Polymer Journal, Vol. 36, No. 12, pp. 999—1003 (2004)

SHORT COMMUNICATIONS

Structure of Characteristic Sequences in *Nephila clavipes* Dragline Silk (MaSp1) Studied with ^{13}C Solid State NMR

Tetsuo ASAKURA,† Mingying YANG, and Taiji KAWASE

Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei 184-8588, Japan

(Received August 16, 2004; Accepted September 17, 2004; Published December 15, 2004)

KEY WORDS *Nephila clavipes* Dragline Silk / Conformation-dependent ^{13}C NMR Chemical Shift / 3_1-Helix / α-Helix / β-Sheet /
[DOI 10.1295/polymj.36.999]

The dragline silk of the golden weaver *Nephila clavipes* is composed of two proteins, designated spidroin 1 (MaSp1) and spidroin 2 (MaSp2), and has been the focus of numerous recent studies because they are among the strongest natural protein fibers.[1–14] The dominant MaSpl protein can be described as a block-copolymer consisting of polyalanine poly(Ala) runs and glycine-rich sections containing relatively high concentrations of amino acids with bulky, hydrophilic side chains. Several solid state NMR studies[3–13] have been applied to clarify the structure for native spider silk fibers. The poly(Ala) region in dragline silk fiber took β-sheet structure, which was derived from the ^{13}C chemical shifts of Ala residues clearly.[4–8] However, the structure of the glycine-rich region is not clear because of heterogeneity of the sequences. Thus, it seems difficult to study the precise structure of the native spider silk fiber using solid state NMR and should consider the large structural distributions, especially in the glycine-rich region.[5–9]

The structural analysis using NMR for the model peptides seems very effective because it is possible to prepare several model peptides with well-defined sequences by reference of the primary structure of MaSp1. However, there are no such researches with appropriate model peptides for the spider silk. In our previous papers,[15–17] the model peptides selected from the primary structure of the silk fibroins from *Bombyx mori* (*B. mori*) and *Samia cynthia ricini* (*S. c. ricini*) were effectively used to obtain the structural information using several solid state NMR methods. Especially, the primary structure of the silk fibroin from *S. c. ricini* is the repeated sequences which consist of alternative appearance of poly(Ala) region and glycine-rich region, which is similar to the primary structure of the MaSpl protein of spider silk.[17–19]

In this paper, 49 mer peptide with the sequence, GGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAAAAAAGGAGQG selected from the primary structure of the MaSp1 protein of from *Nephila clavipes* dragline silk and the ^{13}C isotope-labeled model peptide with the same sequence were synthesized. The structure was studied with ^{13}C CP/MAS NMR, especially, the conformation-dependent ^{13}C chemical shifts. The methods of the three kinds of treatments before NMR observation were changed in order to reproduce the structure before and after spinning of the spider silk. Namely, (1) the peptide was dissolved in trifuoroacetic acid and then dried (TFA treatment), (2) the peptide was dissolved in 9 M LiBr and then dialyzed against water (Dialysis treatment) and (3) the peptide was dissolved in 8 M urea and then precipitated in acetonitrile, and dried (AN treatment).

EXPERIMENTAL

The following peptides, **1**–**3**, were synthesized by the solid-phase method.

Peptide **1**; YGGLGSQGAGRG

Peptide **2**; GGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAAAAAAGGAGQG

Peptide **3**; GGLGGQGAGAAA[3-^{13}C]AAAGG[2-^{13}C]AGQGGYGGL[1-^{13}C]GSQGAGRGGQ[2-^{13}C]GAAAAAAGGAGQG

The ^{13}C CP/MAS NMR spectra were acquired on a Chemagnetics CMX-400 spectrometer operating at 100 MHz, with a CP contact time of 1 ms, TPPM decoupling, and magic angle spinning at 10 KHz. A total of 10,000–20,000 scans for natural abundance peptide samples were collected over a spectral width of 60 kHz, with a recycle delay of 5 s. The chemical shifts

†To whom correspondence should be addressed (E-mail: asakura@cc.tuat.ac.jp).

are reported relative to TMS as a reference. The solution ^{13}C NMR was observed with JEOL α-500 NMR spectrometer.

RESULTS AND DISCUSSION

The primary structure of the MaSp1 from *Nephila clavipes* dragline silk was shown in Figure 1.[1] MaSp1 is composed primarily of a 30-amino acid repetitive motif that is rich in poly(Ala) regions. The poly(Ala) regions are composed of 4–7 consecutive alanine residues between glycine-rich regions. In the process of the design of the model peptide used here, the number of the Ala residues in the poly(Ala) region was set to be 6 as an average and the presence of two poly(Ala) regions was considered. One highly conserved YGGLGSQGAGR sequence which has been proposed as playing a major role in the superconstraction process of the spider silk were included between two poly(Ala) regions.[8] Thus, we synthesized the sequence, GGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAAAAAAGGAGQG. Before the structural analysis of the whole peptide, the structure of the YGGLGSQGAGR sequence was examined. The peptide **1** (YGGLGSQGAGRG) is easily soluble in water, which supports the report that the YGGLGSQGAGR sequence is a major role in the the superconstraction process. The solution ^{13}C NMR was observed in aqueous solution as shown in Figure 2b. All the ^{13}C NMR peaks show random coil chemical shifts. After freeze-drying of the aqueous solution of the peptide **1**, the ^{13}C CP/MAS NMR spectrum was observed (Figure 2a). Both ^{13}C chemical shifts of the peptide **1** in the aqueous solution and in the solid state are essentially the same except for the peak broadening of the solid-state NMR spectrum, indicating that the structure of the peptide **1** in aqueous solution and in the solid state is random coil.

The ^{13}C CP/MAS NMR spectra of the peptide **2** (GGLGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGGQGAAAAAAGGAGQG) after TFA treatment, Dialysis treatment and AN treatment, are shown in Figure 3. As reported previously,[16–19] TFA treat-

N-terminal

	QGAGAAAAAA	GGAGQGGYGGLGGQG
		AGQGGYGGLGGQG
	AGQGAGAAAAAAA	GGAGQGGYGGLGSQG
AGR	GGQGAGAAAAAA	GGAGQGGYGGLGSQG
AGRGGL	GGQGAGAAAAAAA	GGAGQGGYGGLGNQG
AGR	GGQ GAAAAAA	GGAGQGGYGGLGSQG
AGRGGL	GGQ AGAAAAAA	GGAGQGGYGGLGGQG
		AGQGGYGGLGSQG
AGRGGL	GGQGAGAAAAAAA	GGAGQ GGLGGQG
	AGQGAGASAAAA	GGAGQGGYGGLGSQG
AGR	GGEGAGAAAAAA	GGAGQGGYGGLGGQG
		AGQGGYGGLGSQG
AGRGGL	GGQGAGAAAA	GGAGQ GGLGGQG
	AGQGAGAAAAAA	GGAGQGGYGGLGSQG
AGRGGL	GGQGAGAVAAAAA	GGAGQGGYGGLGSQG
AGR	GGQGAGAAAAAA	GGAGQGGYGGLGNQG
AGRGGL	GGQGAGAAAAAAA	GGAGQGGYGGLGNQG
AGR	GGQ GAAAAA	GGAGQGGYGGLGSQG
AGR	GGQGAGAAAAAA	VGAGQEGIR GQG
		AGQGGYGGLGSQG
SGRGGL	GGQGAGAAAAAA	GGAGQ GGLGGQG
	AGQGAGAAAAAA	GGVRQGGYGGLGSQG
AGR		GGAGQGGYGGLGGQG
VGRGGL	GGQGAGAAAA	GGAGQGGYGGV GSG
	ASAASAAAASRLSS	

C-terminal

Figure 1. Primary structure of spidroin 1 (MaSp1) of dragline silk of *Nephila clavipes*.

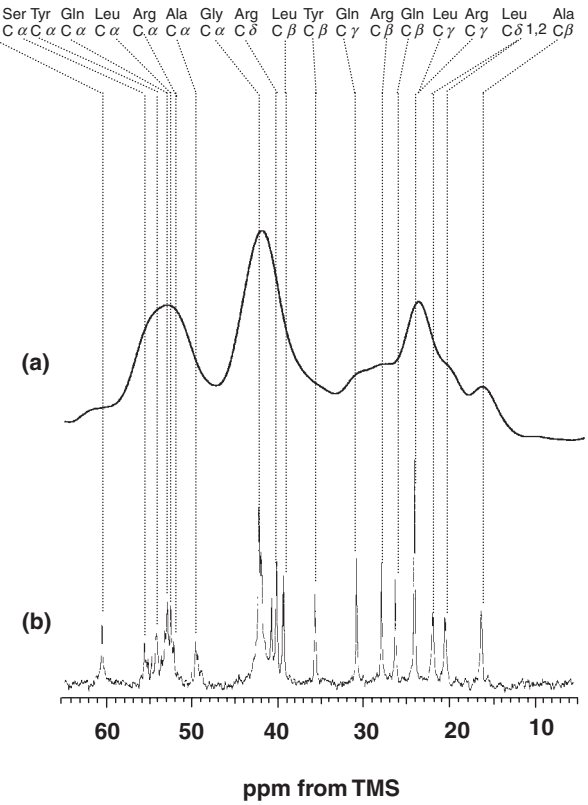

Figure 2. (a) ^{13}C CP/MAS NMR and (b) ^{13}C solution NMR spectra of the peptide **1**, YGGLGSQGAGRG, together with the peak assignment in the solution NMR spectrum. The spectral range, 5–65 ppm, was expanded.

Table I. ^{13}C CP/MAS NMR chemical shifts of the peptide **2**, GGLGGQGAGAAAAAAGGAGQGG-YGGLGSQGAGRGGQGAAAAAAGGAGQG, and the peptide **3**, GGLGGQGAGAAA[3-^{13}C]-AAAGG[2-^{13}C]AGQGGYGGL[1-^{13}C]GSQGAGRGGQ[2-^{13}C]GAAAAAAGGAGQG. The ^{13}C chemical shifts of Ala and Gly residues with typical conformation are also listed

		Ala Cβ	Ala Cα	Gly Cα	Gly C=O
Peptide **2**	(TFA treatment)	15.7	52.1	≈43.0 (broad)	—
	(Dialysis treatment)	16.3	51.2	≈43.0 (broad)	—
	(AN treatment)	20.6	48.9	≈43.0 (broad)	—
Peptide **3**	(AN treatment)	A^{13} 20.4[a]	A^{18} 49.0	G^{37} 42.4	G^{27} 171.5[b]
		A^{13} 23.0			
		A^{13} 16.6			
	Dragline silk	20.6	49.0	≈43.0	—
	Random coil	16.6	50.0	42.9	171.3
	α-helix	15.7	52.5	44.0	172.3
	β-sheet	20.1	48.7	42.4	169.1
	3_1 helix	17.4	48.9	41.6	171.3

[a]The A^{13} peak can be deconvoluted by assuming three components, *i.e.*, 19% at 23.0 ppm, 67% at 20.4 ppm and 14% at 16.6 ppm. [b]A sharp carbonyl carbon peak was observed at 171.5 ppm, although the spectral range from 5 to 65 ppm was shown in Figure 4b.

ment of the the model peptide, GGAGGGYGGDG-G(A)$_{12}$GGAGDGYGAG for *S. c. ricini* silk fibroin induces α-helix conformation and therefore such treatment was also applied to generate α-helix form (Figure 3a). The ^{13}C chemical shifts of Ala Cβ and Ala Cα carbons are 15.7 ppm and 52.1 ppm, respectively as listed in Table I. Since the Ala peaks come mainly from two poly(Ala) regions in the peptide **2**, it is possible to say that two poly(Ala) regions adopt on α-helix structure. On the other hand, β-sheet structure is generated by the AN treatment (Figure 3c). Actually, the peaks of Ala Cβ and Ala Cα carbons shift to 20.6 ppm and 48.9 ppm, respectively, indicating that the structure of poly(Ala) region changes to β-sheet by the AN treatment. In both treatments, the Gly Cα peak does not shift significantly and relatively broard. Judging from the chemical shift, it is likely that Gly residue takes essentially random-coil. The Ala Cα and Ala Cβ chemical shifts of the peptide after dialysis treatment (Figure 3b), are between α-helix and β-sheet chemical shifts, but closer to α-helix (Table I).

In order to obtain further structural information as a model of the spider silk fiber, the peptide **3** (GGLGG-QGAGAAA[3-^{13}C]AAAGG[2-^{13}C]AGQGGYGGL[1-^{13}C]G SQGAGRGGQ[2-^{13}C]GAAAAAAGGAGQG), where four positions were selectively ^{13}C isotope-labeled without changing the sequence of the peptide **2**, was synthesized and the ^{13}C CP/MAS NMR was observed after AN treatment of the peptide (Figure 4b). The Ala residue at the 13th position in the former poly(Ala) region takes β-sheet judging from the chemical shift value 20.4 ppm of the main peak of Ala Cβ carbon, but the line shape is asymmetric. Thus, the peak can be deconvoluted by assuming

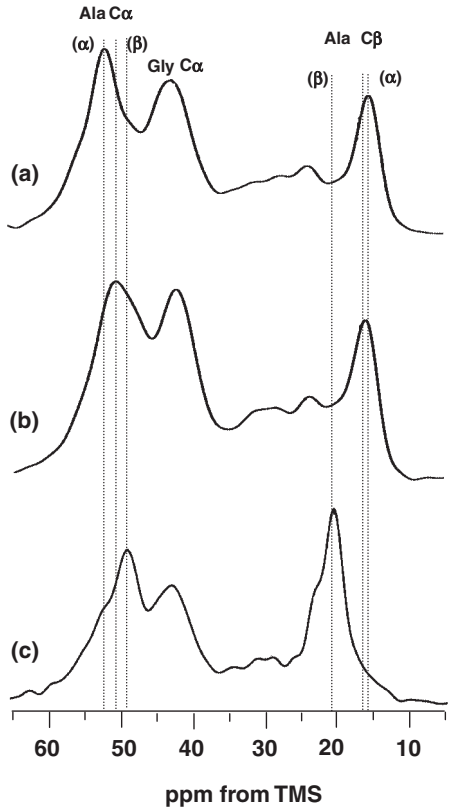

Figure 3. ^{13}C CP/MAS NMR spectra of (a) the peptide **2** after trifluoroacetic acid treatment, (b) the peptide **2** after dialysis treatment and (c) the peptide **2** after acetonitrile treatment. The spectral range, 5–65 ppm, was expanded.

three components, *i.e.*, 19% at 23.0 ppm, 67% at 20.4 ppm and 14% at 16.6 ppm as shown in Figure 5. This is similar to the corresponding spectra of the crystalline fraction of *B. mori* silk fibroin and the model peptide, (AG)$_{15}$ with silk II (β-sheet) form as

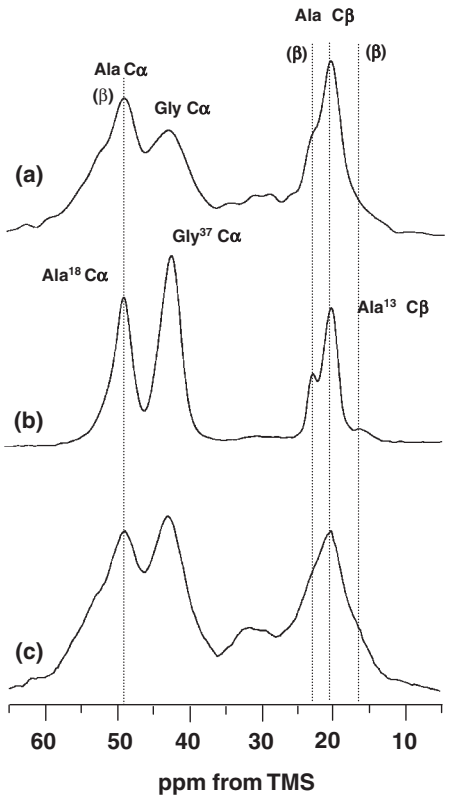

Figure 4. ^{13}C CP/MAS NMR spectra of (a) the peptide **2** after acetonitrile treatment, (b) the peptide **3** after acetonitrile treatment and (c) the dragline silk of *Nephila clavipes*.[4] The spectral range, 5–65 ppm, was expanded.

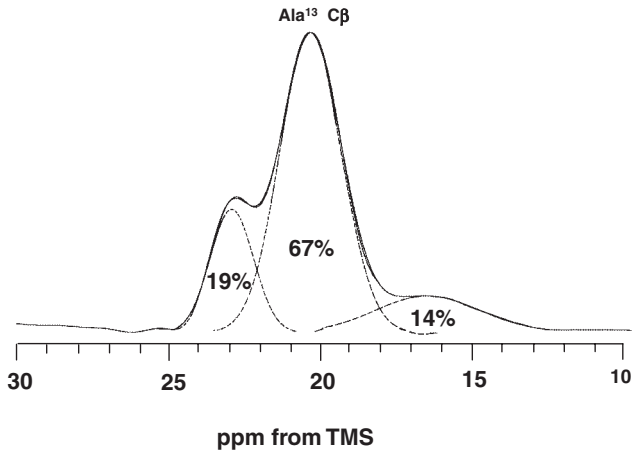

Figure 5. Deconvolution of the Ala^{13} Cβ peak of the peptide **3** after acetonitrile treatment.

reported previously[20,21] although the fraction is different from the case of *B. mori* silk. Thus, the heterogeneous structure can be proposed for the poly(Ala) region with β-sheet structure as well as *B. mori* silk fibroin.

The Cα chemical shift of Ala residue at the 18th position in the GGA sequence followed to the former poly(Ala) region is 49.0 ppm (Figure 4b). However, the chemical shift value corresponds to both β-sheet and 3_1 helix conformations[22] and therefore we cannot distinguish the conformation although it is different from either α-helix (52.5 ppm) or random coil (50.0 ppm) (Table I). The C=O chemical shift at the Gly 27th position in the sequence, LGSQ, is 171.5 ppm (the spectrum is not shown here), which is close to the value of both 3_1 helix and random coil, 171.3 ppm. The Cα chemical shift, 42.4 ppm for 37th Gly residue adjacent to the latter poly(Ala) region is different from the value of either 3_1 helix or α-helix (Figure 4b and Table I). Although it is difficult to distinguish between β-sheet and random coil because the chemical shift is similar for both cases, the value is closer to the β-sheet (Table I). The Ala Cα and Cβ, and Gly Cα chemical shifts of the peptide **2** after AN treatment (Figure 4a), are the same as those of the dragline silk fiber of *Nephila clavipes* reported previously (Figure 4c).[4]

Acknowledgment. TA acknowledges supports from the Insect Technology Project, Japan and Agriculture Biotechnology Project, Japan.

REFERENCES

1. M. B. Hinman and R. V. Lewis, *J. Biol. Chem.*, **267**, 19320 (1992).
2. F. Vollrath and D. P. Knight, *Nature*, **410**, 541 (2001).
3. A. H. Simmons, C. A. Michal, and L. W. Jelinski, *Science*, **271**, 84 (1996).
4. A. H. Simmons, E. Ray, and L. W. Jelinski, *Macromolecules*, **27**, 5235 (1994).
5. J. Kümmerlen, J. D. van Beek, F. Vollrath, and B. H. Meier, *Macromolecules*, **29**, 2920 (1996).
6. C. A. Michal and L. W. Jelinski, *J. Biomol. NMR*, **12**, 231 (1998).
7. A. Seidel, O. Liivak, S. Calve, J. Adaska, G. Ji, Z. Yang, D. Grubb, D. Zax, and L. W. Jelinski, *Macromolecules*, **33**, 775 (2000).
8. Z. Yang, O. Liivak, A. Seidel, G. Laverda, D. Zax, and L. W. Jelinski, *J. Am. Chem. Soc.*, **122**, 9019 (2000).
9. J. D. van Beek, S. Hess, F. Vollrath, and B. H. Meier, *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 10266 (2002).
10. P. T. Eles and C. A. Michal, *Biomacromolecules*, **5**, 661 (2004).
11. G. P. Holland, R. V. Lewis, and J. L. Yarger, *J. Am. Chem. Soc.*, **126**, 5867 (2004).
12. P. T. Eles and C. A. Michal, *Macromolecules*, **37**, 1342 (2004).
13. P. T. Eles and C. A. Michal, *Biomacromolecules*, **5**, 661 (2004).
14. M. Hronska, J. D. van Beek, P. T. F. Williamson, F. Vollrath, and B. H. Meier, *Biomacromolecules*, **5**, 834 (2004).
15. C. Zhao and T. Asakura, *Prog. Nucl. Magn. Reson. Spectrosc.*, **39**, 301 (2001).

16. J. Yao, Y. Nakazawa, and T. Asakura, *Biomacromolecules*, **5**, 680 (2004).
17. T. Asakura and Y. Nakazawa, *Macromol. Biosci.*, **4**, 175 (2004).
18. Y. Nakazawa, M. Bamba, S. Nishio, and T. Asakura, *Protein Sci.*, **12**, 666 (2003).
19. Y. Nakazawa and T. Asakura, *J. Am. Chem. Soc.*, **125**, 7230 (2003).
20. T. Asakura, J. Yao, T. Yamane, K. Umemura, and A. S. Ulrich, *J. Am. Chem. Soc.*, **124**, 8794 (2002).
21. T. Asakura and J. Yao, *Protein Sci.*, **11**, 2706 (2002).
22. J. Ashida, K. Ohgo, K. Komatsu, A. Kubota, and T. Asakura, *J. Biomol. NMR*, **25**, 91 (2003).