Conformation of the Polyalanine Repeats in Minor Ampullate Gland Silk of the Spider *Nephila clavipes*

Oskar Liivak,[†] Anthony Flores,[‡] Randolph Lewis,[‡] and Lynn W. Jelinski*[,†]

Center for Advanced Technology in Biotechnology, Cornell University, Ithaca, New York 14853, and Department of Molecular Biology, University of Wyoming, Box 3944, Laramie, Wyoming 82071-3944

Received June 10, 1997; Revised Manuscript Received September 15, 1997[⊗]

ABSTRACT: Solid state ^{13}C NMR is used to identify the conformation of alanine residues in minor ampullate gland silk from *Nephila clavipes* and in a genetically engineered protein based on the consensus sequence of MaSp2, a protein present in low concentrations in major ampullate gland silk. The results of the NMR on minor ampullate gland silk are compared to previous NMR data on major ampullate gland silk, and the results on the genetically engineered protein are compared to previous NMR data on the gland fibroin from the major ampullate gland. Differences in the secondary structure of major and minor ampullate gland silk are correlated with mechanical properties. The alanine residues of major ampullate gland silk have previously been shown to consist almost entirely of β-sheet conformations. In contrast, the conformations of the alanine residues of minor ampullate gland silk are more heterogeneous, with a larger fraction of the alanine residues in non-β-sheet conformations. It is hypothesized that these non-β-sheet structures in minor ampullate gland silk are part of the cause for its lower tensile strength. In addition, both unprocessed genetically engineered silk and lyophilized major ampullate gland fibroin, have similar alanine C_β carbon chemical shifts yet differ in their alanine C_α carbon chemical shifts. This suggests that this genetically engineered protein is good starting material for biofiber synthesis but that it does not exactly mimic the gland fibroin.

Introduction

The structure, conformation, and mechanical properties of spider silk, particularly the major ampullate dragline silk from the spider *Nephila clavipes*, have been the subjects of intense scrutiny. Spider silk is a biomaterial that has undergone eons of evolutionary optimization, which has resulted in differing properties for each type of silk.[1,2] These silks have unusually good mechanical properties with potential human applications such as seat belts, surgical materials, and parachutes. It is therefore appropriate and timely to understand the molecular basis for its mechanical properties.

Single fibers of spider dragline silk have a tensile strength that rivals that of steel[3] and Kevlar, yet the fibers can stretch to more than 10% elongation before breaking.[4] Dragline silk combines a reasonably high stiffness with a very large extension to break, so that the toughness—the energy required to cause a tensile failure—is among the highest of any known fiber. The fiber does not fail by kinking, a feature that makes spider silk in some ways superior to the highest performance human-made fibers.[5] Furthermore, it is a protein polymer, produced from natural resources. These features make it attractive to investigate the production of high-performance fibers based on proteins. Such investigations are particularly timely, as the tools of biotechnology now make it possible to produce large quantities of silklike proteins.[6,7]

^{13}C NMR is capable of probing the structure of proteins. A number of workers have investigated the relationship between the ^{13}C chemical shifts of amino acids and their local secondary structure[8−10] and have shown that ^{13}C chemical shifts are useful for determining the local secondary structure of proteins.[11] Using NMR, it is possible to deduce the local secondary structure of the alanine residues. In particular, it has been shown that the chemical shift of the alanine C_α peak shifts upfield when the residue is in a β-sheet conformation and shifts downfield when the conformation is α-helical.[9,10] The trend for the alanine C_β chemical shifts is just the opposite; β-sheets shift the C_β peak downfield, while α-helices shift the peak upfield. Table 2 lists the chemical shifts of alanine for various conformations.

Using these techniques, we have previously established that almost all of the alanines of major ampullate gland silk reside in β-sheet conformations. About 40% of the β-sheets are highly oriented (±3°) with respect to the fiber axis, while the other 60% are also crystalline but less well oriented (±60°). It is hypothesized that these poorly oriented β-sheets provide the coupling between the highly oriented crystallites and the glycine-rich regions. It is this combination that is believed to give major ampullate gland silk its exceptional toughness.[12]

Although the structure, conformation, and mechanical properties of *major* ampullate gland silk fibers from *N. clavipes* have been well-studied by solid state NMR[12,13] and by X-ray diffraction,[14] until now NMR studies of *minor* ampullate gland silk have been stymied by several factors, including the difficulty in collecting sufficient amounts of this silk. We report here the first solid state ^{13}C NMR spectra of minor ampullate gland silk. Using a home-built solid state spectrometer optimized for small samples, we now show that it is possible to obtain meaningful spectra from very small amounts (ca. 2.6 mg) of minor ampullate gland silk. By studying the conformation of different kinds of silks (*i.e.*, major and minor ampullate gland silks) and correlating them with amino acid sequence and mechanical properties, it is anticipated that one can obtain a molecular-level understanding of the relationship between silk protein structure and its mechanical properties.

[†] Cornell University.
[‡] University of Wyoming.
[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1997.

S0024-9297(97)00834-6 CCC: $14.00

It is thought that major ampullate silk contains two closely related proteins MaSp1 and MaSp2.[4] These proteins are characterized by the 32 amino acid consensus repeat GGAGQGGYGGLGGQGAGRGGLGGQ-GAGAAAAA of MaSp1 and by the 37 amino acid consensus repeat $(GPGGYGPGQQ)_2GPSGPGS(A)_{10}$ of MaSp2. Genetically engineered proteins can be made on the basis of these consensus repeats.[7] We report here the solid state ^{13}C NMR spectra of a genetically engineered protein based on the consensus sequence of MaSp2. By comparing its spectrum with that of major ampullate gland fibroin, we hope to elucidate the conformational similarities between the genetically engineered protein MaSp2 expressed in *Eschirichia coli* and the fibroin of the major ampullate gland of *N. clavipes*. If the two proteins have similar alanine conformations, it is thought that it may be possible first to understand the spider's natural spinning process and then adapt these methods to process genetically engineered protein into strong, viable biofibers.

Like major ampullate gland silk, minor ampullate gland silk is thought to comprise two proteins, MiSp1 and MiSp2.[15] Both proteins contain highly repetitive alanine and glycine rich regions. The primary difference is that the alanines of minor ampullate silk are not all found in polyalanine runs but instead have very short (3−5 residue) polyalanine runs with glycine−alanine repeats mixed. However, unlike major ampullate gland silk, the minor silk also contains serine rich spacer regions.

Experimental Section

Silk Samples. Major and minor ampullate gland silks were obtained from repeated silking of adult female *N. clavipes* spiders by manual extraction and spooling according the method of Work.[16] The oriented fibers (50 mg of major and 2.6 mg of minor) were randomized for solid state NMR measurements.

Genetically engineered silklike protein corresponding to 16 repeats of the consensus sequence of MaSp2 was purified from *E. coli* according to previously published methods.[7] The protein was flanked by a polyhistidine handle (for purification purposes) on one side and an additional arginine residue on the other. 38 mg of this protein, lyophilized from aqueous solution, was used for NMR measurements.

Characterization. Amino acid analysis for both major and minor ampullate samples were carried out on a Pico-Tag amino acid analysis system.[17] Duplicate samples were hydrolyzed in propionic acid/HCl 50/50 at 150 °C for 90 min. Appropriate blanks, controls, and standards were hydrolyzed in the same vessel as a batch hydrolysis. The amino acids were derivatized with phenylisothiocyanate and analyzed by reversed-phase HPLC.[18,19] The resulting (phenylthio)carbamyl amino acid derivatives were separated on 4.6 × 300 mm Nova Pack C_{18} column employing a modified Pico-Tag buffer system. The mole percent averages of the abundant amino acids for major and minor ampullate gland silk are given in Table 1. The amino acid composition of MaSp2 is taken from previously published data.[7]

Mechanical testing of the minor ampullate gland silk was performed on an Instron in a climate-controlled room (70 F, 65% RH). All samples were equilibrated at least 12 h before testing. All samples have a gauge length of 5 mm and were tested with a crosshead speed of 2 mm/min. The minor ampullate data set consisted of seven single fiber samples.

The errors for each mechanical property are the standard deviation from the mean for each of the samples. It should be noted, as mentioned by Cuniff,[20] that errors in determining fiber diameter have been the largest factor in the uncertainty in the mechanical properties. Fiber diameter was determined by viewing the samples under a phase contrast microscope with a calibrated stage micrometer. The diameter of the minor ampullate gland silk was found to be 2.5 ± 0.2 μm. We note that the uncertainty propagated from the error in diameter is of the same order as the quoted standard deviation in the mechanical properties.

Table 1. Amino Acid Composition (Mole Percent) of Major Ampullate, Minor Ampullate, and the Genetically Engineered Silks

amino acid[a]	major amp silk[b]	minor amp silk[b]	MaSp2[c]
GLY	45	38	34
ALA	29	28	23
SER	6	3	5
GLX	5	10	11
TYR	5	3	6
PRO	1	1	15
LEU	3	4	0

[a] Only those amino acids which are present in greater than 4 mol % for any one sample are listed. [b] Amino acid analysis perfomed in duplicate. Average values are shown. [c] Data taken from ref 7.

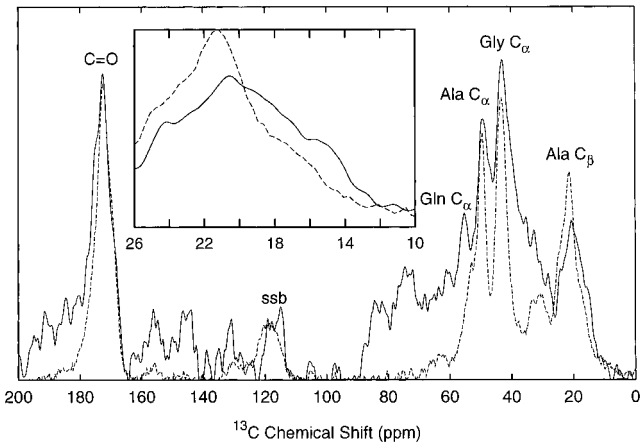

Figure 1. CP-MAS ^{13}C spectra: (solid line) minor ampullate gland silk; (dashed line) major ampullate gland silk. Inset: alanine C_β region. ssb denotes a spinning side band.

NMR Measurements. Solid state ^{13}C spectra were obtained on a home-built spectrometer operating at 90.556 MHz. A Doty magic angle spinning probe was used. A 4.1 mm ZrO_2 rotor was used with Kel-F end caps. All spectra were obtained using a cross polarization pulse sequence with magic-angle spinning at 5 kHz. Cross polarization was achieved by satisfying the Hartman−Hahn condition at 53.9 kHz with a $^{13}C-^{1}H$ contact time of 2.5 ms. The Hartman−Hahn match was optimized using an adamantane sample. A 5.2 μs pulse was used for the initial 90° pulse. During the ^{13}C acquisition, the protons were decoupled at 95 kHz. With a sweep width of 30 kHz and a relaxation delay of 2 s, spectra for the minor ampullate and MaSp2 silks were obtained with 33 000 and 20 000 scans, respectively. Chemical shifts were referenced to adamantane before and after the spectrum of each silk sample was obtained; no drift was observed. The adamantane reference peaks were taken as 29.5 and 38.56 ppm from TMS.

Results and Discussion

Solid state ^{13}C spectra of the minor ampullate silk and that of major ampullate silk are shown in Figure 1. Inspection of the upfield region (ca. 15−30 ppm), which corresponds to the alanine methyl (C_β) carbon, reveals substantial differences between the two types of silks. Because the chemical shifts of amino acid residues are indicative of local conformation,[8,9] it is possible to deduce the local secondary structure for the alanine groups. Table 2 lists the chemical shifts of alanine for various conformations and compares these shifts to those found for the silk proteins. As pointed

Table 2. Chemical Shifts for Carbons in Major Ampullate Silk, Minor Ampullate Silk, and a Genetically Engineered Silk Protein along with Reference Chemical Shifts for Alanine in Both β-Sheets and α-Helices

carbon	α-helix[a]	random coil[a]	β-sheet[a]	major amp silk[b]	minor amp most intense peak[c]	minor amp range[c]	major amp gland fibroin[d]	genetically engineered silk protein[e]
Ala C_β	15.1	17.1	20.1	20.6	20.8	20.8−14.4	17.2	17.7
Ala C_α	52.5	50.6	48.7	49.0	48.5	48.1−51.2	52.4, 48.8	48.1
Ala C′	176.5	175.2	171.9	172.3	172.4	172.9−177	172.2	173.0

[a] Polyalanine chemical shifts for various conformations.[8,9] [b] Solid state ^{13}C NMR of major ampullate silk.[13] [c] Solid state ^{13}C NMR of minor ampullate silk; this data. [d] Solid state ^{13}C NMR of major ampullate gland fibroin.[22] [e] Solid state ^{13}C NMR of genetically engineered protein based on consensus sequence of MaSp2; this data.

Table 3. Mechanical Properties of Major and Minor Ampullate Silk

	stress at max. load (MPa)	modulus (GPa)	strain at max. load (%)
major amp silk[a]	1100	22	9
minor amp silk[b]	346 ± 53	3.02 ± 0.6	30 ± 15

[a] Data taken from Cunniff et al.[20] [b] Average values from seven single fiber samples.

out by Simmons,[13] the chemical shifts for alanine in *major* ampullate silk indicate that almost all the alanines are present in β-sheets. Furthermore, as hypothesized by Simmons,[12] the presence of the alanine rich β-sheets is critical to the exceptional strength properties of the silk fibers.

In contrast to what is observed in major ampullate gland silk, the alanine residues of minor ampullate gland silk are found in several environments (Table 2, Figure 1). Peak assignments were made by comparison of spectra to data on the chemical shifts of peptides.[8,21] In minor ampullate gland silk, the C_β carbons of alanine produce signals ranging from 20.8 to 14.4 ppm. Furthermore, the C_α signal ranges from 48.1 to 51.2 ppm. These shifts for both C_α and C_β alanine carbons give evidence for the presence of both β-sheet and non-β-sheet conformations. The most intense C_α and C_β peaks are found at 48.5 and 20.8 ppm, respectively. These shifts are indicative of β-sheets and suggest that the largest homogeneous population is in β-sheets. However, the existence of a strong signal from 19 to 15 ppm indicates the presence of non-β-sheet conformations in the minor ampullate silk, as well. The analysis of the carbonyl peak furthers this assessment. The majority of the signal appears at 172.9 ppm, but there is also a weaker signal from 173 out to 177.0 ppm. These weaker downfield shifts of the C′ are characteristic of alanine non-β-sheet conformations.[8,9] A rough estimate based on integrated peak areas indicates about half the alanines reside in β-sheets and the rest in non-β-sheet conformations. The secondary structure of minor ampullate silk, stands in stark contrast to major ampullate silk where nearly all the alanines are found in β-sheets.

This result is interesting for two reasons. First, referring to the mechanical data of Table 3, we note that the minor ampullate gland silk has a lower tensile strength than does major ampullate gland silk. It is interesting to speculate that the lower stress and modulus correlate with the appearance of populations of alanines that are not all in β-sheet conformations. Secondly, the existence of conformations other than pure β-sheet conformation could increase the maximum strain of the silk. It may also be that the spacer regions in minor ampullate gland silk also contribute to the strain properties. As seen in Table 3, minor ampullate silk displays a higher maximum strain than that of major ampullate gland silk. This strain is not reversible and thus represents deformation, not elasticity.

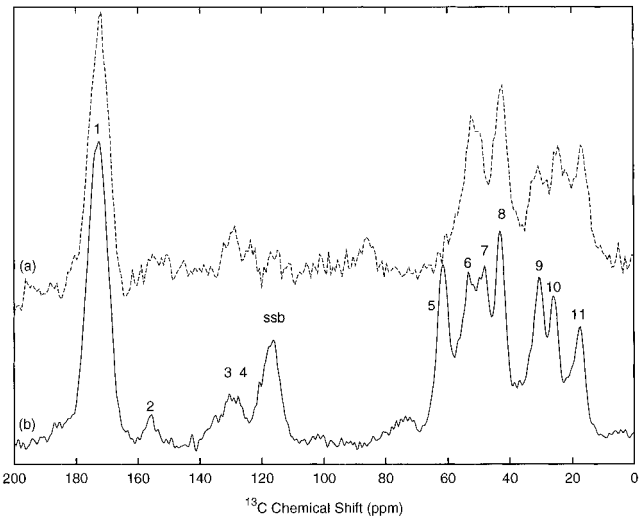

Figure 2. CP-MAS ^{13}C spectra of (a) lyophilized native major ampullate gland fibroin from ref 22 and (b) lyophilized genetically engineered protein based on MaSp2. ssb denotes a spinning side band. Numbered peaks are identified with the following carbons: (1) C═O; (2) Tyr C_ζ; (3) Tyr C_δ; (4) Tyr C_γ; (5) Pro C_α; (6) Gln C_α, Ser C_α; (7) Ala C_α; (8) Gly C_α; (9) Gln C_γ, Pro C_β; (10) Gln C_β; (11) Ala C_β.

Figure 2 shows a comparison between lyophilized MaSp2 and flash-frozen, lyophilized major ampullate gland fibroin. For MaSp2, we find the chemical shifts of the alanine C_α and C_β carbons are at 48.1 and 17.7 ppm (Table 2). The alanine chemical shifts of the genetically engineered silklike protein (Figure 2b) do not suggest any of the three common conformations and are somewhat similar, but not identical to those in the flash-frozen then lyophilized fibroin of the major ampullate gland (Figure 2a). In the fibroin, the splitting in signal from the C_α carbon of alanine is interpreted as frozen-out conformations that were averaging between two states.[21]

It should be noted that there are subtle line-width differences between the lyophilized genetically engineered protein (Figure 2b) and the flash-frozen and then lyophilized fibroin (Figure 2a). While lyophilization can cause broader line widths due to trapped conformations, the sequence homogeneity of the genetically engineered material may produce narrower lines.

The chemical shifts of the C_β carbon of alanine are the same for the two samples, but the chemical shifts of the C_α carbon of alanine differ. In order to faithfully mimic the gland fibroin, fiber engineers may need to learn how to maintain a genetically engineered protein like MaSp2 in a metastable state, just as the spider maintains its fibroin in the major ampullate gland.

Taken together, the data presented here illustrate the role of primary structure, particularly the polyalanine repeats, in producing the secondary structures found in native spider silks. These same secondary structures

are not faithfully reproduced when the native fibroin is flash frozen and lyophilized or when genetically engineered silks are lyophilized. These results underscore the importance of the processing steps.

A correlation between primary and secondary structures and mechanical properties for the native major and minor ampullate gland silks suggests that the polyalanine β-sheet repeats are at least partly responsible for the stress and modulus. It is inferred that the spacer regions in the minor ampullate gland silk contribute to its strain.

Acknowledgment. We gratefully acknowledge the support of the National Science Foundation grants MCB-9601018 and DMR-9708062. In addition, Oskar Liivak acknowledges the support of the NIH Training Grant in Molecular Physics of Biological Systems (NIH T32GM08267).

References and Notes

(1) Amato, I. *Science* **1991**, *253*, 966.
(2) Shear, W. A.; Palmer, J. M.; Coddington, J. A.; Bonamo, P. M. *Science* **1989**, *246*, 479.
(3) Xu, M.; Lewis, R. V. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7120.
(4) Lewis, R. V. *Acc. Chem. Res.* **1992**, *25*, 392.
(5) Mahoney, D. V.; Vezie, D. L.; Eby, R. K.; Adams, W. W.; Kaplan, D. In *Silk Polymers: Materials Science and Biotechnology*; Kaplan, D.; Adams, W. W.; Farmer, B.; Viney, C., Eds.; ACS Symposium Series 544; American Chemical Society: Washington, DC, 1994; p 196.
(6) Prince, J. T.; McGrath, K. P.; DiGirolamo, C. M.; Kaplan, D. L. *Biochemistry* **1995**, *34*, 10879.
(7) Lewis, R. V.; Hinman, M.; Kothakota, S.; Fournier, M. J. *Protein Expression Purif.* **1996**, *7*, 400.
(8) Saito, H. *Magn. Reson. Chem.* **1986**, *24*, 844.
(9) Wishart, D. S.; Sykes, B. D. *J. Biomol. NMR* **1994**, *4*, 171.
(10) Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490.
(11) Oldfield, E. *J. Biomol. NMR* **1995**, *5*, 217.
(12) Simmons, A. H.; Michal, C. A.; Jelinski, L. W. *Science* **1996**, *271*, 84.
(13) Simmons, A.; Ray, E.; Jelinski, L. W. *Macromolecules* **1994**, *27*, 5235.
(14) Grubb, D. T.; Jelinski, L. W. *Macromolecules* **1997**, *30*, 2860.
(15) Colgin, Mark. Ph.D. Thesis, University of Wyoming, 1996.
(16) Work, R. W.; Emerson, P. D. *J. Arachnol.* **1982**, *10*, 1.
(17) Cohen, S. A.; Meys, M.; Tarvin, T. L. In *The Pico-Tag Method: A Manual of Advanced Techniques for Amino Acid Analysis*; Millipore Corp.: Bedford, MA, 1989; p 2.
(18) Heinrikson, R. L.; Meredith, S. C. *Anal. Biochem* **1984**, *136*, 65.
(19) Bidlingmeyer, B. A.; Cohen, S. A.; Tarvin, T. L. *J. Chromatogr* **1984**, *336*, 93.
(20) Cunniff, P. M. In *Silk Polymers: Materials Science and Biotechnology*; Kaplan, D., Adams, W. W., Farmer, B., Viney, C., Eds.; ACS Symposium Series 544; American Chemical Society: Washington, 1994; p 234.
(21) Howarth, O.; D. M. J. Lilley. *Prog. NMR Spectrosc.* **1978**, *12*, 1.
(22) Hijirida, D.; Do; K. G.; Michal; C.; Wong, S.; Zax, D.; Jelinski, L. W. *Biophys. J.* **1996**, *71*, 3444.

MA970834A