providing improved resolution with narrower linewidths [27]. To probe these different resonances, a direct INADEQUATE was collected with a recycle delay of 1 s (Figure 2b), as well as 3 s (Figure 2c). From the direct INADEQUATE it is clear that both Ala and Gly have two distinct carbonyl sites, indicating two different environments. With a shorter recycle delay mobile regions are enhanced, while longer delays enhance rigid regions. The Ala CO and Gly CO that are downfield (higher ppm) are more intense in the 1 s INADEQUATE than the 3 s, indicating that these carbonyls for Ala and Gly adopt a mobile helical conformation. Conversely, the Ala and Gly CO resonances that are upfield (lower ppm) are less intense in the 1 s than 3 s INADEQUATE confirming that these are from the Ala and Gly found in a β-sheet structure. These assignments are consistent with carbonyl chemical shifts from polypeptides with known secondary structures (Table 1). In addition to resolving the chemical shifts for the carbonyls in different conformations, the linewidths from these peaks were also extracted from the $C_{\alpha}$-CO contact (previous data, not shown) [28]. In Ma silk, the full width at half-maximum (fwhm) of the carbonyl peaks for both Ala and Gly in a β-sheet structure are 220 Hz. The carbonyls for Ala and Gly in a helical conformation have narrower fwhm at 160 Hz and 150 Hz, respectively. Mi silk exhibits slightly broader carbonyls from the β-sheet Ala, 240 Hz, and Gly, 250 Hz, than Ma silk. Ala and Gly in Mi silk that are in a helical conformation exhibit carbonyl fwhm of 140 Hz and 160 Hz, respectively. These linewidths confirm the chemical shift assignments as narrower lines are expected for the mobile helical regions.

## Peak Fitting to Extract Percentage of Secondary Structure

After extracting the chemical shift and fwhm for each carbonyl peak, these two parameters were used to fit the carbonyl regions of wetted Ma and Mi silk from a fully relaxed $^{13}C$ direct spectrum (Figure 3). For the fits, the peak position and linewidth were fixed and only the amplitude of the peaks were varied. From the fits, the fraction of Ala and Gly in β-sheet and helical conformation was extracted. In Ma silk (Figure 3a), the Ala CO for the helical portion was found to be 18 ± 4% and the Ala CO of the β-sheet fraction was found to be 82 ± 4% (Table 2). As a self-consistent check, these percentages are compared to the Ala $C_{\beta}$ fit of the helical and β-sheet portions which are 20 ± 5% and 80 ± 5%, respectively (Figure 4a). Gly in Ma silk exhibits a smaller fraction in the β-sheet domain at 28 ± 5%, with 72 ± 5% in a helical structure. The Mi silk carbonyl (Figure 3b) shows a similar, but slightly higher amount of Ala in a helical structure at 23 ± 2% than Ma silk. This too agrees with the Ala $C_{\beta}$ fit for Mi silk, were the helical portion is 27 ± 5% (Figure 4b). Conversely, the fraction of Gly in a β-sheet found in Mi silk is nearly twice that compared to Ma silk at 53 ± 2%, with the helical Gly CO fraction amounting to 47 ± 2%. This is a significant result because it shows that although there is less Ala in a β-sheet in Mi silk than Ma silk, a large fraction of the silk does form a β-sheet structure that is Gly-rich.

In addition to the Ala and Gly carbonyls, the Ser $C_{\beta}$ chemical shift also is dependent on secondary structure (Table 1) [44]. Therefore, the Ser $C_{\beta}$ was fit to extract the fraction of Ser found in a β-sheet structure, as well as a helical conformation (Figure 4). Ser $C_{\beta}$ was assigned using the CP INADEQUATE (Figure 2a). Ser $C_{\beta}$ in a β-sheet is shifted downfield from Ser $C_{\beta}$ in a helical conformation [44]. The assignments of Ser $C_{\beta}$ in a β-sheet and helical region were confirmed by comparing $^{13}C$ CP-MAS and $^{13}C$ DD-MAS of both dry and water wetted Ma and Mi silk (See supplemental, Figure S1). In both cases, the downfield Ser $C_{\beta}$ did not lose

CP signal when wet indicating that it is the rigid β-sheet portion. The upfield Ser $C_\beta$ in both cases is enhanced during $^{13}C$ DD-MAS with a short recycle delay of 1 s, indicating that this resonance can be attributed to a mobile helical conformation. The Ser $C_\beta$ resonances from completely relaxed water wetted spectra of Ma and Mi silk were fit and the fraction of Ser found in a β-sheet and helical structure was extracted. The Ser in Ma silk was found to be 21 ± 2% in a β-sheet secondary structure and 79 ± 2% helical conformation (Figure 4a). The fraction of Ser in a β-sheet in Mi silk is over twice that of Ma silk at 52 ± 1%, with 48 ± 1% in a helical conformation (Table 2). The Ser in Mi exhibits two helical resonances, which have been combined to account for the total helical fraction. Further work is needed to determine the exact structure of these helical regions.

## Correlating Primary Amino Acid Sequence to Percentages Extracted from the NMR Data

The amino acid sequence for both MaSp1 and MaSp2 are used to predict the percent of Ala, Gly, and Ser expected in a β-sheet structure (Figure 5). Although the complete sequences for *N. clavipes* MaSp1 and MaSp2 are not know, recent data has shown that the sequences for Ma silk are highly repetitive [37]. Therefore, the consensus sequences are believed to be an accurate representation of the proteins as a whole. If it is assumed that all the Ala, Gly, and Ser in MaSp1 and MaSp2 found in poly-A or poly-(GA) terminating poly-A runs form a β-sheet, as well as poly-A runs that contain Ser, then the percent of Ala, Gly, and Ser that form a β-sheet structure can be calculated. Taking into account that the ratio of MaSp1 to MaSp2 in Ma silk is 81:19% [45], the amount of Ala, Gly, and Ser in a β-sheet structure predicted by the primary amino acid sequence is found to be 86%, 26%, and 19% respectively (Table 2). Comparing these percentages to the fraction of β-sheet extracted from the fits of the Ma silk, excellent agreement is observed. From the NMR data, the percentage of Ala, Gly, and Ser in a β-sheet structure is 82 ± 4%, 28 ± 5%, and 21 ± 2%, respectively (Figures 3a).

Like Ma silk, MiSp1 and MiSp2 have not been completely sequenced. Although MiSp1 and MiSp2 are only partial sequences, with the MiSp2 sequence being less complete than many other silk sequences, the consensus sequences are still believed to be representative of the proteins as a whole. It should be noted that the use of partial sequences does introduce the possibility of error into the correlation between the primary amino acid sequence and NMR data. The percentage of Ala, Gly, and Ser in Mi silk expected to be in a β-sheet are counted in a similar manner as the Ma silk (Figure 6). However, the ratio of MiSp1 to MiSp2 in Mi silk is unknown. Therefore, when predicting the amount of β-sheet found in the silk from the consensus sequences only a range can be provided. This range spans the possibility of Mi silk being composed only of MiSp1 to solely MiSp2. For MiSp1, Ala expected to be in β-sheet is 83% and for MiSp2 it is 75%. Therefore, although the ratio of MiSp1 to MiSp2 is unknown, the fraction of Ala in a β-sheet in Mi silk is 75-83%. Gly is predicted to be 52% β-sheet structure in MiSp1 and 44% β-sheet in MiSp2. Ser in MiSp1 is counted to be 41% β-sheet and 48% β-sheet in MiSp2. These percentages include Ser found in the “spacer” and carboxy-terminal nonrepetitive (“nonrep”) regions. If these regions do not contain β-sheet, then 7% and 33% of MiSp1 and MiSp2 would be predicted to form β-sheet, respectively.

Comparing the percentage of Ala, Gly, and Ser predicted to be in a β-sheet from the primary amino acid sequences from Mi silk with the fits from the NMR experiments exhibit excellent agreement. For Ala, the NMR data shows that 77±2% is in a β-sheet conformation and the primary amino acid sequence predicts 75-83%. Gly exhibits good agreement with the NMR data indicating 53 ± 2% β-sheet and the primary amino acid sequences predicting 44-52% β-sheet. The NMR data for the fraction of Ser in β-sheet is 52 ± 1%. From this data, it is concluded that the “spacer” and “nonrep” regions must contain β-sheet given that 52% is closer to the 41-48% predicted from the primary amino acid sequence, as opposed to the 7-33% predicted

when these regions are considered amorphous or helical. This is the first time that any characterization has been made for the “spacer” and “nonrep” regions of Mi silk.

## Conclusion

From the data presented, 34% of the Ma silk is present in a β-sheet conformation. This number is much higher than the 10-15% crystallinity reported by X-ray diffraction (XRD).[46] One possible explanation for this discrepancy is that only the poly-A runs in the silk are ordered enough to diffract and subsequently are the only portions of the silk that are observed by XRD. Poly-A makes up ~16% of the entire Ma silk, agreeing very nicely with the 10-15% crystallinity by X-ray. Unfortunately, X-ray diffraction studies of Mi silk are rare and no percent crystallinity has been reported [18]. From the sequences of MiSp1 and 2, 7-10% of the silk is poly-A. The total percent of β-sheet in Mi silk made up of Ala, Gly, and Ser is ~45%. Unlike Ma silk, Mi has a very large percentage of poly-(GA) repeats. In this way, Mi silk is very similar to silkworm silk, which is made up of 55% AGSGAG repeats that also form β-sheet crystalline structure [47]. Ma silk has a smaller total percent β-sheet than Mi silk, but is stronger and has a lower extensibility. Further studies on the non-β-sheet regions of the silk are needed to better correlate the mechanical properties to specific regions of the silk.

Ala, Gly, and Ser from the primary amino acid sequences of both proteins in Ma and Mi silk from the *Nephila clavipes* spider have been quantitatively correlated to the secondary structures present in the silks utilizing solid-state NMR. Taking advantage of the plasticizing effect that water has on silk to increase spectral resolution from a direct INADEQUATE experiment, distinct carbonyl resonances were observed for Ala and Gly in both β-sheet and helical regions. The chemical shifts and linewidths were extracted and used to fit fully relaxed and quantitative 1D spectra of Ma and Mi silk to extract the percentage of Ala and Gly found in β-sheet structure. The Ser $C_\beta$ was also fit to extract β-sheet and helical fractions. This data has been correlated to the primary amino acid sequences and exhibits good agreement with what has previously been speculated [20, 48]. In addition to being the first quantitative correlation between specific amino acids present in the primary amino acid sequence and the secondary structure that they adopt, this is the first data that characterizes portions of the “spacer” and “nonrep” regions of Mi silk. This work has successfully characterized 34% of the total secondary structure in Ma silk and 45% of the secondary structure of Mi silk from the *N. clavipes* spider as β-sheet.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by the National Science Foundation (CHE-0612553 and DMR-0805197) and the NIH (NIBIB-5R01EB000490-05). We thank Dr. Brian Cherry for help with NMR instrumentation, student training and scientific discussion.

## References

1. Shear WA, Palmer JM, Coddington JA, Bonamo PM. Science 1989;246:479–481. [PubMed: 17788699]
2. Lewis R. Chem. Rev 2006;106:3762–3774. [PubMed: 16967919]
3. Vollrath F, Porter D. Soft Matter 2006;2:377–385.
4. Gosline J, DeMont E, Denny M. Endeavour 1986;10:37–43.
5. Swanson BO, Blackledge TA, Beltran J, Hayashi CY. Appl. Phys. A 2006;82:213–218.
6. Blackledge T, Hayashi C. J. Exp. Biol 2006;209:2452–2461. [PubMed: 16788028]

7. Work R. Tex. Res. J 1977;47(10):650–662.
8. Papadopoulos P, Ene R, Weidner I, Kremer F. Macromol. Rapid Comm 2009;30:851–857.
9. Lazaris A, Arcidiacono S, Huang Y, Zhou JF, Duguay F, Chretien N, Welsh EA, Soares JW, Karatzas CN. Science 2002;295:472–476. [PubMed: 11799236]
10. Brooks AE, Stricker SM, Joshi SB, Kamerzell TJ, Middaugh CR, Lewis RV. Biomacromolecules 2008;9:1506–1510. [PubMed: 18457450]
11. Seidel A, Liivak O, Jelinksi L. Macromolecules 1998;31:6733–6736.
12. O'Brian JP, Fahnestock SR, Termonia Y, Gardner KH. Adv. Mater 1998;10(15):1185–1195.
13. Xu M, Lewis R. Proc. Natl. Acad. Sci. USA 1990;87:7120–7124. [PubMed: 2402494]
14. Hinman M, Lewis R. J. Biol. Chem 1992;27:19320–19324. [PubMed: 1527052]
15. Colgin M, Lewis R. Protein Sci 1998;7:667–672. [PubMed: 9541398]
16. Simmons A, Ray E, Jelinski L. Macromolecules 1994;27:5235–5237.
17. Liivak O, Flores A, Lewis R, Jelinski L. Macromolecules 1997;(30):7127–7130.
18. Parkhe A, Seeley S, Gardner K, Thompon L, Lewis R. J. Mol. Rec 1997;10:1–6.
19. Simmons A, Michal C, Jelinski L. Science 1996;271:84–87. [PubMed: 8539605]
20. Hayashi C, Shipley N, Lewis R. Int. J. Biol. Macromol 1999;24:271–275. [PubMed: 10342774]
21. van Beek JD, Hess S, Vollrath F, Meier BH. Proc. Natl. Acad. Sci. USA 2002;99:10266–10271. [PubMed: 12149440]
22. Marcotte I, van Beek B, Meier B. Macromolecules 2007;40:1995–2001.
23. Holland GP, Creager MS, Jenkins JE, Lewis RV, Yarger JL. J. Am. Chem. Soc 2008;130:9871–9877. [PubMed: 18593157]
24. Michal C, Jelinski L. J. Biomol. NMR 1998;12:231–241. [PubMed: 9751996]
25. Work RW, Young CT. J. Arachonol 1987;15(1):65–80.
26. Yang Z, Liivag O, Seidel A, LaVerde G, Zax D, Jelinski L. J. Am. Chem. Soc 2000;122:9019–9025.
27. Holland GP, Jenkins JE, Creager MS, Lewis RV, Yarger JL. Biomacromolecules 2008;9:651–657. [PubMed: 18171016]
28. Holland GP, Jenkins JE, Creager MS, Lewis RV, Yarger JL. Chem. Comm 2008;(43):5568–5570. [PubMed: 18997954]
29. Work RW, Emerson PD. J. Arachnol 1982;10:1–10.
30. Zax DB, Armanios DE, Horak S, Malowniak C, Yang Z. Biomacromolecules 2004;5:732–738. [PubMed: 15132654]
31. Craig CL, Riekel C, Herberstein ME, Weber RS, Kaplan D, Pierce NE. Mol. Biol. Evol 2000;17(12): 1904–1913. [PubMed: 11110907]
32. Takegoshi K, Nakamura S, Terao T. J. Chem. Phys 2003;118(5):2325–2341.
33. Bennett A, Rienstra CM, Auger M, Lakshmi KV, Griffin RG. J. Chem. Phys 1995;103(16):6951–6958.
34. Lesage A, Auger C, Caldarelli S, Emsley L. J. Am. Chem. Soc 1997;119:7867–7868.
35. Cadars S, Sein J, Duma L, Lesage A, Pham TN, Baltisberger JH, Brown SP, Emsley L. J. Magn. Reson 2007;188:24–34. [PubMed: 17588789]
36. Massiot D, Fayon F, Capron M, King I, Le Calvé S, Alonso B, Durand JO, Bujoli B, Gan Z, Hoatson G. Magn. Reson. Chem 2002;40(1):70–76.
37. Ayoub NA, Garb JE, Tinghitella RM, Collin MA, Hayashi CY. Plos One 2007;(6):1–13.
38. Lesage A, Bardet M, Emsley L. J. Am. Chem. Soc 1999;121:10987–10993.
39. Savage K, Gosline J. J. Exp. Biol 2008;211:1937–1947. [PubMed: 18515724]
40. Work R. J. Arachnol 1981;9:299–308.
41. Jelinski LW, Blye A, Liivak O, Michal C, La Verde G, Seidel A, Shah N, Yang Z. Int. J. Biol. Macromol 1999;24(23):197–201. [PubMed: 10342765]
42. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. J. Biomol. NMR 1995;5:67–81. [PubMed: 7881273]
43. Holland GP, Lewis RV, Yarger JL. J. Am. Chem. Soc 2004;126:5867–5872. [PubMed: 15125679]

44. Hronska M, van Beek JD, Williamson PTF, Vollrath F, Meier BH. Biomacromolecules 2004;5:834. [PubMed: 15132669]
45. Brooks AE, Steinkraus HB, Nelson SR, Lewis RV. Biomacromolecules 2005;6(6):3095–3099. [PubMed: 16283732]
46. Grubb DT, Jelinski LW. Macromolecules 1997;30:2860–2867.
47. Yao J, Nakazawa Y, Asakura T. Biomacromolecules 2004;5:680–688. [PubMed: 15132647]
48. Jelisnki LW. Curr. Opin. Solid St. M 1998;3(3):237–245.
49. Saito H. Magn. Reson. Chem 1986;24(10):835–852.
50. Saito H, Tabeta R, Asakura T, Iwanaga Y, Shoji A, Ozaki T, Ando I. Macromolecules 1984;17(7): 1405–1412.
51. Shoji A, Ozaki T, Saito H, Tabeta R, Ando I. Macromolecules 1984;17(8):1472–1479.
52. Asakura T, Yao J. Protein Sci 2002;11(11):2706–2713. [PubMed: 12381852]
53. Asakura T, Yang MY, Kawase T, Nakazawa Y. Macromolecules 2005;38(8):3356–3363.
54. Ashida J, Ohgo K, Komatsu K, Kubota A, Asakura T. J. Biomol. NMR 2003;25(2):91–103. [PubMed: 12652118]
55. Ishida M, Asakura T, Yokio M, Saito H. Macromolecules 1990;23(1):88–94.
56. Kricheldorf HR, Muller E. Macromolecules 1983;16(4):615–623.
57. Murata K, Kuroki S, Ando I. Polymer 2002;43(25):6871–6878.
58. Wildman KAH, Wilson EE, Lee DK, Ramamoorthy A. Solid State Nucl. Mag 2003;24(23):94–109.
59. Kishi S, Santos A, Ishii O, Ishikawa K, Kunieda S, Kimura H, Shoji A. J. Mol. Biol 2003;649(12): 155–167.
60. Asakura T, Yao JM, Yamane T, Umemura K, Ultrich AS. J. Am. Chem. Soc 2002;124(30):8794–8795. [PubMed: 12137522]
61. Asakura T, Nakazawa Y, Ohnishi E, Moro F. Protein Sci 2005;14(10):2654–2657. [PubMed: 16195552]
62. Zhao CH, Asakura T. Prog. Nucl. Mag. Res. Sp 2001;39(4):301–352.
63. Asakura T, Ashida J, Yamane T, Kameda T, Nakazawa Y, Ohgo K, Komatsu K. J. Mol. Biol 2001;306 (2):291–305. [PubMed: 11237601]
64. Asakura T, Demura M, Date T, Miyashita N, Ogawa K, Williamson MP. Biopolymers 1997;41(2): 193–203.

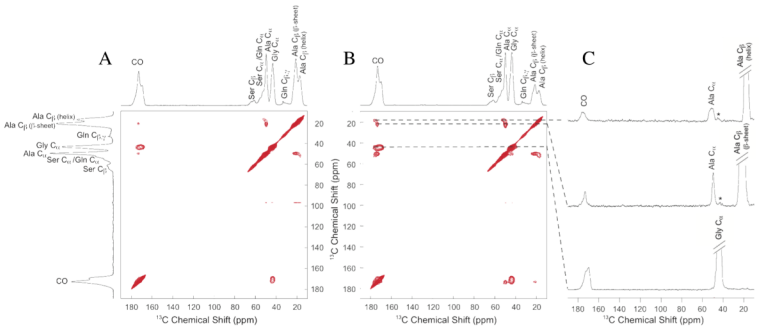


**Figure 1.**

$^{13}$C-$^{13}$C correlation NMR experiments of *N. clavipes* Mi silk with DARR mixing times of (A) 150 ms and (B) 1 s. Projections are taken from the 1 s DARR experiment at the (C, top) $3_1$-helical Ala $C_\beta$ (17.4 ppm), (C, middle) β-sheet Ala $C_\beta$ (21.7 ppm), and (C, bottom) Gly $C_\alpha$ (43.4 ppm). * Indicates weak intermolecular correlations between Gly and Ala.

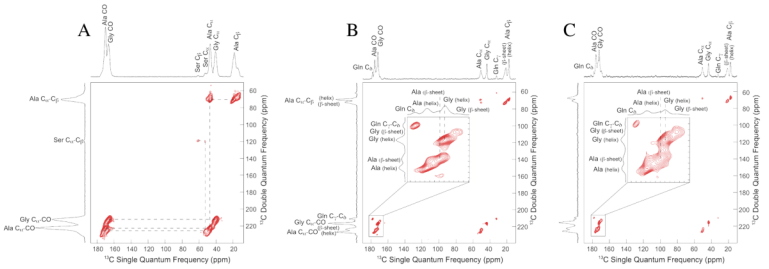


**Figure 2.**

$^{13}$C-$^{13}$C through-bond double quantum/single quantum (DQ/SQ) correlation collected with refocused INADEQUATE NMR experiments of *N. clavipes* Mi silk: (A) An initial cross polarization step was used to enhance carbon magnetization for the dry silk, while a direct carbon INADEQUATE was used for the water wetted silk (B & C). A recycle delay of (B) 1 s enhances mobile components that contain shorter $T_1$ relaxation times, while a (C) 3 s recycle delay emphasizes rigid components. Carbonyl regions are blown up in insets to reveal two distinct CO resonances for Ala and Gly.

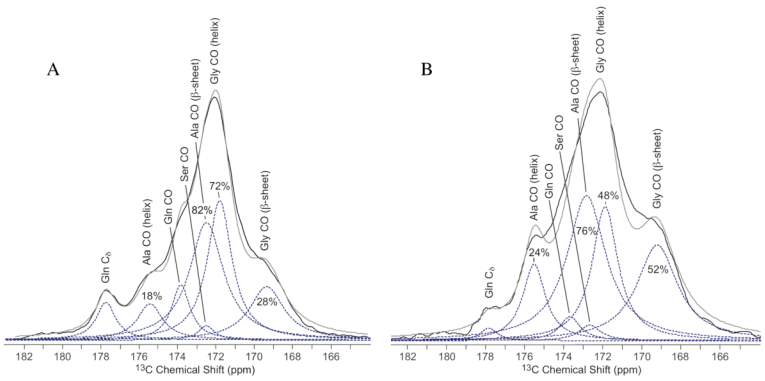


**Figure 3.**

Carbonyl region from fully relaxed $^{13}C$ DD-MAS NMR spectrum of water wetted *N. clavipes* (A) Ma silk and (B) Mi silk. The spectra were fit to extract the percent of Gly and Ala that each adopt both β-sheet and helical conformations (see Table 2). The original data is in black, the sum of the fits are in grey, and the individual fits are dotted lines.

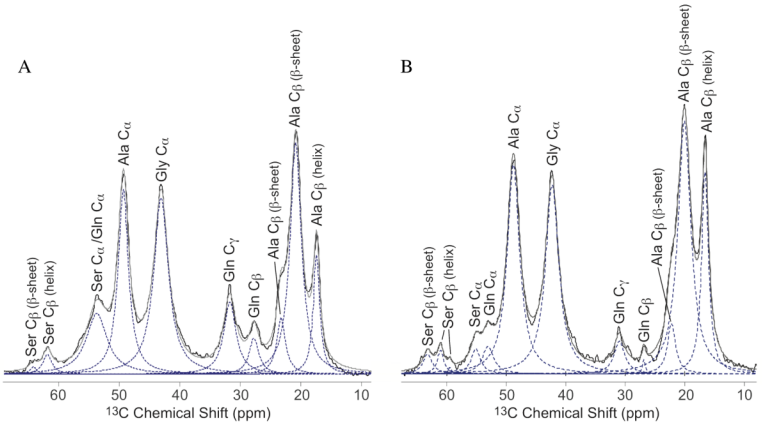


**Figure 4.**

Aliphatic region from fully relaxed $^{13}C$ DD-MAS NMR spectrum of water wetted *N. clavipes* (A) Ma and (B) Mi silk. The Ser $C_\beta$ fit was used to determine the percentage of Ser incorporated into a β-sheet or helical conformation. The Ala $C_\beta$ in both the Ma and Mi silk were fit and compared to carbonyl fits. This provided a self-consistent check for the percentage of Ala in a β-sheet or helical conformation extracted from the carbonyl fits. The original data is in black, the sum of the fits are in grey, and the individual fits are dotted lines.

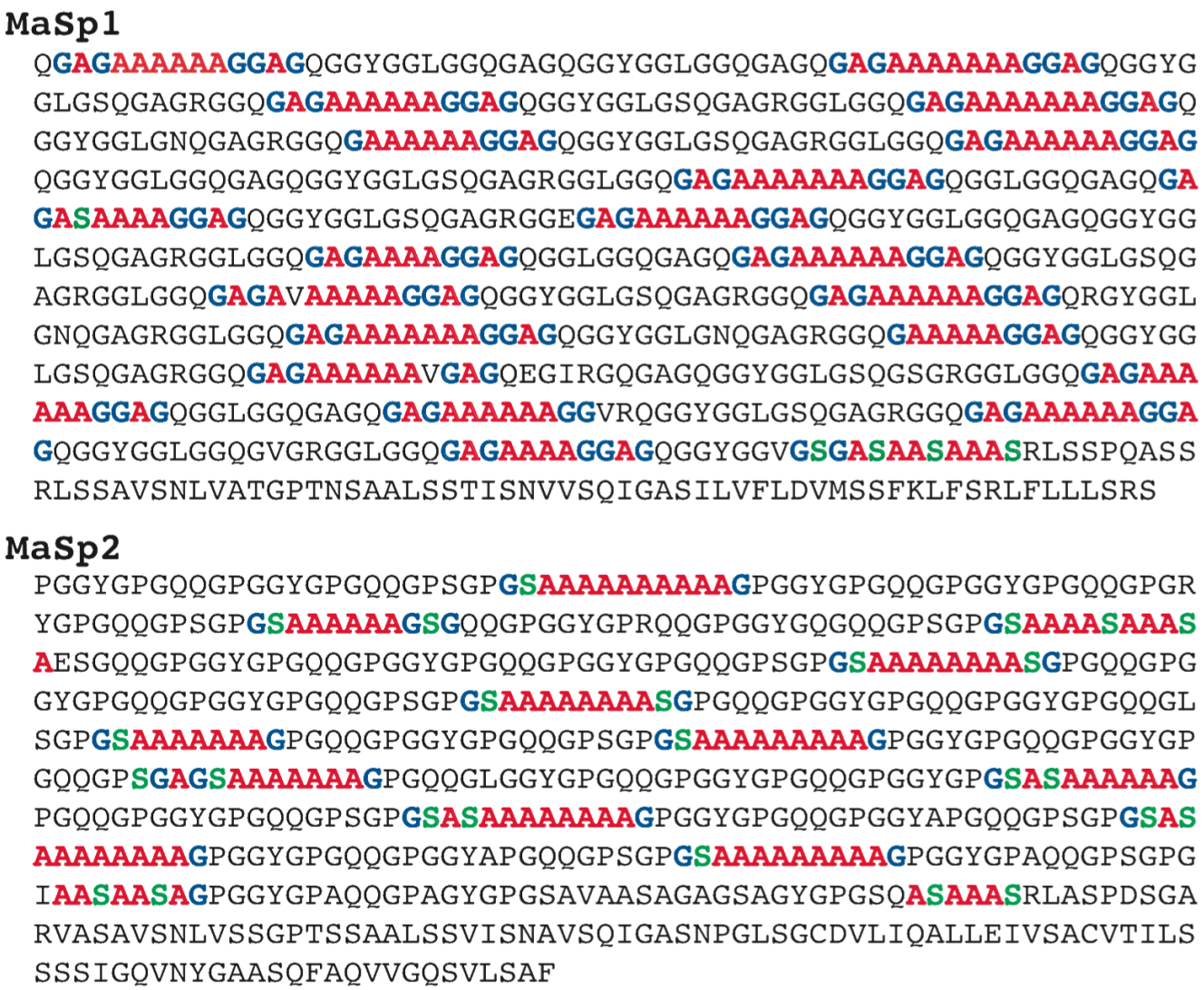


**Figure 5.**

Primary amino acid sequences for *N. clavipes* MaSp1 and MaSp2, the two proteins that make up Ma silk. Colored Ala (red), Gly (blue), and Ser (green) represent β-sheet structure. The fraction of Ala, Gly, and Ser in a β-sheet structure in silk predicted by the primary amino acid sequences are 86%, 26%, and 19%, respectively. Percentage of β-sheet determined experimentally by NMR agrees with the above structural model.

**MiSp1**

YGRGAAAGAGAGAGGSGYGGQGGYGAGAGAGAAAAAGSGAGAGGYGGQAGYGAGAGAG-SPACER-GAGAAA
AAGGAGGYGRGAGAGAGATAGAGAGGYGGQGGYGAGAGAAAASGAGAGGAGGYGRGAGAGAGAAAAGAGGYG
RGAAAGAGAAAGAGAGAGGGGYGQGGYGAGAGAGAAAAAGAGAAAAAAAAAGYGRGGAGAGTGAGAGGYGGQ
GGYGAGAGAG-SPACER-GAGAAAAARGAASGAGAAAGAGAGAGGAGYGGQIGYGAGAGAGAAAAAGAGAGA
GGAAGYGRGAGAGSGAAAGAGSGAGAGGYGGQAGYGAGA-SPACER-GAGAGAVAAAGAAGAGGYGRGAGGY
GGQGGYGAGAGAGAAAAAGAGAGGAGGYGRGAGAGAGAAAGAGAGAGGAGYGGQGGYGAGAGAGAAAAAGAG
AGGAGGYGRGAGAGAGAAAGAGAGGYGGQGGYGAGAGAGAAAAAAGAGSGGAGGYGRGAGAGAGAAAGAGAG
AGSYGGQGGYGAGAGAGAAAAAGAGAGAGGYGRGAGAGAGAGAGAAARAGAGAGGAGYGGQGGYGAGAGAGA
AAAAGAGAGGAGGYGRGAGAGAGAAAGAGAGAGGYGGQSGYGAGAGAAAAAGAGAGGAGGYGRGAGAGAGAA
AGAGAGAAAGAGAGGYGGQGGYGAGAGAGAAAAAGAGAGGAGYGRGAGAGAGVAAGAGAGGYGGQGGYGAGA
GAGAAAAAATGAGGAGGYGRGAGAGAGAAAGAGAGTGGAGYGGQGGYGAGAGAGAAAAAGAGAGGAGYGRGA
GAGAGAAAGAGAGAAAGAGAGAGGYGGQGGYGAGAGAGAAAAAGAGAGGAAGYSRGGRAGAAGAGAGAAAGA
GAGAGGYGGQGGYGAGAGAGAAAAAGAGSGGAGGYGRGAGGAAAGAGAAAGAGAGAGGYGGQGGYGAGAGAA
AAAGAGAGRGGYGRGAGAGGYGGQGGYGAGAGAGAAAAAGAGAGGYGDKEIACWSRCRYTVASTT-NONREP

"MiSp1 NONREP"
SRLSSAEASSRISSAASTLVSSGGYLNTAALPSV
ISDLFAQVGASSPGVSDSDEVLIQVLLEIVSSLI
HILSSSSVGQVDFSSVGSSAAAVGQSMQVVMG

**MiSp2**

GGYDGQGGYGAGAGAAAAAGAGAGASVGGYGRGAVAGSGSAAGAGAGARAGSGGYGGQGGYGAGTGAAAAAG
AGAGAGSAAGYGRGAGGGGEGFGGQGSYGAGAGA-SPACER-GGAGGYERGAGAGPGAAAGSGAAVGAGAGA
GSAGGYGGQRGYGAVTGAGAAAAAGAAGSAGGYGRGAGAGTGAGGYGRGAGAGAGTSAGGYGGQGGYGAGAG
AGAAAGAGDGSAGGYGRGAGAGAAAGAGAGGSTGGFGGQGGYGAGAGAAAAGAFAGRAGGYGRAAGAAAGTG
AAAGAGAGSTGGFGGQRGYGAGRSNGTGAGAAAAAGAGAGSAGGYGRGAGVAAGAGAGGASAGGYGGQGGYG
AVTKAGAAAGGYGRGAGAGPGAAAGSGAAVGAGAAAGSAGGYGGQ-NONREP

"MiSp2 NONREP"
SRLSSAEACSRISAAASTLVSGGSSLNTAALPSV
ISDLFAQVGASSPGVSDSDEVLIQVLLEIVSSLI
HILSSSSVGQVDFSSVGSSAAAVGQSMQVVMG

**SPACER**

GAGSSAGNAFAQSLSSNLLSSGDFVQMISSTTSTDHAVSVATSVAQN
VGSQLGLDANAMNNLLGAVSGYVSTLGNAISDASAYANALSSAIGN
VLANSGSISESTASSAASSAASSVTTTLTSYGPAVFYAPSASSGG

**Figure 6.**

Primary amino acid sequences for *N. clavipes* MiSp1, MiSp2, the Ser-rich nonrepetitive spacer region ("spacer"), and carboxy-terminal nonrepetitive ("nonrep") regions that make up Mi silk. Colored Ala (red), Gly (blue), and Ser (green) represent β-sheet structure. The predicted percentage of Ala, Gly, and Ser that form β-sheet in Mi silk (75-83%, 44-52%, and 41-48%, respectively) agrees with experimental NMR data. The ratio of MiSp1 to MiSp2 is unknown, therefore a range in percentage predicted by the AA sequence is based on 100% MiSp1 or MiSp2.

## Table 1

$^{13}$C chemical shifts of *N. clavipes* major and minor ampullate spider silk along with chemical shifts from polypeptides with known secondary structures

| | $^{13}$C chemical shift (in ppm from TMS) | | | | | | |
|---|---|---|---|---|---|---|---|
| **residue** | **major silk** | **minor silk** | **α-helix** | **β-sheet** | **random coil** | **$3_1$-helix** | **β-turn** |
| Ala $C_\beta$ | 17.5 | 17.4 | 14.8-16.0 | 19.9-20.7 | 19.1 | 17.4 | 16.5-17.4 |
| Ala $C_\beta$ | 20.9 | 20.8 | | | | | |
| Ala $C_\beta$ | 23.3 | 23.1 | | | | | |
| Ala $C_\alpha$ | 49.0 | 49.0 | 52.3-52.8 | 48.2-49.3 | 52.5 | 48.9 | 50.8-51.7 |
| Ala $C_\alpha$ | 49.2 | | | | | | |
| Ala $C_\alpha$ | 50.0 | | | | | | |
| Ala CO | 172.5* | 172.8* | 176.2-176.8 | 171.6-172.4 | 177.8 | 174.6 | 176.8-177.5 |
| Ala CO | 175.4* | 175.5* | | | | | |
| Gln $C_\beta$ | 32 | 28.0 | 25.6-26.3 | 29.0-29.9 | 29.4 | | |
| Gln $C_\gamma$ | 33.2 | 31.9 | 29.7-29.8 | 29.7-29.9 | 33.7 | | |
| Gln $C_\alpha$ | 52.9 | 53.8 | 56.4-57.0 | 51.0-51.4 | 56.2 | | |
| Gln CO | 172.1 | | 175.4-175.9 | 171.9-172.2 | 176.0 | | |
| Gln $C_\delta$ | 176.5 | 177.8* | | | 180.5 | | |
| Gly $C_\alpha$ | 43.3 | 43.1 | | 43.2-44.3 | 45.1 | 41.4-42.5 | 43.8-44.1 |
| Gly CO | 169.3* | 169.2* | | 168.4-169.7 | 174.9 | 171.2-173.1 | 170.6-170.7 |
| Gly CO | 171.8* | 171.9* | | | | | |
| Ser $C_\alpha$ | 55.5* | 55.9* | 59.2 | 54.5-55.0 | 58.3 | | 58.0 |
| Ser $C_\beta$ | 62.0 * | 60.3* | 60.7 | 62.3-63.9 | 63.8 | | 60.7 |
| Ser $C_\beta$ | 64.3* | 61.9* | | | | | |

**$^{13}$C chemical shift (in ppm from TMS)**

| residue | major silk | minor silk | α-helix | β-sheet | random coil | $3_1$-helix | β-turn |
|---|---|---|---|---|---|---|---|
| Ser $C_\beta$ | | 64.1* | | | | | |
| Ser CO | 172.2 | | 170.0-171.2 | 174.6 | 173.7 | | |

*Nephila clavipes* Ma and Mi chemical shifts are from previous work [23, 27], except where noted.

*Chemical shifts extracted from INADEQUATE data of wetted silk in this study. Chemical shifts of known secondary structures from model pepides [49-59], random coil conformation [42], and Silk I *β*-turn structure [60-64]. The $^{13}$C chemical shifts are in parts per million (ppm) with tetramethylsilane (TMS) as the 0 ppm reference. The label 'CO' indicates the peptide backbone carbonyl for a specific amino acid.

**Table 2**

Percent β-sheet and helical structure in *N. clavipes* major (Ma) and minor (Mi) ampullate spider silk predicted from the consensus primary amino acid sequence (AA %) and extracted from the curve fits to NMR data shown in figure 3 and 4 (NMR %)

| | **Major Silk** | | | | **Minor Silk** | | | |
|---|---|---|---|---|---|---|---|---|
| | **β-sheet** | | **helical** | | **β-sheet** | | **helical** | |
| | **AA (%)** | **NMR (%)** | **AA (%)** | **NMR (%)** | **AA (%)** | **NMR (%)** | **AA (%)** | **NMR (%)** |
| Ala | 86 | 82 ± 4 | 14 | 18 ± 4 | 75[‡] - 83[†] | 77 ± 2 | 17[†] - 25[‡] | 23 ± 2 |
| Gly | 26 | 28 ± 5 | 74 | 72 ± 5 | 44[‡]- 52[†] | 53 ± 2 | 48[†]- 56[‡] | 47 ± 2 |
| Ser | 19 | 21 ± 2 | 81 | 79 ± 2 | 41[†]- 48[‡] | 52 ± 1 | 52[‡]- 59[†] | 48 ± 1 |

The amount of MiSp1 and MiSp2 is not known for *Nephila clavipes* minor ampullate silk (Mi). The percent amino acid in pure MiSp1[†] and MiSp2[‡] represent the range of possible concentrations for minor ampullate silk.