Conformation of Spider Silk Proteins In Situ in the Intact Major Ampullate Gland and in Solution

Thierry Lefèvre,[†] Jérémie Leclerc,[†] Jean-François Rioux-Dubé,[†] Thierry Buffeteau,[§] Marie-Claude Paquin,[†] Marie-Eve Rousseau,[†] Isabelle Cloutier,[†] Michèle Auger,[†] Stéphane M. Gagné,[⊥] Simon Boudreault,[‡] Conrad Cloutier,[‡] and Michel Pézolet*[,†]

Department of Chemistry, Department of Biology, and Department of Biochemistry and Microbiology, Université Laval, Québec (Québec) G1K 7P4 Canada, and Institut des Sciences Molèculaires (UMR 5255 CNRS), Université Bordeaux 1, Talence 33405 cedex France

Received May 18, 2007; Revised Manuscript Received June 27, 2007

To understand the spinning process of dragline silk by spiders, the protein conformation before spinning has to be determined. Raman confocal spectromicroscopy has been used to study the conformation of the proteins in situ in the intact abdominal major ampullate gland of *Nephila clavipes* and *Araneus diadematus* spiders. The spectra obtained are typical of natively unfolded proteins and are very similar to that of a mixture of recombinant silk proteins. Vibrational circular dichroism reveals that the conformation is composed of random and polyproline II (PPII) segments with some α-helices. The α-helices seem to be located in the C-terminal part whereas the repetitive sequence is unfolded. The PPII structure can significantly contribute to the efficiency of the spinning process in nature.

The spinning process of dragline silk by spiders consists of the conversion at ambient temperature and pressure of a highly viscous and concentrated protein aqueous solution into an insoluble fiber. This filament is produced by the abdominal major ampullate (MA) gland of orb-weaving spiders and exhibits remarkable mechanical properties.[1] There is currently a great interest in understanding this natural transformation to help in the development of bioinspired materials having comparable performances. However, to understand this process, it is crucial to have as much information as possible on the structure of the proteins before and after spinning. Whereas the final dragline fiber is relatively well characterized, little is known about the native protein conformation in the MA gland.

A key problem stems from the fact that the dope solution stored in the MA gland is metastable so that the proteins are readily converted to β-sheets upon handling. It is thus of paramount importance to find a technique that allows the direct investigation of the protein conformation in the feedstock solution contained in the intact gland. To date, this problem has not been solved, and the conformation of the silk proteins before spinning remains unclear. Hijirida et al.[2] used ^{13}C NMR, electronic circular dichroism (ECD), and infrared spectroscopy to examine the conformation of MA gland silk fibroin within the glands of the spider *Nephila clavipes* (*N. clavipes*) and concluded that silk fibroin in the glands exists in dynamically averaged helical conformations. More recent ^{1}H and ^{13}C NMR results on the native liquid silk stored within the lumen of a MA silk gland led to the conclusion that each amino acid of a given type lies in an identical environment, which was found to be consistent with a random coil structure.[3] ECD results of the spinning dope diluted in water and silk fibers dissolved in aqueous solution indicate that the protein conformation is mainly unordered but may contain various secondary structures such as α-helix, β-sheet, turn, and polyproline II (PPII) helix.[4]

Here, Raman confocal spectromicroscopy has been used successfully to study the native protein conformation in situ in intact MA glands of the spiders *N. clavipes* and *Araneus diadematus* (*A. diadematus*). The spectra obtained were compared with that of a mixture of the proteins MaSpI and MaSpII found in the dragline silk of *N. clavipes* using the recombinant proteins produced by Nexia Biotechnologies, Inc.[5] The secondary structure of these two proteins was further analyzed by vibrational circular dichroism (VCD) and ^{1}H NMR.

Figure 1 shows the Raman spectra in the amide I and amide III regions of the MA gland lumen of *N. clavipes* and *A. diadematus* collected with a 50× objective, with the laser beam being focused 50−100 μm below the surface of the gland epithelium. It is remarkable that spectra with very good signal-to-noise ratios could be obtained in situ through the gland membrane. For both spider species, the spectra have identical shapes in the amide I and amide III regions, indicating very similar protein conformations. They are totally different from those of a dragline silk fiber (Figure 1) in terms of position, shape, and bandwidth. For the fiber, the amide I and amide III bands appear at 1670 and 1229/1241 cm^{-1}, respectively, which is characteristic of β-sheets,[6] the dominant secondary structure in dragline silk. In addition, the amide I band is narrow, showing a rather high degree of homogeneity of the secondary structures present. On the contrary, the amide I band of the spinning dope is broad, showing the presence of various secondary structures before spinning. For the spinning dope, the amide I and amide III bands are located at 1660 and 1260 cm^{-1}, respectively, and are typical of unfolded proteins.[7] Other smaller differences can also be observed. The bands at 1303 and 1338 cm^{-1} in the dope are associated with the α-helix conformation,[8,9] indicating that the spidroins adopt a small but significant amount of α-helices. These bands are weak or absent in dragline fibers.

* Corresponding author.
[†] Department of Chemistry, Université Laval.
[‡] Department of Biology, Université Laval.
[⊥] Department of Biochemistry and Microbiology, Université Laval.
[§] Université Bordeaux 1.

10.1021/bm7005517 CCC: $37.00
Published on Web 07/21/2007

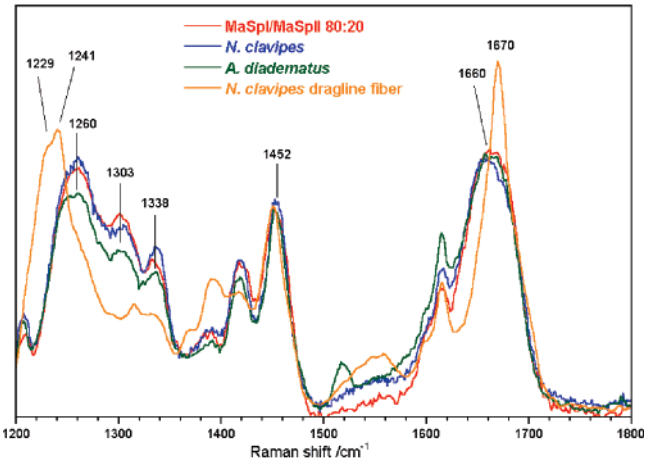

Figure 1. Raman spectra of the MA gland lumen of *N. clavipes* and *A. diadematus*, compared with a 80:20 MaSpI/MaSpII aqueous mixture (5% w/v of proteins, pH 9) and the dragline silk fiber of *N. clavipes*. The latter spectrum is the orientation-insensitive one calculated from polarized spectra as described elsewhere.[10] The spectra are normalized with respect to the peak maximum of the 1452 cm^{-1} band.[8,11]

To mimic the spinning dope of the MA gland, the recombinant proteins MaSpI and MaSpII have been studied in water at a molar ratio of 80:20 (the ratio found for *N. clavipes* dragline fiber by Brooks et al.[12] and from our own measurements using NMR spectroscopy of solubilized *N. clavipes* dragline fibers (unpublished results)). The spectrum recorded in vitro is almost identical to that of the gland lumen, particularly in the amide I and III regions. Therefore, although there is a dramatic difference in the protein concentration between the aqueous solution (5% w/v) and the MA gland (minimum 30%), the recombinant proteins appear to be a good model to further investigate the native structure of silk proteins before spinning.

Figure 2A presents the VCD spectra in the amide I′ region of MaSpI and MaSpII in D_2O at pD 7 and 11, respectively. MaSpII was investigated at high pD to avoid spontaneous protein aggregation. The spectra exhibit a strong negative component at 1639 cm^{-1} (MaSpI) or 1630 cm^{-1} (MaSpII), clearly revealing a significant contribution from the PPII structure (left-handed 3_1 helix).[13] The expected positive contribution at ~1660 cm^{-1} typical of PPII is hidden for MaSpI by another band due to α-helices (see below). For comparison, the spectrum of the seven Ala peptide acetyl-XXAAAAAAAOO-amide (XAO) is presented (X refers to diaminobutyric acid, and O refers to ornithine). This peptide is a model of the PPII structure[14] and is very representative of the Ala segments of the silk proteins (see Supporting Information). The comparison with XAO further supports the propensity of MaSpI and MaSpII to form PPII. As seen in Figure 2A, the magnitude of the VCD spectra of MaSpI and MaSpII is about one-third that of XAO, indicating that part of the backbone of these proteins does not contribute to the VCD signal. Therefore, some amino acid residues seem to be totally unordered. The negative band at 1639 cm^{-1} for MaSpI is also similar to that of long homopolypeptides such as poly-L-lysine and poly-L-glutamic acid, which are known to adopt the PPII structure.[13,15] The lower position of this feature for MaSpII is attributed to the presence of 15% of Pro residues in the sequence since the position of this band for poly-L-proline is 1615 cm^{-1} (MaSpI is devoid of Pro).[13] This result shows that the Pro residues of MaSp2, as well as other residues, form PPII segments.

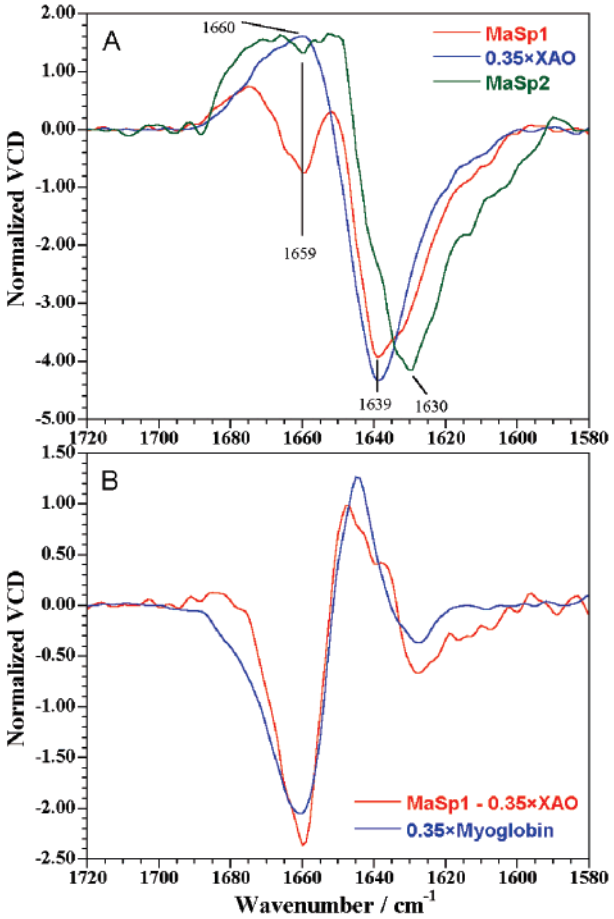

Figure 2. (A) VCD spectra in the amide I′ region of MaSpI, MaSpII, and the peptide XAO in D_2O solution (5%, 5%, and 3% w/v, respectively). The VCD spectra were normalized so that all samples have their amide I′ absorbance maximum equal to 1. (B) Comparison between the VCD spectrum of myoglobin and the difference spectrum obtained by subtracting the spectrum of XAO from that of MaSpI.

Besides this spectral difference, the overall conformations of MaSpI and MaSpII are basically the same and contain mostly random and PPII structures with some α-helices. Interestingly, 3_1 helices have also been detected in the amorphous region of the solid fibers.[16] Since the PPII and such 3_1 helices correspond to the same structure, the presence of the PPII segments in the spinning dope suggests that residual 3_1 helices are conserved after the spinning process. The presence of random and PPII structures is the basic characteristic of natively unfolded proteins.[17] The VCD spectrum of MaSpI exhibits a small negative feature near 1659 cm^{-1}. As seen in Figure 2B, the difference spectrum calculated by subtracting the XAO spectrum from that of MaSpI is typical of the α-helix and is very similar to that of the highly helical protein myoglobin. Figure 2A also shows that, for MaSpII, the α-helix feature is less well-defined due to the basic pH. On the basis of the secondary structure prediction of the C-terminal regions of MaSpI and MaSpII (data not shown) and on ECD results on recombinant dragline proteins of the spider *A. diadematus*,[18] it appears that the α-helices of the spidroins are mainly located in the C-terminal region of the spidroins.

The H^{α} chemical shifts of different amino acid residues of MaSpI/MaSpII in aqueous solution (0.5% w/v) were also determined by ^1H NMR spectroscopy (Table 1).

These chemical shifts are very different from those found for the α-helices and β-sheets and are close to those of PPII

Table 1. Chemical Shifts of the $^1H^\alpha$ Atom of Recombinant Proteins and Models of Secondary Structure[a]

	GGXGG δ (ppm)[b]	α-helix δ (ppm)[c]	β-sheet δ (ppm)[c]	MaSpI δ (ppm)	MaSpII δ (ppm)	MaSpI δ (ppm)	MaSpII δ (ppm)	XAO δ (ppm)
Ala	4.34	4.05	4.67	4.25	4.25	4.33	4.33	4.26
Arg	4.34	3.99	4.65	4.34		4.34		
Gln	4.36	4.02	4.88	4.34	4.34	4.35	4.37	
Pro trans	4.44	4.12	4.79		4.42		4.44	
Leu	4.35	3.98	4.69	4.35		4.35		
Tyr	4.56	4.15	5.03	4.55	4.57	4.55	4.58	

[a] Some of the chemical shifts were corrected (δ_R) using the sequence-dependent correction factors given by Schwarzinger et al.[19] The details of the chemical shift corrections are given in the Experimental Section. [b] Reference 20. [c] Reference 21.

measured on GGXGG peptides (X refers to the 20 natural amino acid residues).[20] Therefore, it appears that these silk protein residues have a strong tendency to adopt the PPII conformation, in agreement with the VCD data. Finally, the chemical shifts of the H^α atom of Ala residues in MaSpI and MaSpII are very close to that of the central Ala residue of XAO, indicating that the polyalanine motifs of the proteins are clearly involved in PPII. The formation of PPII helices by the Ala segments of the proteins has been suggested previously on the basis of molecular simulations on Ala-based peptides.[22]

Taken together, our results show that the silk proteins in the sac of the MA gland are typical of natively unfolded proteins and have a propensity for the PPII conformation, especially in the repetitive sequences. This conclusion is in line with recent findings relating unfolded proteins and the PPII helix.[17] The dragline silk proteins are subject to two opposite constraints. On one hand, they have to be stable and avoid premature spontaneous self-aggregation in the storage MA gland (in spite of temperature variations due to outdoor conditions), and, on the other hand, they have to readily form β-sheets during the spinning process to optimize the mechanical properties of the fibers. We believe that the presence of the PPII structure may be crucial to solve this conflicting situation. First, the presence of the PPII structure and the strong propensity to be natively unfolded can account for the high concentration of silk proteins and the long-term stability of the spinning dope. The extended nature of the PPII helix indeed precludes intrachain hydrogen bonds, with the structure being stabilized instead by main-chain hydrogen bonding with water molecules. Second, the PPII conformation may play a major role in the spinning process. The dihedral angles of PPII being very close to those of a β-strand,[23] the energy barrier for the formation of β-sheets may be small. Thus, silk protein backbones, particularly the Ala sequences, can readily convert into β-sheet and then undergo self-assembly via β-aggregation as the dope flows into the duct and as shear forces and extensional flow occur.

Acknowledgment. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Fonds Québécois de Recherche sur la Nature et les Technologies (FQRNT). We thank Nexia Biotechnologies, Inc. for the generous supply of the recombinant proteins.

Supporting Information Available. Experimental section, repetitive sequence of MaSpI and MaSpII (Figure S1), and the corresponding references. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Denny, M. In *The Mechanical Properties of Biological Materials*; Vincent, J. F. V., Currey, J. D., Eds.; Cambridge University Press: Cambridge, U.K., 1980; Vol. 34, pp 246−272. Gosline, J. M.; DeMont, M. E.; Denny, M. *Endeavour* **1984**, *10*, 37−43.
(2) Hirijida, D. H.; Do, K. G.; Wong, S.; Zax, D.; Jelinski, L. W. *Biophys. J.* **1996**, *71*, 3442−3447.
(3) Hronska, M.; van Beek, J. D.; Willimason, P. T. F.; Vollrath, F.; Meier, B. H. *Biomacromolecules* **2004**, *5*, 834−839.
(4) Dicko, C.; Knight, D.; Kenney, J. M.; Vollrath, F. *Biomacromolecules* **2004**, *5*, 2105−2115. Dicko, C.; Knight, D.; Kenney, J. M.; Vollrath, F. *Biomacromolecules* **2004**, *5*, 758−767. Shao, Z.; Vollrath, F.; Yang, Y.; Thøgersen, H. C. *Macromolecules* **2003**, *36*, 1157−1161.
(5) Lazaris, A.; Arcidianacono, S.; Huang, Y.; Zhou, J.-F.; Duguay, F.; Chretien, N.; Welsh, E. A.; Soares, J. W.; Karatzas, C. N. *Science* **2002**, *18*, 472−476.
(6) Rousseau, M.-E.; Lefèvre, T.; Beaulieu, L.; Asakura, T.; Pézolet, M. *Biomacromolecules* **2004**, *5*, 2247−2257. Dong, J.; Wan, Z.; Popov, M.; Carey, P. R.; Weiss, M. A. *J. Mol. Biol.* **2003**, *330*, 431−442. Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181−364.
(7) Maiti, N. C.; Apetri, M. M.; Zargoski, M. G.; Carey, P. R.; Anderson, V. E. *J. Am. Chem. Soc.* **2004**, *126*, 2399−2406.
(8) Frushour, B. G.; Koenig, J. L. *Biopolymers* **1974**, *13*, 1809−1819.
(9) Pezolet, M.; Pigeon, M.; Ménard, D.; Caillé, J.-P. *Biophys. J.* **1988**, *53*, 319−325. Rousseau, M. E.; Beaulieu, L.; Lefevre, T.; Paradis, J.; Asakura, T.; Pézolet, M. *Biomacromolecules* **2006**, *7*, 2512−2521.
(10) Lefevre, T.; Rousseau, M. E.; Pézolet, M. *Appl. Spectrosc.* **2006**, *60*, 841−846.
(11) Yu, T.-J.; Lippert, J. L.; Peticolas, W. L. *Biopolymers* **1973**, *12*, 2161−2176.
(12) Brooks, A. E.; Steinkraus, H. B.; Nelson, S. R. *Biomacromolecules* **2005**, *6*, 3095−3099.
(13) Dukor, R. K.; Keiderling, T. A. *Biopolymers* **1991**, *31*, 1747−1761.
(14) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190−9195.
(15) Yasui, S. C.; Keiderling, T. A. *J. Am. Chem. Soc.* **1986**, *108*, 5576−5581.
(16) Kümmerlen, J.; van Beek, J. D.; Vollrath, F.; Meier, B. H. *Macromolecules* **1996**, *29* (8), 2920−2928. Eles, P. T.; Michal, C. A. *Biomacromolecules* **2004**, *5*, 661−665. Lefevre, T.; Rousseau, M.-E.; Pezolet, M. *Biophys. J.* **2007**. Marcotte, I.; van Beek, J. D.; Meier, B. H. *Macromolecules* **2007**, *40*, 1995−2001.
(17) Rose, G. D. *Unfolded Proteins*; Academic Press: San Diego, CA, 2002; Vol. 62.
(18) Huemmerich, D.; Helsen, C. W.; Quedzuweit, S.; Oschmann, J.; Rudolph, R.; Scheibel, T. *Biochemistry* **2004**, *43*, 13604−13612.
(19) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, J. *J. Am. Chem. Soc.* **2001**, *123*, 2970−2978.
(20) Merutka, G.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **1995**, *5*, 14−24. Shi, Z.; Chen, K.; Liu, Z.; Ng, A.; Bracken, W. C.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17964−17968.
(21) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311−333.
(22) Lewis, R. V. *Chem. Rev.* **2006**, *106*, 3762−3774.
(23) Blanch, E. W.; Gill, A. C.; Rhie, A. G. O.; Hope, J.; Hecht, L.; Nielsen, K.; Barron, L. D. *J. Mol. Biol.* **2004**, *343*, 467−476.

BM7005517