^{13}C Solid-State NMR Study of Structural Heterogeneity in Peptides Containing Both Polyalanine and Repeated GGA Sequences as a Local Structural Model of *Nephila clavipes* Dragline Silk (Spidroin 1)

Tetsuo Asakura,* Mingying Yang, Taiji Kawase, and Yasumoto Nakazawa

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

Received November 14, 2004; Revised Manuscript Received January 5, 2005

ABSTRACT: To avoid potential ambiguity in the structural determination of native spider silk, we prepared both a nonlabeled peptide with a sequence containing both the polyalanine and the repeated GGA regions, QGAG(A)$_6$GGAGA(GGA)$_3$GAGRGGLGG (**I**), and the ^{13}C-labeled peptides QGAGAAA[1-^{13}C]A^{8}AAGG[2-^{13}C]A^{13}GAGGAG[2-^{13}C]G^{20}[3-^{13}C]A^{21}GGAGAGRGGLGG (**Ia**) and QGAGAAAAAAGGAGAG-GAG[1-^{13}C]G^{20}[1-^{13}C]A^{21}GGAGAGRGGLGG (**II**) as a local structural model of spidroin 1 (MaSpl) protein in spider dragline silk of *Nephila clavipes.* Solvent treatments prior to the NMR measurements induce a structural change of these model peptides and provide a model to reproduce the structure of the silk fiber. Conformation-dependent ^{13}C NMR chemical shifts were mainly used to determine the local structure, including the evaluation of the fraction of several conformations. The characteristic structure, 65% β-sheet for the Ala^8 residue in the poly-Ala region, and 70% 3_1-helix for the Ala^{21} residue and mainly 3_1-helix for the Gly^{20} residue in the GG^{20}A^{21} sequence was observed after peptides **Ia** and **II** were dissolved in 9 M LiBr followed by dialysis against water. The 2D spin diffusion ^{13}C solid-state NMR spectrum of the Ala^{21} residue of peptide **II** after this treatment was also reproduced by 70% 3_1-helix (ϕ, φ = −90°, 120°) and 30% β-sheet (ϕ, φ = −150°, 150°) structure. However, the Ala C$_\beta$ peak assigned to the 3_1-helix in the spectrum of **Ia** is broad, implying that the torsion angles of the Ala^{21} residue are distributed, but with an average that corresponds approximately to the torsion angles of the 3_1-helix. An increase in the fraction of β-sheet in both the poly-Ala and GG^{20}A^{21} regions was observed for **Ia** after it was dissolved in formic acid and then dried in air. Moreover, after **Ia** was dissolved in formic acid and then precipitated in methanol, the spectrum showed a tightly packed β-sheet structure with a further increase in the fraction of β-sheet although 15% 3_1-helix still remained in the GG^{20}A^{21} region. The β-sheet structure of the poly-Ala region and both 3_1-helix and β-sheet structures in the repeated GGA sequence are in agreement with the structural model for the native spider dragline silk fiber from *N. clavipes* from a previous NMR study (van Beek, J. D.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10266−10271). On the other hand, the α-helical conformation was found to be dominant for the peptide treated with trifluoroacetic acid together with a significant contribution from other structures. The fraction of the other structures was 20−40% depending on the position of the ^{13}C-labeled Ala residue in the chain.

Introduction

The dragline filaments produced by orb weaving spiders have been the focus of numerous recent studies because they are the toughest protein fibers known, and the selected examples have stress−strain behaviors attractive for use in strength and toughness critical applications.[1−3] The dragline silk of the golden orb web spider *Nephila clavipes* contains two structural proteins, designated spidroin 1 (MaSp1) and spidroin 2 (MaSp2).[4,5] The dominant MaSpl protein can be described as a block copolymer consisting of poly-Ala and Gly-rich regions.

Several kinds of solid-state NMR[6−16] and X-ray diffraction[17,18] methods have been applied to clarify the detailed structure and dynamics of native spider silk fibers. It has been shown that silk fibroins undergo a substantial structural change from gland silk to native dragline silk. The liquid silk stored in the spider silk gland has been reported in different studies to be in a dynamic loose helical structure[19] or a random coil conformation[20] on the basis of the conformation-dependent Ala C$_\beta$ chemical shift in the ^{13}C solution NMR spectrum of native liquid silk. The Ala C$_\beta$ chemical shifts in both these studies are essentially the same. In addition, the same chemical shift has also been observed for the liquid silk stored in the silk gland of the *Bombyx mori* silkworm and regenerated silk fibroin in aqueous solution.[21−24]

On the other hand, the poly-Ala regions in the dragline silk fiber were found to be predominantly in β-sheet conformation as characterized by the conformation-dependent ^{13}C NMR chemical shifts of Ala residues[6−11] as well as the DOQSY (double-quantum single-quantum correlation experiment) and 2D spin diffusion NMR spectra of [1-^{13}C]Ala spider dragline silk.[8,13] The longer poly-Ala domains in *Samia cynthia ricini* silk fibers also adopt a predominantly β-sheet conformation.[25−29] However, there is more disagreement about the local structure of the Gly-rich region. It has been suggested that the Gly-rich region in the final silk is either a random coil[30] or amorphous.[17] These suggestions were made to account respectively for the rubberlike character or low crystallinity (10−15%) of spider silk determined by X-ray scattering. Theil et al.[31] also speculated that the Gly-rich regions may be incorporated into β-sheet crystals in the spider dragline silk fiber, though his evidence for the existence of large crystallites from TEM studies is open to question. However, van Beek et al. proposed recently that the Gly-rich region of the native spider silk fiber is partly

* To whom correspondence should be addressed. Phone and Fax: +81-42-383-7733. E-mail: asakura@cc.tuat.ac.jp.

10.1021/ma047660z CCC: $30.25
Published on Web 03/16/2005

incorporated into β-sheets and otherwise forms helical structures with approximately 3-fold symmetry from the DOQSY and DECODER (direction exchange with correlation for orientation-distribution evaluation and reconstruction) observations of a [1-^{13}C]Gly spider dragline silk fiber.[13] They had previously provided evidence for the presence of a 3_1-helical structure in the Gly-rich region using the 2D spin diffusion solid-state NMR spectrum of [1-^{13}C]Gly-labeled spider dragline silk.[8] The presence of turnlike structures for the LGXQ (X = S, G, or N) motif was also proposed on the basis of REDOR (rotational-echo double-resonance) measurements of isotopically labeled dragline silk fibers.[9] The fraction of fast-moving and immobilized Gly residues was determined by Yang et al. for a wet [1-^{13}C]Gly-labeled spider dragline silk.[12] These NMR studies give direct evidence about the local structure and dynamics of the Gly-rich region. However, the difficulties in the structural determination of the Gly-rich region are due to the heterogeneity in the repeated sequences of dragline silk. Therefore, model peptides with defined primary structures selected from the native spider silk sequences should be very valuable for structural determinations because they avoid the large variations in structural distributions resulting from the heterogeneity in the primary structure.

In our previous papers,[32−39] ^{13}C-labeled model peptides for *B. mori* and *S. cynthia ricini* silk fibroins coupled with several solid-state NMR techniques have proved to be valuable for the precise determination of local structure in silk fibroins. The structural characterization of several repeated sequences has been performed by combined evaluation of the conformation-dependent ^{13}C CP/MAS NMR chemical shift and 2D spin diffusion ^{13}C solid-state NMR.[37,38] Torsion angle determination has also been reported for a peptide, $(AGG)_{10}$ selected as one of the sequence models of spider silk.[40]

In this paper, we describe the use of solid-state NMR methods to analyze pertinent model peptides. We obtained further detailed structural information for MaSp1 by focusing on the Gly-rich region. Previous studies[37,38] indicated that a single repeat of poly-Ala and Gly-rich blocks from the primary structure of the *S. cynthia ricini* silk fibroin will fold into a secondary structure, closely similar to that of the whole native fibroin. Therefore, it was expected that a single repeat can also imitate the structure of MaSp1 in spider dragline silk, although a possible conformational heterogeneity depending upon the natural variation of the repetitive blocks should be taken into account. Generally, the Gly-rich region is constructed from the GGX repeat, with X being L, Y, and Q. Recent studies show that the peptides $(AGG)_{10}$ and $(LGG)_{10}$ as models for the Gly-rich region in spider silk both adopt the same 3_1-helix structure.[40] The local structural information can also be obtained from the line shape analysis of the Ala peaks, mainly from Ala C_β peaks, as reported previously.[33−35] Therefore, we placed Ala in the X position in the Gly-rich blocks to obtain more detailed structural information relatively easily. Accordingly, the following peptides were synthesized: QGAGAAAAAAGGAGA$(GGA)_3$GAGRGGLGG (**I**), QGAGAAA[1-^{13}C]A^{8}AAGG[2-^{13}C]A^{13}GAGGAG[2-^{13}C]G^{20}[3-^{13}C]A^{21}GGAGAGRGGLGG (**Ia**), QGAGAAAAAAGGAGAGGAG[1-^{13}C]G^{20}[1-^{13}C]A^{21}GGAGAGRGGLGG (**II**).

Solvent treatments prior to the NMR measurements induce a structural change of these model peptides and provide a model for the conformation changes in the formation of the silk fiber. The model peptide containing poly-Ala crystalline regions of the *S. cynthia ricini* silk fibroin, GGAGGGYGGDGG$(A)_{12}$GGAGDGYGAG, was shown to take an α-helical conformation after trifluoroacetic acid treatment corresponding to the prespun state.[37,38] In contrast, after being dissolved in formic acid (FA) and then dried in air, it adopted a mainly β-sheet structure, corresponding to that of the final silk fiber. Further, the model peptide $(AG)_{15}$ after FA treatment also adopts a β-sheet structure corresponding to the silk II form of the *B. mori* silk thread.[24,25] Although $(AG)_{15}$ takes exclusively a silk I like structure after dissolution in 9 M LiBr, followed by dialysis against water, the poly-Ala model peptide retains a β-sheet structure after 9 M LiBr/dialysis treatment. We discuss the extent to which the conformational changes during the formation of spider silk can be mimicked by treating our model spider peptide with different reagents.

Materials and Methods

Model Peptides. The model peptides **I**, **Ia**, and **II** were synthesized by the F-moc solid-phase method as described previously[37,38] (Pioneer peptide synthesizer, PE Biosystems Co. Ltd.). After synthesis, the peptides were (1) dissolved in 9 M LiBr followed by dialysis against water and lyophilization (9 M LiBr/dialysis treatment), (2) dissolved in trifluoroacetic acid (TFA) followed by precipitation in diethyl ether and drying in air (TFA/D treatment), or (3) dissolved in formic acid followed by precipitation by methanol and drying in air (FA/M treatment). Model peptide **Ia** was also dissolved in formic acid and then dried in air (FA treatment) without methanol treatment.

^{13}C CP/MAS NMR Observation. Solid-state ^{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 10000−25000 scans were collected over a spectral width of 60 kHz, with a recycle delay of 3 s. Chemical shifts were indirectly calibrated using the adamantane methylene peak observed at 28.8 ppm relative to the TMS peak.

2D Spin Diffusion NMR Using Off Magic Angle Spinning. The 2D spin diffusion NMR spectrum was obtained at room temperature with a Varian Unity INOVA 400 NMR spectrometer equipped with a 7 mm ∅ Jakobsen-type double-tuned MAS probe using off magic angle spinning ($\theta_m - 7°$) with sample spinning at 6 kHz. The scaling factor of the 2D spin diffusion spectra is then $(1/2)(3\cos^2(\theta_m - 7°) - 1) = 0.198$. Other NMR experimental conditions and the method of the simulation are as described previously.[32]

Results

Conformation of Peptides I, Ia, and II after 9 M LiBr/Dialysis Treatment. Figure 1 shows the ^{13}C CP/MAS NMR spectra of the model peptides **I**, **Ia**, and **II** after 9 M LiBr/dialysis treatment. Peptide **I** gave an intense Ala C_β ^{13}C chemical shift at 17.2 ppm which can be ascribed to the 3_1-helix, close to the value of 17.4 ppm of a reference system, as summarized in Table 1. Site-specific assignment of peaks was possible by comparing ^{13}C-labeled preparations with the unlabeled peptide. The ^{13}C chemical shifts of [1-^{13}C]A^8-, [2-^{13}C]A^{13}-, [3-^{13}C]-A^{21}-, and [2-^{13}C]G^{20}-labeled **I** (**Ia**) and [1-^{13}C]G^{20}- and [1-^{13}C]A^{21}-labeled peptide (**II**) are shown in parts B and C of Figure 1, respectively. The latter peptide was also used for the 2D spin diffusion experiment for determination of the torsion angle at the Ala^{21} residue in the sequence $(GGA)_3$.

As shown in Figure 1B, the Ala^8 residue in the poly-Ala region takes mainly the β-sheet conformation as judged from the chemical shift of 171.8 ppm for the

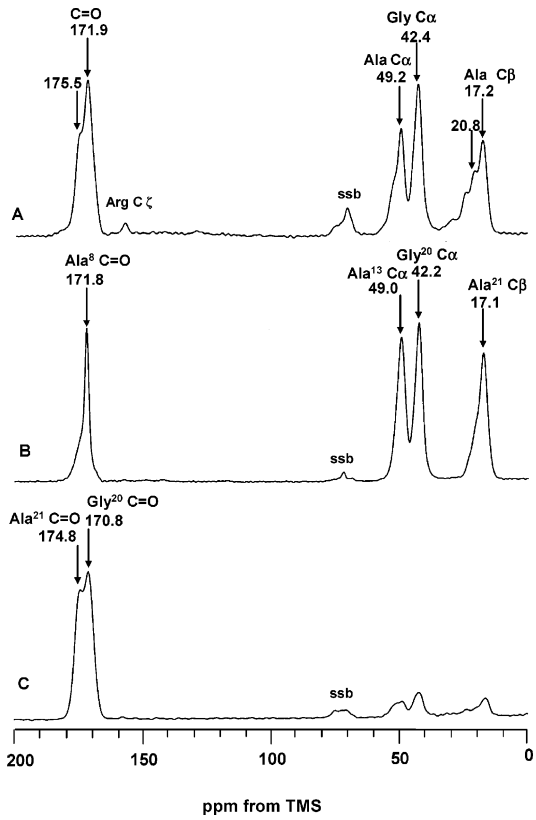

Figure 1. ^{13}C CP/MAS NMR spectra of **I** (A), **Ia** (B), and **II** (C) after they were dissolved in 9 M LiBr and then dialyzed against water (ssb means spinning sideband).

[1-^{13}C]Ala^{8} peak. However, there remains a shoulder at the lower field of the main peak, indicating the presence of another structure. The Ala^{13} C_α chemical shift of GGA adjacent to the poly-Ala region in peptide **Ia** resonated at 49.0 ppm. This chemical shift value is ascribed to either a 3_1-helix or a β-sheet. Further distinction between these conformations is not possible from the Ala^{13} C_α chemical shift; they need to be distinguished by means of the ^{13}C chemical shift of Ala^{13} Cβ or carbonyl carbons. The local conformation of **Ia** at the Gly^{20} residue in the central GGA sequence of $(GGA)_3$ can be assigned to a 3_1-helix instead of a β-sheet using the C_α and carbonyl chemical shifts at 42.2 and 170.8 ppm for **Ia** and **II**, respectively. The Ala^{21} C_β ^{13}C NMR peak of **Ia** resonated at 17.1 ppm, indicating that the Ala^{21} residue of this peptide adopts a 3_1-helix, consistent with the data of the Ala^{21} carbonyl chemical shift at 174.8 ppm. However, it is noteworthy that the expanded Ala C_β peak of **Ia** (Figure 2A) is rather broadened as compared with that of the Ala C_β carbon in $(AGG)_{10}$, which also adopts a 3_1-helix conformation as reported previously (Figure 2B).[40] This broadening in the former peptide could in theory result from a shoulder at 20.8 ppm characteristic of the β-sheet, the deconvolution of which indicates that the β-sheet accounts for 30% of the total conformations. However, after deconvolution, the main peak of Ala^{21} C_β at 17.1 ppm still remained broad, and the peak position was shifted to high field by 0.3 ppm as compared with that of the Ala C_β carbon in $(AGG)_{10}$ (Figure 2B). This implies that the torsion angles of the Ala^{21} residue are distributed, but with an average that corresponds approximately to the torsion angles of the 3_1-helix. Thus, the 3_1-helix here appears to be slightly distorted.

Table 1. ^{13}C CP/MAS NMR Chemical Shifts (ppm from the TMS Peak) of the Peptides QGAGAAAAAAGGAGAGGAGGAGGAGAGAGRGGLGG (I), QGAGAAA[1-^{13}C]AAAGG[2-^{13}C]AGAGGAG[2-^{13}C]G[3-^{13}C]AGGAGAGAGRGGLGG (Ia), and QGAGAAAAAAGGAGAGGAG[1-^{13}C]G[1-^{13}C]AGGAGAGAGRGGLGG (II) Treated in Four Ways[a]

	^{13}C-labeled peptides **Ia** and **II**												
	9 M LiBr/dialysis		FA		FA/methanol		TFA/diethyl ether		nonlabeled peptide				
carbon	observation	deconvolution	observation	deconvolution	observation	deconvolution	observation	deconvolution	9 M LiBr/ dialysis	TFA/ diethyl ether	FA/ methanol	α-helix[41]	3_1-helix[42]	β-sheet[34]
Ala Cβ	17.1 (A^{21})	17.1 [70%]	17.2 (A^{21})	17.2 [45 %]	17.2 (A^{21})	17.2 [15%]	16.1 (A^{21})		17.2	15.9	21.0	15.7	17.4	(16.7)[b]
		20.8 [30%]	20.8	20.8 [40%]	20.8	20.8 [45%]			20.8					19.6
		22.8 [15%]		22.8 [40%]										22.2
Cα	49.0 (A^{13})[c]	52.3 [60%]	49.0 (A^{13})		49.1 (A^{13})	52.0 (A^{13})		52.3 [60%]	49.2	52.1	49.1	52.5	48.9	48.7
		48.8 [40%]						48.8 [40%]						
C=O	171.8 (A^{8})	171.8 [65%]	172.0 (A^{8})		172.1 (A^{8})		176.5 (A^{8})	176.5 [65%]	171.9	172.0	172.4	176.5	174.6	171.8
		175.0 [35%]						172.0	175.5	176.0				
	174.8 (A^{21})													
Gly Cα	42.2 (G^{20})		42.3 (G^{20})		42.4 (G^{20})		43.0 (G^{20})	44.2 [70%]	42.4	42.9	42.8	44.0	41.6	42.4
								42.2 [30%]						
C=O	170.8 (G^{20})								171.9	172.0	168.8	172.3	171.3	169.1

[a] The ^{13}C chemical shifts of Ala and Gly residues with the typical conformation are also listed. [b] The peak at 16.7 ppm was assigned to a distorted β-turn. [c] The chemical shift of 49.0 ppm indicates the absence of the α-helix.

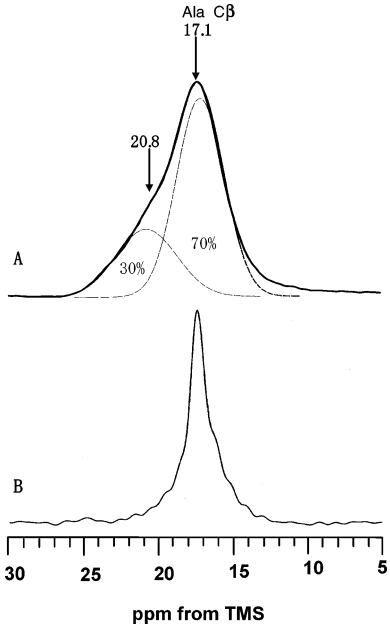

Figure 2. (A) Expanded C_β peaks of the Ala^{21} residue in peptide **II** after 9 M LiBr/dialysis treatment. The relative proportion of the 3_1-helix and β-sheet was determined as 70% and 30%, respectively. (B) Peak for Ala residues in $(Gly\text{-}GlyAla)_{10}$ adopting the 3_1-helix form reported previously.[30]

Torsion Angles for the Ala^{21} Residue after 9 M LiBr/Dialysis Treatment. The 2D spin diffusion ^{13}C solid-state spectrum of peptide **II** after 9 M LiBr/dialysis treatment was used to determine the local torsion angles at the Ala^{21} residue as shown in Figure 3A. For comparison, the ^{13}C spin diffusion spectrum of the Ala residue in $(AGG)_4AG[1\text{-}^{13}C]G[1\text{-}^{13}C]AGG(AGG)_4$ with the 3_1-helix form reported previously[29] is shown in Figure 3B. The observed spectral pattern at the Ala^{21} residue is slightly different from the latter pattern of the Ala residue in $(AGG)_{10}$. Figure 4 shows the 2D spin diffusion patterns calculated with observed chemical shift tensors as a function of the torsion angles ϕ and φ of the Ala residue.[21] The spectral patterns change significantly on changing the torsion angles. As pointed out above, 70% of the Ala^{21} residues take the 3_1-helix conformation, while the remaining 30% take the β-sheet structure (Figure 2). Therefore, the spectral pattern was calculated by adding 30% of the calculated typical spectral pattern of the β-sheet (with the torsion angles $\phi, \varphi = -150°, 150°$ [34]) to 70% of the calculated 3_1-helix pattern (with the torsion angles $\phi, \varphi = -90°, 150°$ [40]). The calculated pattern (Figure 5B) is slightly different from the observed one (Figure 5A). When the slightly deviated torsion angles $\phi, \varphi = -90°, 120°$ were used for the 70% 3_1-helix pattern, a better agreement between the observed and calculated patterns was obtained (Figure 5C).

Conformation of Peptide Ia Treated by 9 M LiBr/Dialysis, TFA/Diethyl Ether, FA, or FA/methanol. Figure 6 illustrates the ^{13}C CP/MAS NMR spectra of the $[1\text{-}^{13}C]A^8$-, $[2\text{-}^{13}C]A^{13}$-, $[3\text{-}^{13}C]A^{21}$-, and $[2\text{-}^{13}C]G^{20}$-labeled version of **I** (peptide **Ia**) treated in four ways: 9 M LiBr/dialysis, FA, FA/methanol (M), and TFA/diethyl ether (D). Treatment with formic acid produced a marked structural change in the peptide, an increase in the fraction of β-sheet structure compared with that from 9 M LiBr/dialysis treatment (Figures 1A and 6A). In this paper, two kinds of treatments with formic acid were performed (see the Materials and Methods). Figure 6B shows an increase in the fraction of β-sheet structure produced by FA treatment as determined by deconvolution of the Ala^{21} C_β peaks. Slightly narrower peaks were also observed for Ala^8 C=O, Ala^{13} C_α, and Gly^{20} C_α carbons compared with the line widths of these peaks of the peptide after 9 M LiBr/dialysis treatment. FA/M treatment compared with FA treatment produced a further increase of β-sheet structure as shown in Figure 6C. The sharp peak at 172.1 ppm for Ala^8 C=O in Figure 6D shows that the poly-Ala region in peptide **Ia** takes exclusively the β-sheet structure after FA/M treatment. The intense Ala^{21} C_β peak at around 20.8 ppm is ascribed to the predominant β-sheet, while a minor component at 17.2 ppm is ascribed to the 3_1-helix. Further, the sharp Gly^{20} C_α peak at 42.4 ppm and the Ala C^{13} C_α peak at 49.1 ppm are also consistent with the β-sheet as the major component. The fraction of the conformation determined by deconvolution from ^{13}C-labeled signals was determined as summarized in Table 2.

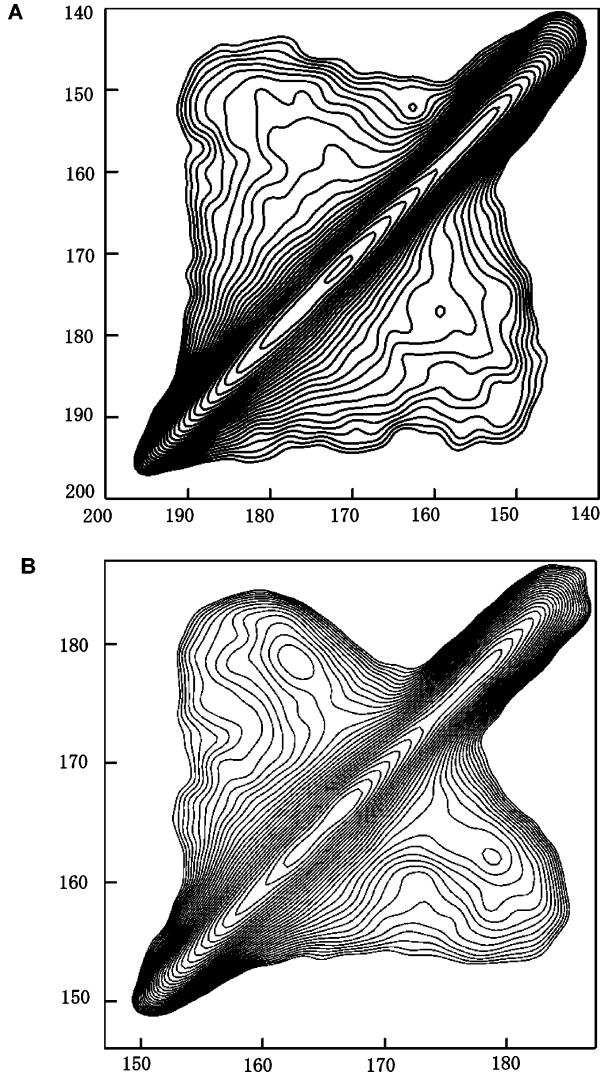

Figure 3. (A) 2D spin diffusion ^{13}C solid-state NMR spectra (carbonyl carbon region) of peptide **II** after 9 M LiBr/dialysis treatment. (B) As for (A) but $(AGG)_4AG[1\text{-}^{13}C]G[1\text{-}^{13}C]AGG(AGG)_4$ in the 3_1-helix form published previously.[30] The latter spectrum was obtained using the off magic angle condition $\theta_m + 9°$, which is different from the condition $\theta_m - 7°$ in (A). Thus, the direction of the scale is reversed.

The most intense Ala^8 C=O peak in the poly-Ala region of the TFA-treated peptide **Ia** resonated at 176.5

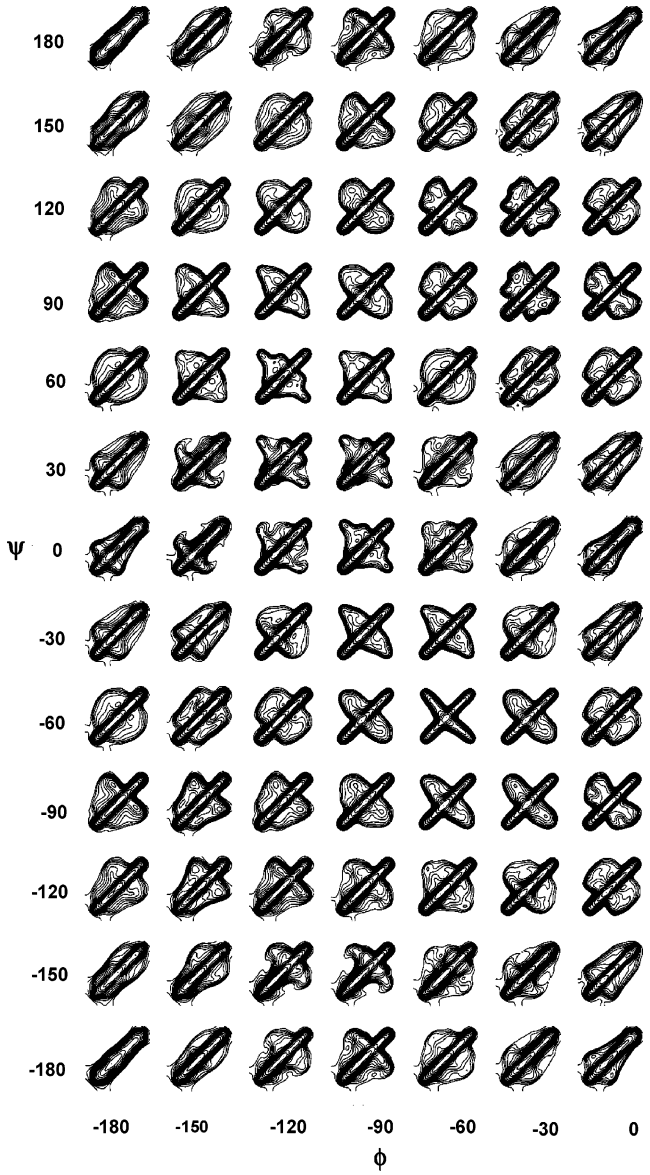

Figure 4. 2D ^{13}C spin diffusion NMR patterns calculated as a function of the torsion angles ϕ and φ of the Ala residue ($-180° < \phi < 0°$, $-180° < \varphi < +180°$).

ppm. This indicates that the Ala[8] residue takes mainly an α-helical conformation after TFA/D treatment (Figure 6D and Table 1). However, an upfield shoulder of the Ala[8] peak was also observed at 172.0 ppm, which was assigned to the β-sheet structure. The latter fraction accounted for only 35%. Thus, the poly-Ala region in this peptide takes mainly an α-helical conformation, but there are also significant amounts of β-sheet structure. It appears that the Ala[13] residue in the GGA sequence immediately before the poly-Ala region adopts mainly the α-helical conformation, judging from the Ala[13] C_α chemical shift of 52.0 ppm. Thus, the GGA sequence followed by the poly-Ala region is incorporated into the α-helix of the poly-Ala region. There is an obvious upfield shoulder at the Ala[13] C_α peak, and the fraction of this was 40%, which is a peak pattern similar to that of the upfield shoulder at the Ala[8] C═O peak in the poly-Ala region. The Gly[20] C_α peak at 43.0 ppm was also broad and could be deconvoluted to 70% α-helix (44.2 ppm) and 30% another conformation (42.2 ppm). The chemical shift of 16.1 ppm for the Ala[21] C_β peak is close to the α-helix chemical shift of the Ala C_β carbon, 15.7

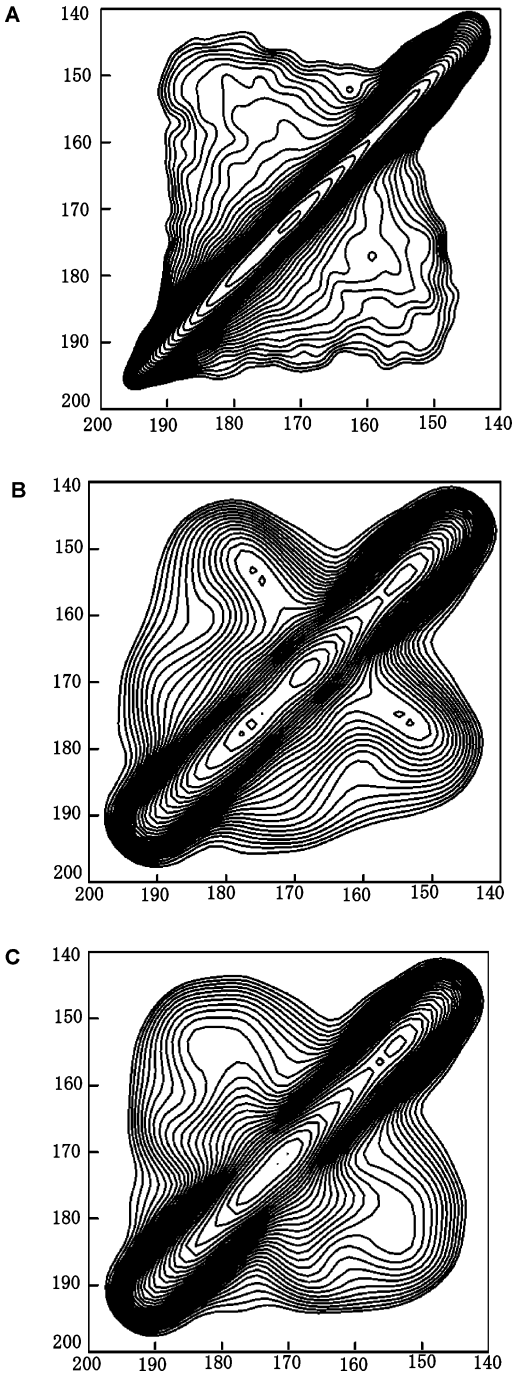

Figure 5. (A) 2D spin diffusion ^{13}C solid-state NMR spectrum of peptide **II** after 9 M LiBr/dialysis treatment. 2D spin diffusion ^{13}C solid-state NMR spectra calculated by assuming (B) 30% β-sheet (torsion angles ϕ, $\varphi = -150°$, 150°) and 70% 3_1-helix (torsion angles ϕ, $\varphi = -90°$, 150°) and (C) 30% β-sheet (torsion angles ϕ, $\varphi = -150°$, 150°) and 70% 3_1-helix (torsion angles ϕ, $\varphi = -90°$, 120°).

ppm, and therefore, the Ala[21] residue takes mainly the α-helix conformation. Thus, peptide **Ia** takes mainly an α-helix conformation in every residue (Table 2). In general, the length of the poly-Ala region, $(Ala)_6$, is too short to form an α-helix.[43] However, the presence of other sequences such as GGA after $(Ala)_6$ and AG prior to $(Ala)_6$ in the chain in the poly-Ala region may allow $(Ala)_6$ to form an α-helix although there are significant amounts of other conformations, mainly β-sheet structure at about 30−40% (Table 2).

Conformation of Peptide I after Treatment with TFA/Diethyl Ether or FA/Methanol. So far we have

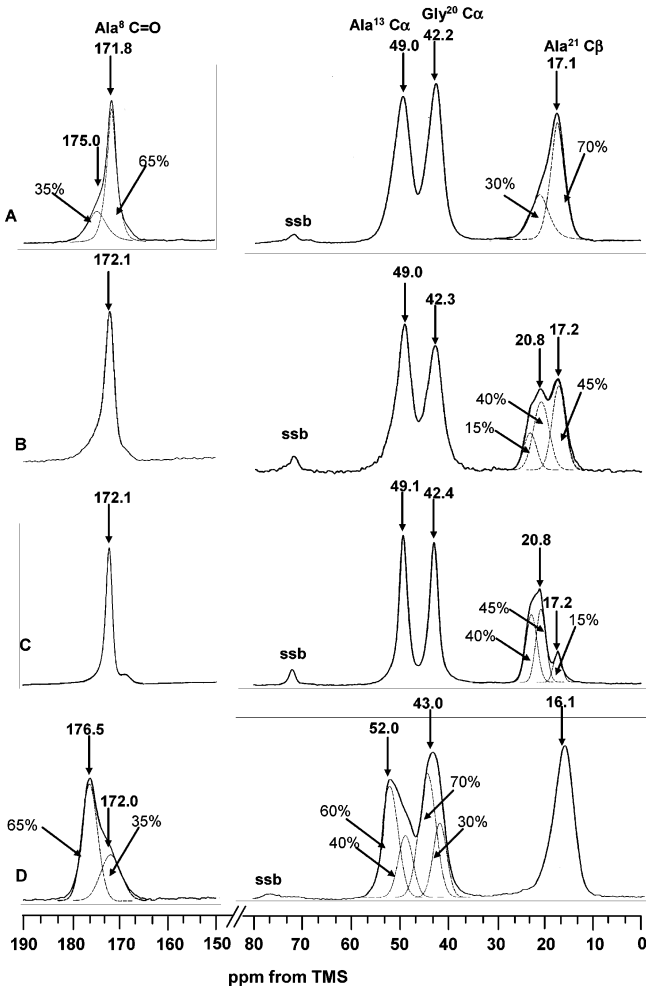

Figure 6. ^{13}C CP/MAS NMR spectra of **Ia** after treatment by (A) 9 M LiBr/dialysis, (B) FA, (C) FA/methanol, and (D) TFA/diethyl ether. The peak deconvolution was performed for each ^{13}C-labeled carbon, and the results are listed in Table 2. Ssb means spinning sideband.

discussed conformational features of ^{13}C-labeled **I** treated by TFA/D, 9 M LiBr/dialysis, FA, and FA/M. It is now feasible to interpret the ^{13}C NMR spectra of the unlabeled peptide **I** in terms of the respective conformations involved after taking into account a plausible contribution of peaks from the unlabeled Ala and Gly residues (residues 13 and 17, respectively) and from Glu, Arg, and Leu carbonyl signals. Only the spectra after treatment with TFA/D or FA/M are shown here to illustrate the clear difference of the secondary structures between the α-helix and β-sheet.

The sharp peak at 24.0 ppm of **I** treated by TFA/D can be assigned to both Leu Cγ and Arg Cγ carbons. The presence of the α-helix conformation as a dominant form is obvious as judged from the chemical shifts of Ala C$_\alpha$, Ala C$_\beta$, and Gly C$_\alpha$ carbons (Figure 7A). After FA/M treatment, the Ala C$_\beta$ peak is displaced to 21.0 ppm, while the sharp C$_\alpha$ peak is displaced to 49.1 ppm, indicating the formation of a tightly packed β-sheet structure (Figure 7B).[23] The carbonyl peak at 172.0 ppm in Figure 7A can be assigned to mainly a mixture of the Gly peak and β-sheet peak of Ala residues, and the 176.0 ppm peak can be assigned to the α-helix peak of Ala residues. After FA/M treatment (Figure 7B), the carbonyl peak for the Gly residue of 168.8 ppm is clearly observed, indicating that the Gly residues are incorporated into β-sheet structure.

Discussion

Poly-Ala regions in the dragline silk fiber have been found to be predominantly in β-sheet conformation as characterized by the conformation-dependent ^{13}C NMR chemical shifts of Ala residues[6−11] as well as the DOQSY and 2D spin diffusion NMR spectra of [1-^{13}C]Ala spider dragline silk.[8,13] On the other hand, the structure of the Gly-rich region has been shown to be important for the rubberlike properties[2,8,30] and supercontraction[10,12,15,16] of the spider dragline silk fiber, but there are arguments about the structure of this region. NMR studies give direct evidence about the local structure of the Gly-rich region. More recently, van Beek et al. suggested that the Gly-rich region of the native spider silk fiber is partly incorporated into β-sheets and otherwise forms 3$_1$-helical structure as indicated by 2D spin diffusion solid-state NMR, DOQSY, and DECODER.[8,13] However, the difficulties in the structural determination of the Gly-rich region may be due to the heterogeneity in the repeated sequences of dragline silk. The use of appropriate model peptides with defined primary structures selected from the native spider silk sequences should be very valuable for structural determinations because the large variations in structural distributions resulting from the heterogeneity in the primary structure can be avoided. Thus, we prepared both a nonlabeled peptide with a sequence containing both the polyalanine and the repeated GGA regions, QGAG-(A)$_6$GGAGA(GGA)$_3$GAGRGGLGG (**I**), and the ^{13}C-labeled peptides QGAGAAA[1-^{13}C]A^{8}AAGG[2-^{13}C]A^{13}G-AGGAG[2-^{13}C]G^{20}[3-^{13}C]A^{21}GGAGAGRGGLGG (**Ia**) and QGAGAAAAAAGGAGAGGAG[1-^{13}C]G^{20}[1-^{13}C]A^{21}GG-AGAGGGLGG (**II**) as a local structural model of MaSpl protein in spider dragline silk of *N. clavipes*. Solvent treatments prior of these model peptides provide a structural model for the conformation changes involved in the formation of the silk fiber.

As summarized in Table 2, the characteristic structure, 65% β-sheet for the Ala^8 residue in the poly-Ala region, and 70% 3$_1$-helix for the Ala^{21} residue and mainly 3$_1$-helix for the Gly^{20} residue in the GG^{20}A^{21} sequence, was observed after peptides **Ia** and **II** were dissolved in 9 M LiBr followed by dialysis against water. Thus, we observed a large difference in the local conformation of Ala residues depending on their position in the peptide. Even in the case of a given Ala residue, the conformation was heterogeneous. In addition, from the detailed analysis of the Ala^{21} C$_\beta$ peak, two kinds of β-sheet structures were observed. This reflects the difference in the intermolecular arrangement of the Ala^{21} residues with β-sheet structure as reported for the *B. mori* silk fibroin fiber previously.[34,35] Thus, useful structural information was obtained by combination of the use of selectively ^{13}C labeled model peptides and conformation-dependent ^{13}C chemical shifts. The 2D spin diffusion ^{13}C solid-state NMR spectrum of the Ala^{21} residue of peptide **II** after this treatment was also reproduced by a combination of 70% 3$_1$-helix (ϕ, φ = −90°, 120°) and 30% β-sheet (ϕ, φ = −150°, 150°) structure. Thus, the presence of the 3$_1$-helix in the sequence (GGA)$_n$ was clearly shown in both the 2D spin diffusion NMR study and the chemical shift analysis of model peptides with defined primary structures selected from the native spider silk sequences. However, the Ala C$_\beta$ peak assigned to the 3$_1$-helix in the spectrum of **Ia** is broad, implying that the torsion angles of the Ala^{21} residue are distributed, but with an average that

Table 2. Fraction of Each Conformation for ^{13}C-Labeled Carbon in QGAGAAA[1-^{13}C]AAAGG[2-^{13}C]AGAGGAG[2-^{13}C]G[3-^{13}C]AGGAGAGRGGLGG (Ia) or QGAGAAAAAAGGAGAGGAG[1-^{13}C]G[1-^{13}C]AGGAGAGRGGLGG (II) Treated in Four Ways

	Ala^{8} C=O		Ala^{13} Cα		Gly^{20} (Cα and C=O)		Ala^{21} (Cβ and C=O)	
treatment	conformation	fraction (%)	conformation	fraction (%)	conformation	fraction (%)	conformation	fraction (%)
9 M LiBr/dialysis	others	35	*a*		3_1-helix[b]		3_1-helix	70
	β-sheet	65					β-sheet	30
FA	β-sheet		β-sheet		β-sheet		β-sheet	55
							3_1-helix	45
FA/methanol	β-sheet		β-sheet		β-sheet		β-sheet	85
							3_1-helix	15
TFA/diethyl ether	α-helix	65	α-helix	60	α-helix	70	α-helix	
	β-sheet	35	others	40	others	30		

[a] 9 M LiBr/dialysis. [a] The 49.0 ppm peak of the Ala^{13} Cα carbon indicated the absence of the α-helix. [b] The chemical shift, 42.2 ppm, of the Gly^{20} Cα carbon and that, 170.8 ppm, of the Gly^{20} C=O carbon suggest that there are small amounts of β-sheet structure in addition to 3_1-helix structure.

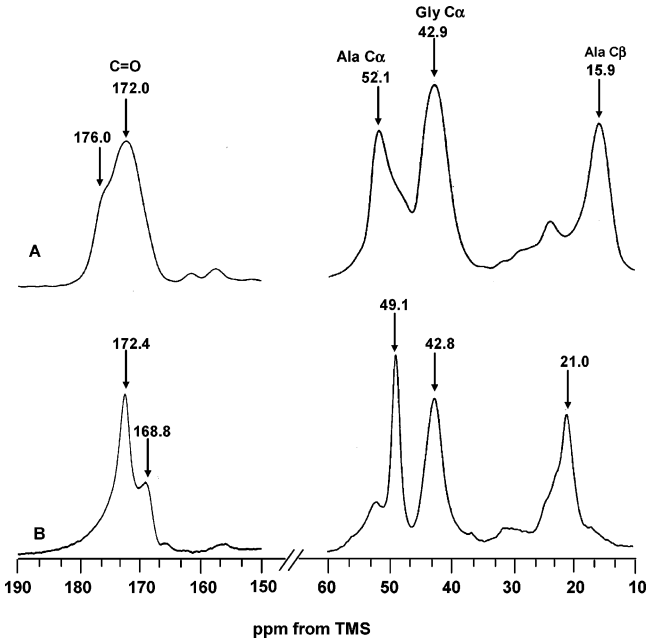

Figure 7. ^{13}C CP/MAS NMR spectra of peptide **I** after treatment with (A) TFA/diethyl ether and (B) FA/methanol.

corresponds approximately to the torsion angles of the 3_1-helix. If the peptide $(GGA)_n$ exclusively adopts repeated 3_1-helical conformation, the Ala C_β peak should be considerably narrower as reported previously.[40] The two-component nature of the mobility of Ala residues in spider dragline silk has been reported in the dried fiber.[7] A significant change in the motion of the Ala C_β carbon by hydration was also reported recently with WISE (wide-line separation) NMR of spider dragline silk.[16]

An increase in the fraction of β-sheet in both the poly-Ala and GG^{20}A^{21} regions was observed for **Ia** after it was dissolved in formic acid and then dried in air. Moreover, after **Ia** was dissolved in formic acid and then precipitated in methanol, the spectrum showed a tightly packed β-sheet structure with a further increase in the fraction of β-sheet although 15% 3_1-helix is still found in the GG^{20}A^{21} region. This seems due to the ability of methanol to promote the β-sheet structure by removing water from the peptide chains.[26] The β-sheet structure of the poly-Ala region and both 3_1-helix and β-sheet structures in the GGA sequence are in agreement with the structural model for the native spider dragline silk fiber from *N. clavipes* as showin in previous NMR studies. However, it was difficult to distinguish the peaks of Ala residues with 3_1-helix structure in GGA sequences from the poly-Ala peaks with β-sheet structure in the [1-^{13}C]Ala-labeled silk fiber samples because all Ala residues labeled uniformally The structure of the Gly-rich region could be deduced from [1-^{13}C]Gly-labeled silk fiber samples. In addition, the local structure of the Gly residues was defined in the GGA sequence.

TFA/D treatment induces a predominantly α-helical structure in the poly-Ala region as well as in the Gly-rich region. This indicated that the Gly residues in GGA sequences can be incorporated within the α-helical conformation of the peptide for a sequence containing both the polyalanine and the repeated GGA regions, QGAG(A)$_6$GGAGA(GGA)$_3$GAGRGGLGG synthesized here. The liquid silk stored in the spider dragline silk gland has been reported to be in an essentially random coil conformation[20] on the basis of the conformation-dependent Ala C_β chemical shift in the ^{13}C solution NMR spectrum of native liquid silk. Therefore, the predominantly α-helical peptide after TFA treatment indicates that it cannot be considered as a model for the spider silk fibroin in the silk gland. The situation is different, however, for the *S. cynthia ricini* silk fibroin after TFA treatment.

Thus, the conformation of the spider silk fibroin depends significantly on its environment. However, the primary sequence may help to define important features in the molecular structure, for example, giving elasticity to the fiber, controlling the formation of the β-sheet in the repeated sequence $(GGA)_n$, and defininig the 3_1-helix.

Acknowledgment. T.A. acknowledges support from the Insect Technology Project, Japan, and Agriculture Biotechnology Project, Japan. We also thank Dr. David Knight at the University of Oxford and Dr. Hajime Saito for useful discussions.

References and Notes

(1) O'Brien, J. P.; Fahnestock, S. R.; Termonia, Y.; Gardner, K. H. *Adv. Mater.* **1998**, *10*, 1185−1195.
(2) Gosline, J. M.; Guerette, P. A.; Ortlepp, C. S.; Savage, K. N. *J. Exp. Biol.* **1999**, *202*, 3295−3303.
(3) Vollrath, F.; Knight, D. P. *Nature* **2001**, *410*, 541−548.
(4) Xu, M.; Lewis, R. V. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7120−7124.
(5) Hinman, M. B.; Lewis, R. V. *J. Biol. Chem.* **1992**, *267*, 19320−19324.
(6) Simmons, A. H.; Ray, Ed.; Jelinski, L. W. *Macromolecules* **1994**, *27*, 5235−5337.
(7) Simmons, A. H.; Michal, C. A.; Jelinski, L. W. *Science* **1996**, *271*, 84−87.

(8) Kümmerlen, J.; van Beek, J. D.; Vollrath, F.; Meier, B. H. *Macromolecules* **1996**, *29*, 2920−2928.
(9) Michal, C. A.; Jelinski, L. W. *J. Biolmol. NMR* **1998**, *12*, 231−241.
(10) van Beek, J. D.; Kümmerlen, J.; Vollrath, F.; Meier, B. H. *Int. J. Biol. Macromol.* **1999**, *24*, 173−178.
(11) Seidel, A.; Liivak, O.; Calve, S.; Adaska, J.; Ji, G.; Yang, Z.; Grubb, D.; Zax, D. B.; Jelinski, L. W. *Macromolecules* **2000**, *33*, 775−780.
(12) Yang, Z.; Liivak, O.; Seidel, A.; Laverda, G.; Zax, D. B.; Jelinski, L. W. *J. Am. Chem. Soc.* **2000**, *122*, 9019−9025.
(13) van Beek, J. D.; Hess, S.; Vollrath, F.; Meier, B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10266−10271.
(14) Eles, P. T.; Michal, C. A. *Biomacromolecules* **2004**, *5*, 661−665.
(15) Eles, P. T.; Michal, C. A. *Macromolecules* **2004**, *37*, 1342−1345.
(16) Holland, G. P.; Lewis, R. V.; Yarger, J. L. *J. Am. Chem. Soc.* **2004**, *126*, 5867−5872.
(17) Grubb, D. T.; Jelinski, L. W. *Macromolecules* **1997**, *30*, 2860−2867.
(18) Riekel, C.; Bränden, C.; Craig, C.; Ferrero, C.; Heidelbach, F.; Müller, M. *Int. J. Biol. Macromol.* **1999**, *24*, 179−186.
(19) Hijirida, D. H.; Do, K. G.; Michal, C.; Wong, S.; Zax, D.; Jelinski, L. W. *Biophys. J.* **1996**, *71*, 3442−3447.
(20) Hronska, M.; van Beek, J. D.; Williamson, P. T. F.; Vollrath, F.; Meier, B. H. *Biomacromolecules* **2004**, *5*, 834−839.
(21) Asakura, T.; Sizuki, H.; Watanabe, Y. *Macromolecules* **1983**, *16*, 1024−1026.
(22) Asakura, T.; Watanabe, Y.; Uchida, A.; Minagawa, H. *Macromolecules* **1984**, *17*, 1075−1081.
(23) Asakura, T.; Watanabe, Y.; Itoh, T. *Macromolecules* **1984**, *17*, 2421−2426.
(24) Asakura, T. *Makromol. Chem., Rapid Commun.* **1986**, *7*, 775.
(25) Saito, H.; Iwanaga, Y.; Tabeta, R.; Narita, M.; Asakura, T. *Chem. Lett* **1983**, 431.
(26) Ishida, M.; Asakura, T.; Yokoi, M.; Saito, H. *Macromolecules* **1990**, *23*, 88−94.
(27) Yao, J.; Nakazawa, Y.; Asakura, T. *Biomacromolecules* **2004**, *5*, 680−688.
(28) Asakura, T.; Nakazawa, Y. *Macromol. Biosci.* **2004**, *4*, 175−185.
(29) Van Beek, J. D.; Beaulieu, L.; Schafe, r H.; Demura, M.; Asakura, T.; Meier, B. H. *Nature* **2000**, *405*, 1077−1079.
(30) Gosline, J. M.; Denny, M. W.; DeMont, M. E. *Nature* **1984**, *309*, 551−552.
(31) Thiel, B. L.; Guess, K. B.; Viney, C. *Biopolymers* **1997**, *41*, 703−719.
(32) Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komastu, K. *J. Mol. Biol.* **2001**, *306*, 291−305.
(33) Asakura, T.; Sugino, R.; Yao, J.; Takashiama, H.; Kishore, R. *Biochemistry* **2002**, *41*, 4415−4424.
(34) Asakura, T.; Yao, J.; Yamane, T.; Umemra, K.; Ulrich, A. S. *J. Am. Chem. Soc.* **2002**, *124*, 8794−8795.
(35) Asakura, T.; Yao, J. *Protein Sci.* **2002**, *11*, 2706−2713.
(36) Asakura, T.; Sugino, R.; Okumura, T.; Nakazawa, Y. *Protein Sci.* **2002**, *11*, 1873−1877.
(37) Nakazawa, Y.; Bamba, M.; Nishio, S.; Asakura, T. *Protein Sci.* **2003**, *12*, 666−671.
(38) Nakazawa, Y.; Asakura, T. *J. Am. Chem. Soc.* **2003**, *125*, 7230−7237.
(39) Yao, J.; Sugino, R.; Kishore, R.; Asakura, T. *Biomacromolecules* **2004**, *5*, 1763−1769.
(40) Ashida, J.; Ohgo, K.; Komastu, K.; Kubota, A.; Asakura, T. *J. Biol. NMR* **2003**, *25*, 91−103.
(41) Asakura, T.; Iwadate, M.; Demura, M.; Williamson, M. P. *Int. J. Biol. Macromol.* **1999**, *24*, 167−171.
(42) Saito, H.; Tabeta, R.; Shoji, A.; Ozaki, T.; Anod, I.; Miyata, T. *Biopolymers* **1984**, *23*, 2279−2297.
(43) Blondelle, S. E.; Forood, B.; Houghten, R. A.; Pérez-Payá, E. *Biochemistry* **1997**, *36*, 8393−8400.
(44) Demura, M.; Minami, M.; Asakura, T.; Cross, T. *J. Am. Chem. Soc.* **1998**, *120*, 1300−1308.

MA047660Z