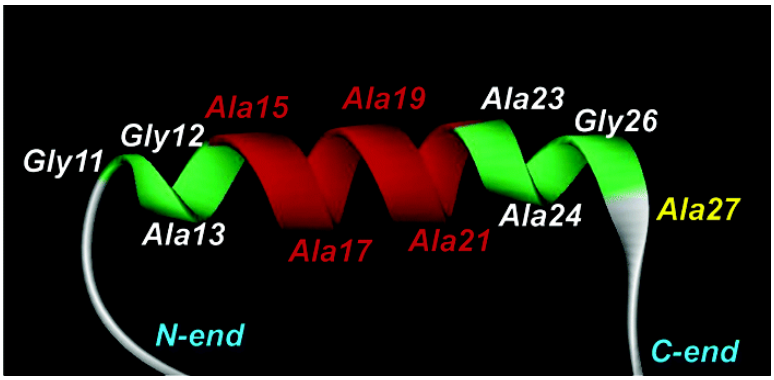
Article

Structure Determination of a Peptide Model of the Repeated Helical Domain in *Samia cynthia ricini* Silk Fibroin before Spinning by a Combination of Advanced Solid-State NMR Methods

Yasumoto Nakazawa, and Tetsuo Asakura

J. Am. Chem. Soc., **2003**, 125 (24), 7230-7237• DOI: 10.1021/ja0300721 • Publication Date (Web): 21 May 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 05/21/2003

Structure Determination of a Peptide Model of the Repeated Helical Domain in *Samia cynthia ricini* Silk Fibroin before Spinning by a Combination of Advanced Solid-State NMR Methods

Yasumoto Nakazawa and Tetsuo Asakura*

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

Received January 31, 2003; Revised Manuscript Received March 14, 2003; E-mail: Asakura@cc.tuat.ac.jp

Abstract: Fibrous proteins unlike globular proteins, contain repetitive amino acid sequences, giving rise to very regular secondary protein structures. Silk fibroin from a wild silkworm, *Samia cynthia ricini*, consists of about 100 repeats of alternating polyalanine (poly-Ala) regions of 12−13 residues in length and Gly-rich regions. In this paper, the precise structure of the model peptide, GGAGGGYGGDGG$(A)_{12}$GGAGDGYGAG, which is a typical repeated sequence of the silk fibroin, was determined using a combination of three kinds of solid-state NMR studies; a quantitative use of ^{13}C CP/MAS NMR chemical shift with conformation-dependent ^{13}C chemical shift contour plots, 2D spin diffusion ^{13}C solid-state NMR under off magic angle spinning and rotational echo double resonance. The structure of the model peptide corresponding to the silk fibroin structure before spinning was determined. The torsion angles of the central Ala residue, Ala_{19}, in the poly-Ala region were determined to be $(\phi, \psi) = (-59°, -48°)$ which are values typically associated with α-helical structures. However, the torsion angles of the Gly_{25} residue adjacent to the C-terminal side of the poly-Ala chain were determined to be $(\phi, \psi) = (-66°, -22°)$ and those of Gly_{12} and Ala_{13} residues at the N-terminal of the poly-Ala chain to be $(\phi, \psi) = (-70°, -30°)$. In addition, REDOR experiments indicate that the torsion angles of the two C-terminal Ala residues, Ala_{23} and Ala_{24}, are $(\phi, \psi) = (-66°, -22°)$ and those of N-terminal two Ala residues, Ala_{13} and Ala_{14} are $(\phi, \psi) = (-70°, -30°)$. Thus, the local structure of N-terminal and C-terminal residues, and also the neighboring residues of α-helical poly-Ala chain in the model peptide is a more strongly wound structure than found in typical α-helix structures.

Introduction

Recently developed solid-state NMR methodology can provide accurate structural constraints on peptides and proteins of substantial complexity in noncrystalline solids.[1] Atomic-level information, such as the torsion angles, ϕ and ψ, has been obtained for membrane-bound peptide systems,[2−6] amyloid fibrils,[1,7−9] and silk fibers from silkworms or spiders.[10−18]

There are many kinds of silks from silkworms and spiders with different structures and properties.[19] The molecular design of fibers with high strength and high elasticity[20] can be aided by studying the structure−property relationships of natural silks. The silk fibroin from the domesticated silkworm, *Bombyx mori* (*B. mori*), is a well-known fibrous protein whose amino acid composition (in mol %) is 42.9% Gly, 30.0% Ala, 12.2% Ser, 4.8% Tyr, and 2.5% Val.[21,22] The details of the secondary structures before and after the spinning process have recently

(1) Tycko, R. *Annu. Rev. Phys. Chem.* **2001**, *52*, 575−606.
(2) Grage, S. L.; Wang, J.; Cross, T. A.; Ulrich, A. S. *Biophys. J.* **2002**, *83*, 3336−3350.
(3) Salgado, J.; Grage, S. L.; Kondejewski, L. H.; Hodges, R. S.; McElhaney, R. N.; Ulrich, A. S. *J. Biomol. NMR* **2001**, *21*, 191−208.
(4) Kumashiro, K. K.; Schmidt-Rohr, K.; Murphy, O. J.; Ouellette, K. L.; Cramer, W. A.; Thompson, L. K. *J. Am. Chem. Soc.* **1998**, *120*, 5043−5051.
(5) Huster, D.; Yao, X.; Hong, M. *J. Am. Chem. Soc.* **2002**, *124*, 874−883.
(6) Hori, Y.; Demura, M.; Iwadate, M.; Ulrich, A. S.; Niidome, T.; Aoyagi, H.; Asakura, T. *Eur. J. Biochem.* **2001**, *268*, 302−309.
(7) Antzutkin, O. N.; Balbach, J. J.; Leapman, R. D.; Rizzo, N. W.; Reed, J.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13045−13050.
(8) Balbach, J. J.; Petkova, A. T.; Oyler, N. A.; Antzutkin, O. N.; Gordon, D. J.; Meredith, S. C.; Tycko, R. *Biophys. J.* **2002**, *83*, 1205−1216.
(9) Benzinger, T. L.; Gregory, D. M.; Burkoth, T. S.; Miller-Auer, H.; Lynn, D. G.; Botto, R. E.; Meredith, S. C. *Biochemistry* **2000**, *39*, 3491−3499.
(10) Demura, M.; Minami, M.; Asakura, T.; Cross, T. A. *J. Am. Chem. Soc.* **1998**, *120*, 1300−1308.
(11) Asakura, T.; Ito, T.; Okudaira, M.; Kameda, T. *Macromolecules* **1999**, *32*, 4940−4946.
(12) van Beek, J. D.; Beaulieu, L.; Schafer, H.; Demura, M.; Asakura, T.; Meier, B. H. *Nature* **2000**, *405*, 1077−1079.
(13) Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komatsu, K. *J. Mol. Biol.* **2001**, *306*, 291−305.
(14) Asakura, T.; Yamane, T.; Nakazawa, Y.; Kameda, T.; Ando, K. *Biopolymers* **2001**, *58*, 521−525.
(15) Ashida, J.; Ohgo, K.; Asakura, T. *J. Phys. Chem. B.* **2002**, *106*, 6434−9439.
(16) van Beek, J. D.; Hess, S.; Vollrath, F.; Meier, B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10266−10271.
(17) Nakazawa, Y.; Bamba, M.; Nishio, S.; Asakura, T. *Protein Sci.* **2003**, *12*, 666−671.
(18) Ashida, J.; Ohgo, K.; Komatsu, K.; Kubota, A.; Asakura, T. *J. Biomol. NMR* **2003**, *25*, 91−103.
(19) Gosline, J. M.; Guerette, P. A.; Ortlepp, C. S.; Savage, K. N. *J. Exp. Biol.* **1999**, *202*, 3295−3303.
(20) Heslot, H. *Biochimie* **1998**, *80*, 19−31.
(21) Shimura, K. In *Zoku Kenshi no Kozo (Structure of Silk Fibers)*; Hojyo, N., Ed.; Shinshu University: Ueda, Japan, 1980; pp 335−352.
(22) Zhou, C.; Confalonieri, F.; Medina, N.; Zivanovic, Y.; Esnault, C.; Yang, T.; Jacquet, M.; Janin, J.; Duguet, M.; Perasso, R.; Li, Z. *Nucleic. Acids. Res.* **2000**, *28*, 2413−2419.

10.1021/ja0300721 CCC: $25.00

been reported with several solid-state NMR methods including ^{13}C two-dimensional (2D) spin-diffusion solid-state NMR, rotational-echo double-resonance (REDOR) NMR, and the quantitative use of the conformation-dependent ^{13}C CP/MAS chemical shifts.[13,14] Two-dimensional spin-diffusion solid-state NMR is a powerful method to obtain the relative orientation of two chemical shift tensors of ^{13}C-labeled sites in the local molecular framework,[13,23] and REDOR is a solid-state recoupling NMR technique for the determination of the internuclear distances between ^{13}C and ^{15}N nuclei.[13,24−27] The ^{13}C CP/MAS NMR chemical shifts of the amino acid residues in peptides and proteins are conformation-dependent and can be used to restrict allowed regions of the backbone torsion angles. A combination of these three solid-state NMR is a very powerful way for determination of the structure of the amino acid residues in proteins, including silk fibroin.[28]

The amino acid composition of silk fibroin from a wild silkworm, *Samia cynthia ricini* (*S. c. ricini*), is considerably different from that of *B. mori* silk fibroin. The sum of Gly and Ala residues is 82% which is basically the same as that of *B. mori* silk (73%), but the relative composition of Ala and Gly is reversed.[21] The proportion of Gly residues is greater in *B. mori* silk fibroin, while the content of Ala residues is greater in *S. c. ricini* silk fibroin. The primary structure of the silk fibroin from *S. c. ricini* has recently been determined by Yukihiro et al.[29] and was found to be very similar to the structure of silk fibroin from another wild silkworm, *Antheraea pernyi*.[30,31] In particular, the silk mainly consists of similar sequences repeated about 100 times. The basic repeat sequence is made of alternating $(Ala)_{12-13}$ region and the Gly-rich regions similar to spider (major ampullate) silk. Here, the polyalanine sequences such as $(Ala)_{12-13}$ or $(Ala)_{12}$ are defined as poly-Ala in this paper. From the solution-state ^{13}C and ^{15}N NMR studies of *S. c. ricini* silk fibroin in aqueous solution, it is found that about 70% of Ala residues form α-helices, while the conformation of the other Ala residues is random coil.[32−35] The fast exchange between helix and coil forms in the poly-Ala region has been observed because of the gradual high field shift of the main peak assigned to the carbonyl carbons of poly-Ala during helix-to-coil transition with increasing temperature.[32] The helicity of each Ala residue in poly-Ala region could be calculated theoretically according to the Bixon−Scheraga−Lifson theory[36] for the helix-to-coil transition of poly-Ala including the hydrophobic side-chain interactions. The change in the NMR spectra of the Ala carbonyl region due to the temperature-induced helix-to-coil transition was interpreted in terms of the change in the statistical weight parameter ω, where ω is related to the formation of an intramolecular hydrogen bond. On the other hand, most of the glycine carbonyl peaks in the ^{13}C solution NMR spectrum of [1-^{13}C]Gly-silk fibroin could be assigned to the primary structure by comparison with the ^{13}C chemical shifts of seven glycine-containing tripeptides.[35] The slow-exchange between helix and coil forms in the NMR time scale was observed with increasing temperature exclusively for the underscored glycine residue in the Gly-Gly-$(Ala)_{12-13}$ sequence during fast helix-to-coil transition of the $(Ala)_{12-13}$ region. For the determination of the structure of *S. c. ricini* silk fibroin fiber, the ^{15}N- and ^{13}C-labelings of *S. c. ricini* silk fiber were performed, and the blocks of the oriented silk fibers were observed by changing the orientations of the oriented silk fiber axis with respect to the magnetic field.[11] All spectra of [^{15}N]Gly or [1-^{13}C]Gly silk fibroin fibers were slightly broader than the corresponding spectra of ^{15}N- and ^{13}C-Ala-labeled silk fibers. The fraction of the oriented Ala domain in the silk fiber was determined as 75%, while that of the oriented Gly domain 65%. The torsion angles of the oriented Ala and Gly residues were determined to be $(\phi, \psi) = (-130° \pm 10°, 140° \pm 10°)$ for both residues.

In this contribution, three kinds of solid-state NMR methods were used to determine the precise structure of the model peptide, GGAGGGYGGDGG$(A)_{12}$GGAGDGYGAG, which is a typical repeated sequence of *S. c. ricini* silk fibroin. Namely, a combination of 2D spin-diffusion ^{13}C solid-state NMR under off magic angle spinning (OMAS) and a quantitative use of ^{13}C CP/MAS NMR chemical shift with conformation-dependent ^{13}C chemical shift contour plots was used to determine the torsion angles, ϕ and ψ, of the N-terminal and the central Ala residues in the $(Ala)_{12}$ chain and of two Gly residues adjacent to the $(Ala)_{12}$ chain. Rotational-echo double-resonance was also used to determine the helical structures of two Ala residues at both N- and C-terminal sites of the $(Ala)_{12}$ chain. The structure of the silk fibroin before spinning in the solid state is the focus of this work.

Experimental Section

1. Preparation of the Stable Isotope-Labeled Model Peptides. The ^{13}C-labeled and ^{15}N-labeled model peptides, GGAGGGYGGDGG-$(A)_{12}$GGAGDGYGAG, which were selected as a typical repeated sequence of *S. c. ricini* silk fibroin, were synthesized by the solid-phase method (Table 1). After synthesis, the samples were dissolved in 9 M LiBr and then dialyzed against water for 4 days. The precipitated samples were collected and dried. However, the structure was β-sheet by judging from the ^{13}C CP/MAS NMR chemical shifts of the Ala residue. To transform the β-sheet form into the α-helical form as a structural model for the silk fibroin before spinning, the precipitated samples were dissolved in TFA and followed by the addition of diethyl ether. After drying, the precipitated samples formed an α-helix as judged from the ^{13}C CP/MAS NMR chemical shifts of the Ala residues.

2. ^{13}C CP/MAS NMR. The ^{13}C CP/MAS NMR experiments were performed on a Chemagnetics Infinity 400 MHz spectrometer with a ^{13}C operating frequency of 100.04 MHz. Samples were spun at a rate of 5 kHz. The number of acquisitions was 8000, and the recycle delays were 5 s. A 50 kHz radio frequency field strength was used for $^{1}H-^{13}C$ decoupling the acquisition period of 12.8 ms. A 90° pulse width of 5 μs with 1 ms CP contact time was employed. Phase cycling was used to minimize artifacts. ^{13}C chemical shifts were calibrated indirectly through the adamantane methylene peaks observed at 28.8 ppm relative to tetramethylsilane at 0 ppm.

3. 2D-Spin Diffusion ^{13}C Solid-State NMR under off Magic Angle Condition. The 2D spin-diffusion NMR spectra were obtained using a Varian Unity INOVA 400 NMR spectrometer with a 7 mm Jakobsen-

(23) Kummerlen, J.; vanBeek, J. D.; Vollrath, F.; Meier, B. H. *Macromolecules* **1996**, *29*, 2920−2928.
(24) Gullion, T.; Schaefer, J. *Adv. Magn. Reson.* **1989**, *13*, 57−83.
(25) Gullion, T. *J. Magn. Reson.* **2000**, *146*, 220−222.
(26) Kameda, T.; Nakazawa, Y.; Kazuhara, J.; Yamane, T.; Asakura, T. *Biopolymers* **2002**, *64*, 80−85.
(27) Michal, C. A.; Jelinski, L. W. *J. Biomol. NMR* **1998**, *12*, 231−241.
(28) Zhao, C.; Asakura, T. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 301−352.
(29) Yukuhiro, K. Personal communication.
(30) Sezutsu, H.; Yukuhiro, K. *J. Mol. Evol.* **2000**, *51*, 329−338.
(31) Nakazawa, Y.; Asakura, T. *Macromolecules* **2002**, *35*, 2393−2400.
(32) Asakura, T.; Murakami, T. *Macromolecules* **1985**, *18*, 2614−2619.
(33) Asakura, T.; Kashiba, H.; Yoshimizu, H. *Macromolecules* **1988**, *21*, 644−648.
(34) Asakura, T.; Yoshimizu, H.; Yoshizawa, F. *Macromolecules* **1988**, *21*, 2038−2041.
(35) Nakazawa, Y.; Asakura, T. *FEBS Lett.* **2002**, *529*, 188−192.
(36) Bixon, M.; Scheraga, H. A.; Lifson, S. *Biopolymers* **1963**, *1*, 419−429.

Table 1. Several ^{13}C- and ^{15}N-Labeled Model Peptides of Silk Fibroin from *S. c. ricini*, $GGAGGGYGGD_{10}GGAAAAAAAA_{20}AAAAGGAGDG_{30}YGAG$ Synthesized by the Solid-Phase Method for 2D Spin-Diffusion and REDOR Experiments

peptide	method	information
1;$GGAGGGYGGDGG(A)_{12}GGAGDGYGAG$	^{13}C CP/MAS NMR	secondary structure
2;$GGAGGGYGGDGG(A)_5[1\text{-}^{13}C]A_{18}\,[1\text{-}^{13}C]A_{19}(A)_5GGAGDGYGAG$	spin-diffusion	Ala^{19} (ϕ, ψ) torsion angle
3;$GGAGGGYGGDGG(A)_{11}[1\text{-}^{13}C]A_{24}\,[1\text{-}^{13}C]G_{25}GAGDGYGAG$	spin-diffusion	Gly_{25} (ϕ, ψ) torsion angle
4;$GGAGGGYGGDG[1\text{-}^{13}C]G_{12}[1\text{-}^{13}C]A_{13}(A)_{11}GGAGDGYGAG$	spin-diffusion	Ala_{13} (ϕ, ψ) torsion angle
5;$GGAGGGYGGD[1\text{-}^{13}C]G_{11}[1\text{-}^{13}C]G_{12}(A)_{12}GGAGDGYGAG$	spin-diffusion	Gly_{12} (ϕ, ψ) torsion angle
6;$GGAGGGYGGDGG(A)_8[1\text{-}^{13}C]A_{21}AAA\,[^{15}N]G_{25}GAGDGYGAG$	REDOR	$[1\text{-}^{13}C]Ala_{21}$-$[^{15}N]Gly_{25}$ distance
7;$GGAGGGYGGDG[1\text{-}^{13}C]G_{12}AAA[^{15}N]A_{16}(A)_8GGAGDGYGAG$	REDOR	$[1\text{-}^{13}C]Gly_{12}$-$[^{15}N]Ala_{16}$ distance

type double-tuned MAS probe at off magic angle condition ($\theta_m - 7°$) at room temperature. The sample spinning rate was 6 kHz (±3 Hz). The scaling factor of the 2D spin-diffusion spectra is $^1/_2(3\cos^2(\theta_m - 7°) - 1) = 0.198$. The mixing times were set to be 2 s. They were optimized for spin diffusion between intramolecular specific carbon atoms of selectively isotope-labeled Ala and Gly residues or both, but with no spin diffusion among carbon atoms in difference molecules. The contact time was set to 2 ms using the variable-amplitude CP technique.[37] About 400 scans with a recycle delay of 2 s were accumulated for every T_1 value in the 2D experiment. The recycle delay of 2 s was carefully determined from the 1D OMAS spectra under several recycle times. The principal values of the chemical shift tensors for the carbonyl carbon nuclei of the ^{13}C-labeled Ala and Gly residues were determined by analysis of the spinning sidebands under slow MAS conditions[13,38] using a Chemagnetics Infinity 400 spectrometer.

4. REDOR. REDOR experiments were performed on a Chemagnetics Infinity 400 MHz spectrometer equipped with solid-state accessories and a triple-channel magic-angle probe with a 5 mm APEX triple-resonance MAS probe. Samples were spun at a rate of 6 kHz (±3 Hz). ^{13}C REDOR NMR spectra were obtained at 100.0 MHz following matched, $^1H-^{13}C$ cross-polarization contacts of 50 kHz. Contact times of 1.0 ms were used. The $\pi/2$ pulses for ^{13}C and ^{15}N channels were 3.0 and 4.1 ms, respectively. A recycle delay of 3.0 s was used throughout these experiments, and data were obtained with 64 transients. Phases of ^{15}N π pulses were cycled according to the XY-8 scheme to minimize off-resonance and pulse error effects.[39,40] REDOR evolution times ranged up to 24 ms (64 rotor cycles). Values of $\Delta S/S_0 = 1 - S/S_0$ were computed as ratios of the peak heights in the REDOR spectra.

Results

2D Spin-Diffusion ^{13}C Solid-State NMR Spectra under off Magic Angle Condition. As shown in Figure 1, the chemical shifts of Ala residues in the ^{13}C CP/MAS NMR spectrum of the model peptide, $GGAGGGYGGDGG(A)_{12}GGAGDGYGAG$ after TFA treatment were 15.3 ppm for Ala Cβ, 52.3 ppm for Ala Cα and 175.9 ppm for Ala carbonyl carbons. These are essentially the same as the chemical shift values of *S. c. ricini* silk fibroin film in α-helical conformation;15.2 ppm for Ala Cβ, 52.4 ppm for Ala Cα and 176.2 ppm for Ala carbonyl carbons.[32,33] Thus, it is concluded that both conformations of the silk fibroin film and the sequential model peptide after TFA treatment are α- helices.

For 2D spin-diffusion experiments, the following ^{13}C double-labeled model peptides were synthesized:

2;$GGAGGGYGGDGG(A)_5[1\text{-}^{13}C]A_{18}[1\text{-}^{13}C]A_{19}(A)_5GGAGDGYGAG$, **3**;$GGAGGGYGGDGG(A)_{11}[1\text{-}^{13}C]A_{24}[1\text{-}^{13}C]G_{25}GAGDGYGAG$, **4**;$GGAGGGYGGDG[1\text{-}^{13}C]G_{12}[1\text{-}^{13}C]A_{13}(A)_{11}GGAGDGYGAG$, **5**;$GGAGGGYGGD[1\text{-}^{13}C]G_{11}[1\text{-}^{13}C]G_{12}(A)_{12}GGAGDGYGAG$

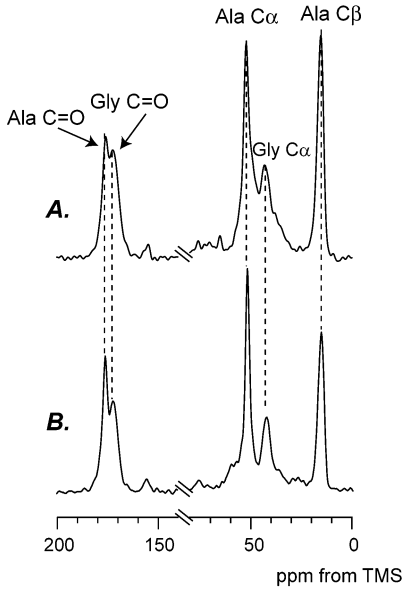


Figure 1. ^{13}C CP/MAS NMR spectra of the model peptide, **1**;$GGAGGGYGGDGG(A)_{12}GGAGDGYGAG$ after TFA treatment (A) and *S. c. ricini* silk fibroin film prepared from the silk fibroin stored in the silk gland (B).

The results are shown in Figure 2, where the carbonyl carbon regions are expanded. The torsion angles of Ala_{19}, Gly_{25}, Ala_{13} and G_{12} residues, respectively, can be determined by simulation of the observed spectra. The experimental spectra of 2−5 are similar with each other, but subtle differences are present. The 2D spin-diffusion spectra of the Ala residues were calculated with the principal values of the chemical shift tensors of the carbonyl carbon atoms of the Ala_{19} residue and shown in Figure 3. The conformations of both Ala_{19} and Ala_{13} residues are in the α-helix region of the Ramachandran plot. The shaded regions of Figure 3 are consistent with the measured 15.3 ppm chemical shift of the Ala Cβ carbon. Thus, both Ala_{13} and Ala_{19} residues are in the α-helix conformation, but there are still small differences between their respective spin-diffusion patterns. Contrary to the case of Ala residue, there are four candidates which satisfy the observed spectral patterns of both Gly_{25} and Gly_{12} residues, as shown in Figure 4. However, the observed ^{13}C chemical shift values of the Gly residues, 43.9 ppm, also satisfy only the α-helix region as shown in Figure 4 (grey regions), and their analysis will be performed below to determine the torsion angles.

2. Determination of the Torsion Angles of Ala_{13}, Ala_{19}, Gly_{12}, and Gly_{25} Residues. The χ^2 deviations between observed and calculated 2D spin-diffusion spectra were introduced to

(37) Peersen, O. B.; Wu, X. L.; Kustanovich, I.; Smith, S. O. *J. Magn. Reson., Ser. A* **1993**, *104*, 334−339.
(38) Herzfeld, J.; Berger, A. E. *J. Chem. Phys.* **1980**, *73*, 6021−6030.
(39) Gullion, T.; Schaefer, J. *J. Magn. Reson.* **1991**, *92*, 439−442.
(40) Gullion, T.; Baker, D. B.; Schaefer, J. *J. Magn. Reson.* **1990**, *89*, 479−484.

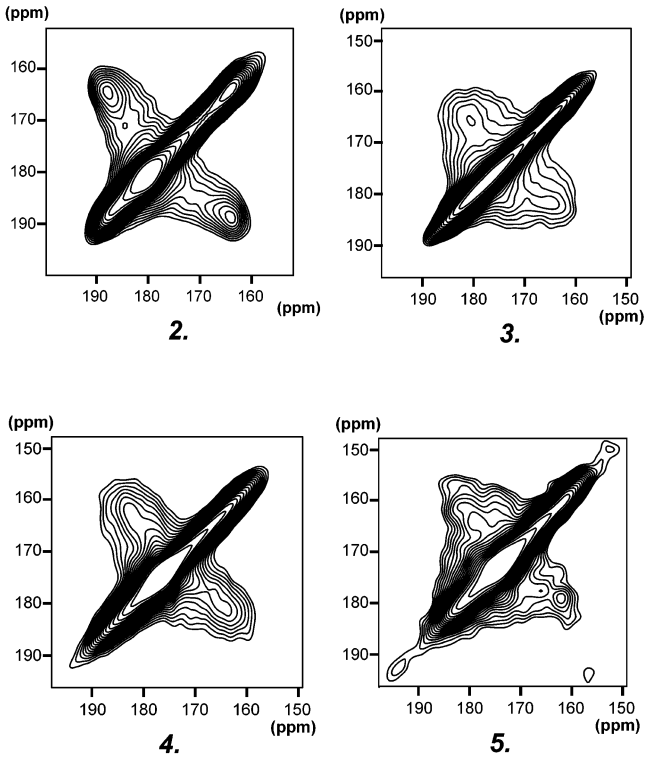


Figure 2. Two-dimensional spin-diffusion ^{13}C solid state NMR spectra (only the carbonyl region is expanded) of four kinds of the model peptides, **2**;GGAGGGYGGDGG(A)$_5$[1-^{13}C]A$_{18}$[1-^{13}C]A$_{19}$-(A)$_5$GGAGDGYGAG, **3**;GGAGGGYGGDGG(A)$_{11}$[1-^{13}C]A$_{24}$[1-^{13}C]G$_{25}$GAGDGYGAG,**4**;GGAGGGYGGDG[1-^{13}C]G$_{12}$[1-^{13}C]A$_{13}$(A)$_{11}$GGAGDGYGAG, **5**;GGAGGGYGGD[1-^{13}C]G$_{11}$[1-^{13}C]G$_{12}$(A)$_{12}$GGAGDGYGAG. The structures of the polyalanine region were basically helix judging from the ^{13}C CP/MAS NMR chemical shifts of the Ala residue.

determine the torsion angles quantitatively.[41] The definition of the χ^2 is as follows.

$$\chi^2(\phi,\psi) = \frac{1}{\sigma^2}\sum_{i=1}^{N}[E_i - S_i(\phi,\psi)]^2$$

where σ is the root-mean-squared noise in the experimental spectrum, N is the number of intensities analyzed, E_i are the experimental intensities, $S_i(\phi, \psi)$ are the calculated intensities. The χ^2 deviations were calculated as a function of the torsion angles, ϕ and ψ for Ala$_{19}$ residue in the α-helix region (ϕ, ψ) = (−50° to −70°, −35° to −55°). The results are shown in Figure 5. The minimum corresponds to (ϕ, ψ) = (−59°, −48°), indicating that a typical α-helix conformation of Ala$_{19}$ residue. The torsion angles of the typical α-helix structure are $\phi = -57°$ and $\psi = -47°$.[42] Thus, the torsion angles of the Ala residue determined here are in agreement with the typical angles of α-helix and also the angles, ((ϕ, ψ) = (−60°, −45°), reported for the 2D DOQSY NMR of the silk fibroin film from *S. c. ricini* by Jacco et al.[12]

The χ^2 deviations were also calculated as a function of the torsion angles, ϕ and ψ, for Ala$_{13}$ residue in the α-helical region (ϕ, ψ) = (−50° to −90°, −10° to −50°) and shown in Figure 6. There are two minima at (ϕ, ψ) = (−70°, −30°) and (ϕ, ψ) = (−80°, −33°). It seems difficult to determine the torsion angles from only these data and therefore the distance constraint determined by a REDOR experiment (see below) was used to distinguish between the two minima. We selected (ϕ, ψ) = (−70°, −30°) from analysis of the REDOR data.

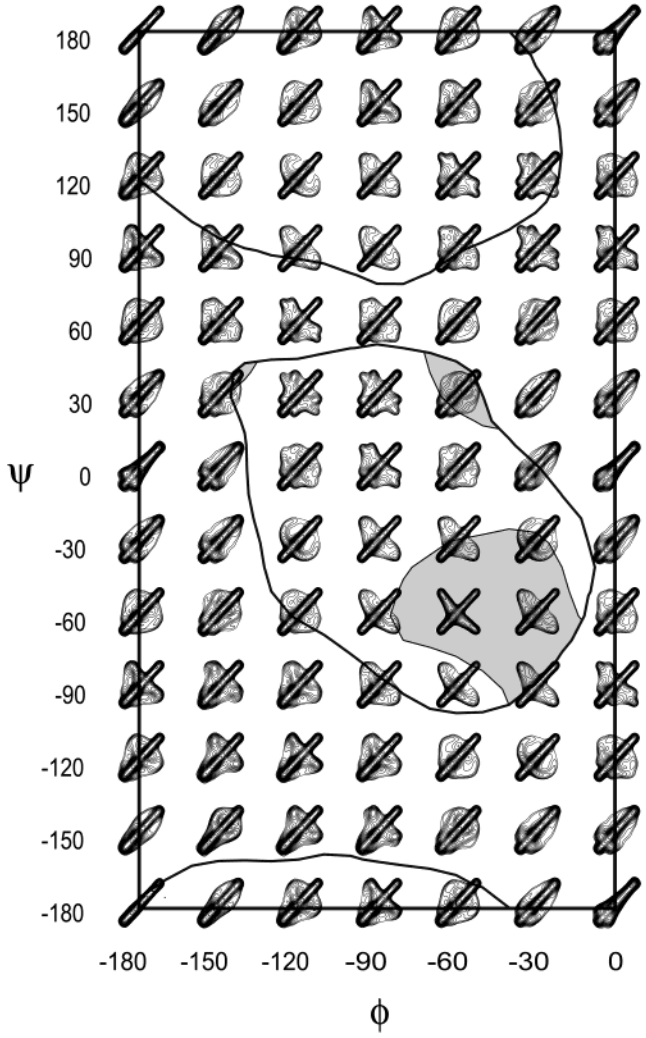


Figure 3. Ramachandran map of the calculated 2D spin-diffusion NMR spectra as a function of the torsion angles ϕ and ψ of Ala residue for each 30° in the region of $-180° < \phi < 0°$ and $-180° < \psi < 180°$ in which the density function is greater than 1, are surrounded by the solid lines. The area which satisfies the observed Ala Cβ chemical shift of 15.3 ppm, was shown as a shadow region, where random coil chemical shift of Ala Cβ carbon is 16.6 ppm.

Determination of the torsion angles using the χ^2 calculation was also performed for Gly residue. Figure 7 shows χ^2 plot for Gly$_{25}$ residue in the α-helical region (ϕ, ψ) = (−50° to −90°, −10° to −50°). The torsion angles were determined to be (ϕ, ψ) = (−66°, −22°). For the Gly$_{12}$ residue, the observed 2D spin-diffusion spectral pattern appear slightly diffuse as is shown in Figure 2. Unfortunately, we could not obtain the small χ^2 deviation region in the χ^2 plot. We assumed here that the torsion angles are the same as the angles of Ala$_{13}$ residue, (ϕ, ψ) = (−70°, −30°).

3. Determination of the Helical Structure at the C-Terminal Side of the Polyalanine Region. The torsion angles of the central Ala residue, Ala$_{19}$, in the poly-Ala chain were determined as (ϕ, ψ) = (−59°, −48°) and those of the Gly$_{25}$ residue, which is adjacent to the C-terminal Ala residue, were found to be (ϕ, ψ) = (−66°, −22°). To explore the region where

(41) Long, H. W.; Tycko, R. *J. Am. Chem. Soc.* **1998**, *120*, 7039−7048.
(42) Creighton, T. E. In *Proteins: Structures and Molecular Properties*, 2nd ed.; W. H. Freeman: New York, 1993.

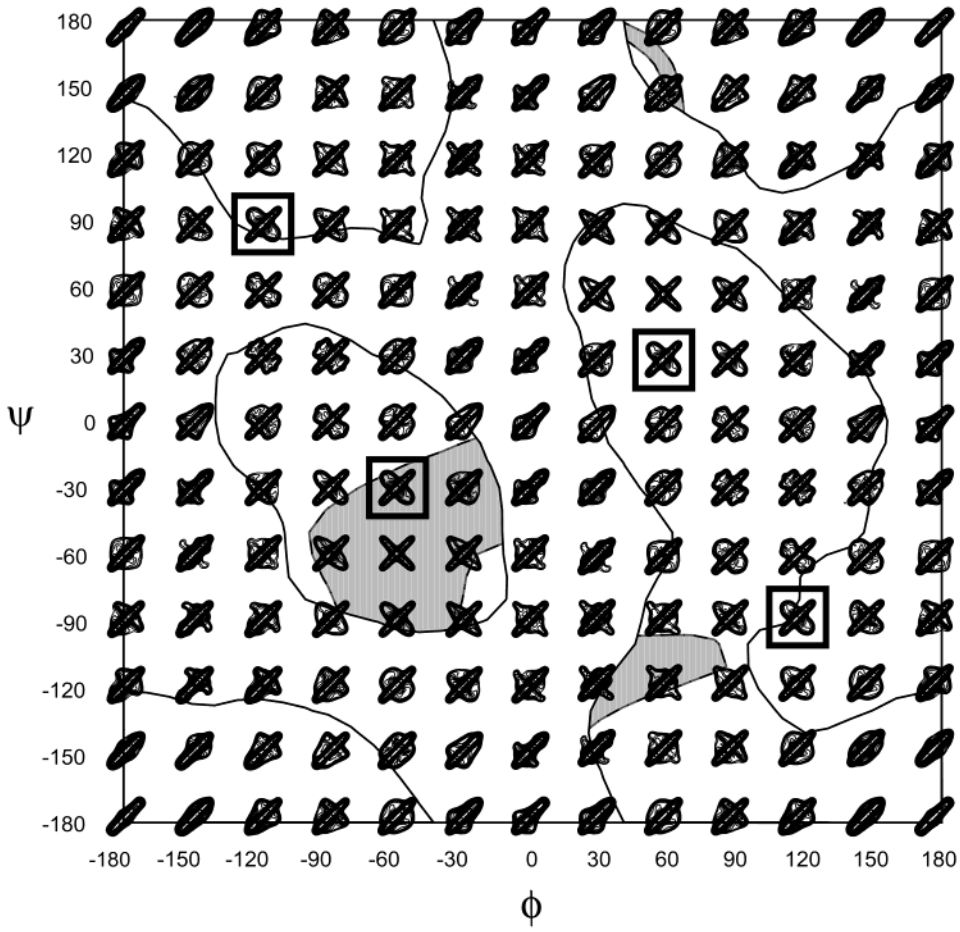


Figure 4. Ramachandran map of the calculated 2D spin-diffusion NMR spectra as a function of the torsion angles ϕ and ψ of Gly residue for each 30° in the region of $-180° < \phi < 180°$ and $-180° < \psi < 180°$ in which the density function is greater than 1, are surrounded by the solid lines. The area which satisfies the observed Gly Cα chemical shift of 43.9 ppm, was shown as a shadow region, where random coil chemical shift of Gly Cα carbon is 42.7 ppm.

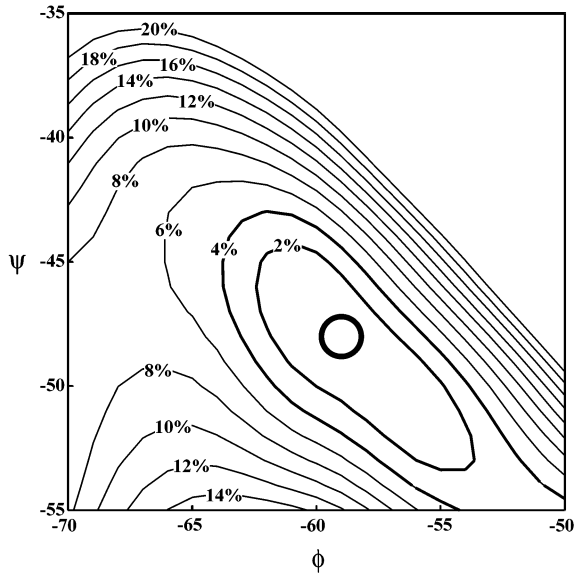


Figure 5. Contour plot of χ^2 deviation (in %) between observed and calculated 2D spin diffusion NMR spectra of **2**;GGAGGGYGGDGG(A)$_5$-[1-^{13}C]A$_{18}$[1-^{13}C]A$_{19}$(A)$_5$GGAGDGYGAG in the region of $-70° < \phi < -55°$ and $-55° < \psi < -35°$. The circle means the smallest χ^2 deviation. The torsion angles were determined as $(\phi, \psi) = (-59°, -48°)$.

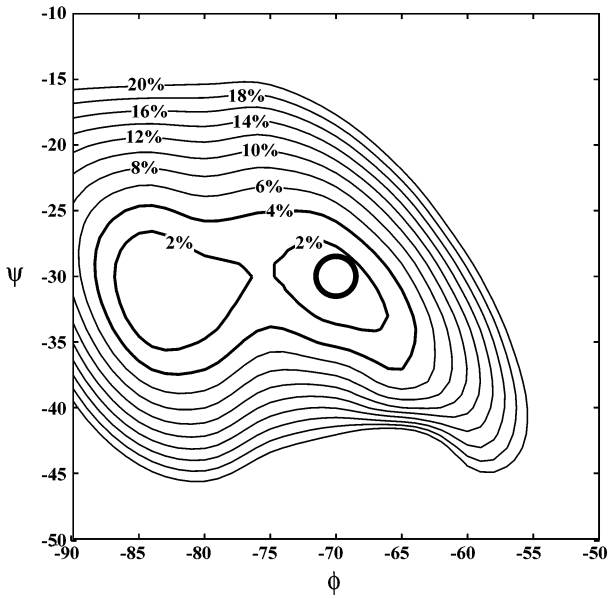


Figure 6. Contour plot of χ^2 deviation (in %) between observed and calculated 2D spin diffusion NMR spectra of **4**;GGAGGGYGGDG[1-^{13}C]-G$_{12}$[1-^{13}C]A$_{13}$(A)$_{11}$GGAGDGYGAG in the region of $-90° < \phi < -50°$ and $-50° < \psi \leq 10°$. There are two minima at $(\phi, \psi) = (-70°, -30°)$ and at $(\phi, \psi) = (-80°, -33°)$. However, by additional information on the distance constraint determined using REDOR method, the torsion angles were determined as $(\phi, \psi) = (-70°, -30°)$. Details are described in the text.

the torsion angle changes from $(\phi, \psi) = (-59°, -48°)$ to $(\phi, \psi) = (-66°, -22°)$ along the poly-Ala chain, ^{13}C-^{15}N REDOR experiments were performed on GGAGGGYGGDGG(A)$_8$[1-

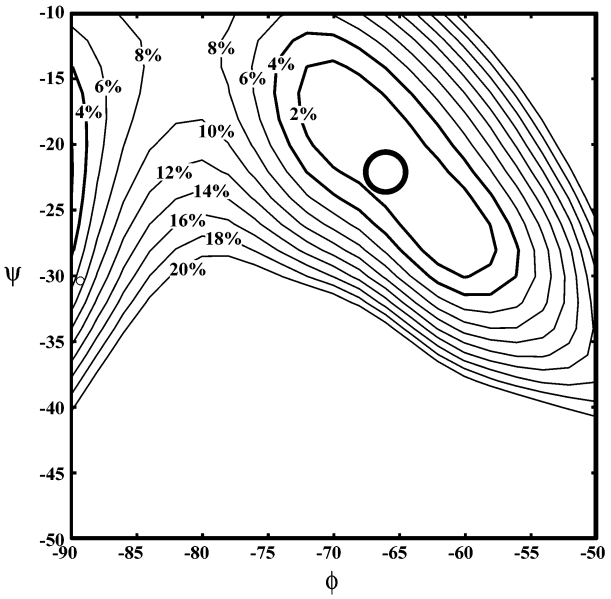


Figure 7. Contour plot of χ^2 deviation (in %) between observed and calculated 2D spin diffusion NMR spectra of **3**;GGAGGGYGGDGG(A)$_{11}$-[1-^{13}C]A$_{24}$[1-^{13}C]G$_{25}$GAGDGYGAG in the region of $-90° < \phi < -50°$ and $-50° < \psi < -10°$. The circle means the smallest χ^2 deviation. The torsion angles were determined as (ϕ, ψ) = ($-66°$, $-22°$).

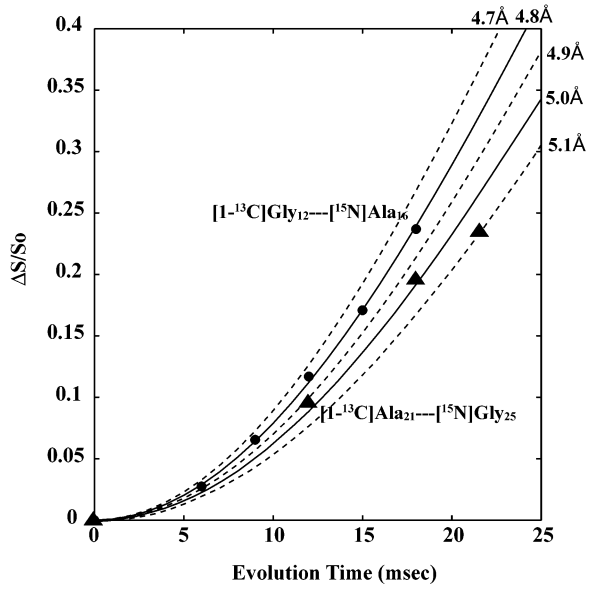


Figure 8. Observed REDOR plots, $\Delta S/S_0 = (1 - S/S_0)$ vs evolution time (ms), for **6**;GGAGGGYGGDGG(A)$_8$[1-^{13}C]A$_{21}$AAA[^{15}N]G$_{25}$GAGDGYGAG (circles) and **7**;GGAGGGYGGDG[1-^{13}C]G$_{12}$AAA[^{15}N]A$_{16}$(A)$_8$-GGAGDGYGAG (triangles). ΔS and S_0 are the REDOR amplitudes and full echo amplitudes, respectively. Solid and dashed lines show the calculated dephasing curves corresponding to the designated distances.

^{13}C]A$_{21}$AAA[^{15}N]G$_{25}$GAGDGYGAG (**6**). Figure 8 shows the experimental ^{13}C REDOR difference curve $\Delta S/S_0$ for the ^{13}C,^{15}N-labeled sample, **6**; continuous and broken lines show theoretical dephasing curve corresponding to the indicated distances. By comparing the REDOR data with the theoretical dephasing curve, the ^{13}C$-^{15}$N interatomic distance was determined as 4.8 Å. The broken lines show that the experimental error is ±0.1 Å. The distance determined here gives us the information for the torsion angles of three Ala residues, Ala$_{22}$, Ala$_{23}$, and Ala$_{24}$. If the three Ala residues, Ala$_{22}$, Ala$_{23}$, and Ala$_{24}$, are in the α-helix conformation with the same torsion angle (ϕ, ψ) = ($-59°$,$-48°$), the distance between [1-^{13}C]A$_{21}$ and [^{15}N]G$_{25}$ atoms should be 4.1 Å which is considerably different from the observed distance, 4.8 Å. Letting these three Ala residues at C-terminal side of the poly-Ala chain take either (ϕ, ψ) = ($-59°$,$-48°$) or (ϕ, ψ) = ($-66°$,$-22°$) provides a range of ^{13}C$-^{15}$N distances, which are listed in Table 2. When the torsion angles of Ala$_{22}$ residue are (ϕ, ψ) = ($-59°$, $-48°$) and those of the other two Ala residues, Ala$_{23}$ and Ala$_{24}$,are (ϕ, ψ) = ($-66°$, $-22°$), the distance between [1-^{13}C]A$_{21}$ and [^{15}N]-G$_{25}$ atoms was calculated to be 4.8 Å which is in agreement with the observed distance. Thus, the REDOR experiment is also effective to determine the precise structure of the helix.

Table 2. Calculated ^{13}C$-^{15}$N Interatomic Distance between [1-^{13}C]Ala$_{21}$ and [^{15}N]Gly$_{25}$ Atoms by Assuming that Ala Residues at C-Terminal Side in Poly-Ala Chain Take Either (ϕ, ψ) = ($-59°$, $-48°$) or ($-66°$, $-22°$)

Ala$_{22}$	Ala$_{23}$	Ala$_{24}$	distance (Å)
($-66°$, $-22°$)	($-66°$, $-22°$)	($-66°$, $-22°$)	5.4
($-66°$, $-22°$)	($-66°$, $-22°$)	($-66°$, $-22°$)	5.4
($-59°$, $-48°$)	($-66°$, $-22°$)	($-66°$, $-22°$)	4.8
($-59°$, $-48°$)	($-59°$, $-48°$)	($-66°$, $-22°$)	4.1
($-59°$, $-48°$)	($-59°$, $-48°$)	($-59°$, $-48°$)	4.1

Table 3. Calculated ^{13}C$-^{15}$N Interatomic Distance between [1-^{13}C]Gly$_{12}$ and [^{15}N]Ala$_{16}$ Atoms by Assuming that Ala Residues at N-Terminal side in Poly-Ala Chain Take Either (ϕ, ψ) = ($-59°$, $-48°$), ($-70°$, $-30°$), or ($-80°$, $-33°$)

Ala$_{13}$	Ala$_{14}$	Ala$_{15}$	distance (Å)
($-70°$, $-30°$)	($-59°$, $-48°$)	($-59°$, $-48°$)	4.6
($-70°$, $-30°$)	($-70°$, $-30°$)	($-59°$, $-48°$)	5.0
($-70°$, $-30°$)	($-70°$, $-30°$)	($-70°$, $-30°$)	4.6
($-80°$, $-33°$)	($-59°$, $-48°$)	($-59°$, $-48°$)	4.6
($-80°$, $-33°$)	($-80°$, $-33°$)	($-59°$, $-48°$)	4.6
($-80°$, $-33°$)	($-80°$, $-33°$)	($-80°$, $-33°$)	4.0

4. Determination of the Helical Structure at the N-Terminal Side of the Polyalanine Region. The REDOR experiment was also applied to determine the helical structure at the N-terminal side of the poly-Ala chain. The distance of [1-^{13}C]Gly$_{12}$ and [^{15}N]Ala$_{16}$ atoms in **7**;GGAGGGYGGDG[1-^{13}C]G$_{12}$AAA[^{15}N]A$_{16}$(A)$_8$GGAGDGYGAG, was observed with REDOR and determined to be 5.0 ± 0.2 Å. Results from the spin diffusion experiment described above provide torsion angles for the Ala$_{13}$ residue equal to either (ϕ, ψ) = ($-70°$, $-30°$) or (ϕ, ψ) = ($-80°$, $-33°$). If both Ala residues, Ala$_{14}$ and Ala$_{15}$ are in α-helix conformation with the same torsion angle (ϕ, ψ) = ($-59°$, $-48°$), the distance between [1-^{13}C]Gly$_{12}$ and [^{15}N]-Ala$_{16}$ atoms should be 4.6 Å for either case of Ala$_{13}$ (ϕ, ψ) = ($-70°$, $-30°$) or ($-80°$, $-33°$). ^{13}C$-^{15}$N distances were calculated by letting the Ala residues Ala$_{14}$ and Ala$_{15}$ take either (ϕ, ψ) = ($-59°$, $-48°$) or (ϕ, ψ) = ($-70°$, $-30°$). The distances are summarized in Table 3. Similarly, when these Ala residues take either (ϕ, ψ) = ($-59°$, $-48°$) or (ϕ, ψ) = ($-80°$, $-33°$), the calculated distances are also listed in Table 3. Only the combination of Ala$_{14}$;(ϕ, ψ) = ($-70°$, $-30°$) and Ala$_{15}$; (ϕ, ψ) = ($-59°$, $-48°$) satisfy the observed distance of 5.0 Å.

Discussion

The primary structure of *S. c. ricini* silk fibroin mainly consists of the 100-fold repeated sequences where there are alternating poly-Ala regions and Gly-rich regions, like spider

(major ampullate) silk. However, the length of the poly-Ala region is longer than that found in spider silk. The main characteristics of the silk structure seem to come from the repeated poly-Ala regions, and therefore it seems important to select the model peptide systems containing the $(Ala)_{12}$ region. Thus, we selected the peptide with the following sequence, $GGAGGGYGGDGG(A)_{12}GGAGDGYGAG$ as a typical sequence of *S. c. ricini* silk fibroin for the structural analysis.

Solid-state NMR methods were used to determine the precise structure because it was too difficult to obtain a single crystal of the peptide suitable for X-ray diffraction experiments. Solution NMR was not performed since an appropriate solvent was not found. A combination of 2D spin-diffusion ^{13}C solid-state NMR under OMAS and a quantitative use of ^{13}C CP/MAS NMR chemical shift with conformation-dependent ^{13}C chemical shift contour plots was successfully used to determine the torsion angles of the N-terminal and the central Ala residues in the $(Ala)_{12}$ chain and of two Gly residues adjacent to the $(Ala)_{12}$ chain. Rotational-echo double-resonance NMR was also used to determine the helical structures of two Ala residues at the N-terminal site and three Ala residues at the C-terminal site of the $(Ala)_{12}$ chain. Thus, the torsion angles of Gly_{12}, Ala_{13}, Ala_{19}, and Gly_{25} together with Ala_{14}, Ala_{15}, and $Ala_{22}-Ala_{24}$ in the model peptide, $GGAGGGYGGDGG(A)_{12}GGAGDGYGAG$, could be determined by a combination of advanced solid-state NMR methods. The torsion angles $(\phi, \psi) = (-59°, -48°)$ determined for Ala_{19} seem typical torsion angles of α-helix structure ($\phi = -57°$ and $\psi = -47°$).[42] On the other hand, the ψ angles in $(\phi, \psi) = (-66°, -22°)$ for Gly_{25} and $(\phi, \psi) = (-70°, -30°)$ for Ala_{13} are closer to those of 3_{10}-helical structure $(\phi, \psi) = (-49°, -26°)$ rather than α-helix structure.

The precise structure of the model peptide determined by this work is shown in Figure 9. As shown in the left side, the local structure of N- and C-terminal residues of the α-helical poly-Ala chain in the model peptide is more strongly wound than that found for a typical α-helix. The torsion angles and hydrogen-bonding patterns are summarized in Table 4. At the terminals of the helical region, five residues, Gly_{12}, Ala_{21}, Ala_{22}, Ala_{23}, and Ala_{24} contribute to the formation of $i \rightarrow i + 3$ hydrogen bonding (The right side of Figure 9). Thus, this structure will stabilize the α-helix of the poly-Ala region and prevent the structural transition to a β-sheet. It appears that there is a mechanism that prevents the structural change from α-helix to β-sheet in the silkworm.

There are highly repeating primary structures not only in the poly-Ala region but also in the Gly-rich region.[29] Interestingly, more than 65% of the Gly-rich regions have the DGGA sequence repeatedly adjacent to the N-terminal side of poly-Ala chain. Also on the C-terminal side of the poly-Ala region, the GGA sequences constitute approximately 75% of Gly-rich regions. It is interesting to point out that such Gly-Gly sequences appear in the 3_{10}-helical structure. For example, in Leu-enkephalin, the peptide backbone is held together by $i \rightarrow i + 3$ hydrogen bonds resulting in partial 3_{10}-helical structure for the Gly-Gly-Phe sequence in the solid state.[43] In addition, it is well-known that the Gly-Gly segment can easily form a tightly wound helical structure rather than an α-helix.[44−46]

Moreover, it is noted that approximately 80% of Asp residues in *S. c. ricini* silk fibroin are located at the N-terminus of the poly-Ala region. The Asp residue is a good helix-capping motif,

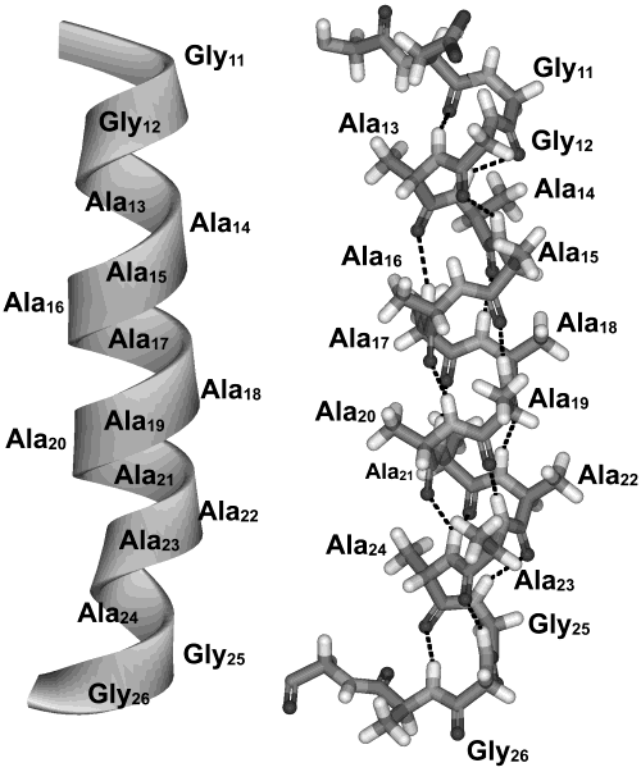


Figure 9. The structure of polyalanine region of the model peptide, $GGAGGGYGGDGG(A)_{12}GGAGDGYGAG$ of polyalanine region of *S. c. ricini* silk fibroin before spinning. Both structures are the same, but different presentation. The left side presentation shows that α-helix structure of polyalanine region tend to be wound strongly at the both terminal ends. The right side presentation shows the corresponding intramolecular hydrogen-bonding pattern by broken lines.

Table 4. Torsion Angles and Hydrogen Bonding of Each Amino Acid Residue of Model Peptide from *S. c. ricini* Silk Fibroin

	torsion angles (deg)		
residue	ϕ	ψ	hydrogen-bonding pattern
Gly_{12}	−70	−30	$i, i+3$
Ala_{13}	−70	−30	$i, i+4$
Ala_{14}	−70	−30	$i, i+4$
Ala_{15}	−59	−48	$i, i+4$
Ala_{16}	−59	−48	$i, i+4$
Ala_{17}	−59	−48	$i, i+4$
Ala_{18}	−59	−48	$i, i+4$
Ala_{19}	−59	−48	$i, i+4$
Ala_{20}	−59	−48	$i, i+4$
Ala_{21}	−59	−48	$i, i+3$
Ala_{22}	−59	−48	$i, i+3$
Ala_{23}	−66	−22	$i, i+3$
Ala_{24}	−66	−22	$i, i+3$
Gly_{25}	−66	−22	−

(43) Aubry, A.; Birlirakis, N.; Sakarellos-Daitsiotis, M.; Sakarellos, C.; Marraud, M. *Biopolymers* **1989**, *28*, 27−40.

(44) van Beek, J. D.; Kummerlen, J.; Vollrath, F.; Meier, B. H. *Int. J. Biol. Macromol.* **1999**, *24*, 173−178.

(45) Hayashi, C. Y.; Shipley, N. H.; Lewis, R. V. *Int. J. Biol. Macromol.* **1999**, *24*, 271−275.

(46) Karle, I. L.; Banerjee, A.; Bhattacharjya, S.; Balaram, P. *Biopolymers* **1996**, *38*, 515−526.

in particular the N-cap motif in globular proteins.[47,48] The helix-capping motifs[49] are specific patterns of hydrogen bonding and hydrophobic interactions found at or near the ends of helices in both globular proteins and peptides. In an α-helix, the first four amino groups lack intramolecular hydrogen bond formation. Instead, such groups are often capped by alternative hydrogen bond partners.

Acknowledgment. T.A. thanks Professor Terry Gullion (West Virginia University) for discussion and acknowledges the financial support of the Grant from the Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan.

JA0300721

(47) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648−1652.
(48) Wan, W. Y.; Milner-White, E. J. *J. Mol. Biol.* **1999**, *286*, 1651−1662.
(49) Penel, S.; Hughes, E.; Doig, A. J. *J. Mol. Biol.* **1999**, *287*, 127−143.