Structure of *Bombyx mori* Silk Fibroin Based on Solid-State NMR Orientational Constraints and Fiber Diffraction Unit Cell Parameters

Makoto Demura,[†] Masashi Minami,[†] Tetsuo Asakura,*[,†] and T. A. Cross[‡]

Contribution from the Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184, Japan, Department of Chemistry & National High Magnetic Field Laboratory, Florida State University, Tallahassee, Florida 32306

Received June 23, 1997. Revised Manuscript Received October 16, 1997

Abstract: Isotopic labeling of *Bombyx mori* silk fibroin was achieved biosynthetically using two approaches. First, labeled fibroin was achieved by feeding silk worms [$^{13}C_1$]Gly and [$^{13}C_1$]Ala along with an artificial diet. Second, in vitro production of [^{15}N]Gly and [^{15}N]Ala labeled *B. mori* silk fibroin was accomplished by culturing the posterior silk glands isolated from five-day-old silkworm larvae in the fifth instar stage. Orientation-dependent ^{15}N and ^{13}C solid-state NMR spectra of these isotope-labeled silk fibroin fibers were observed when the fiber axis was arranged at various angles between 0° and 90° to the magnetic field direction. Such data from the silk II structure was simulated on the basis of the chemical shift anisotropy to determine the Euler angles that relate the principal axis system (PAS) to the fiber axis coordinate system. The dipolar modulated ^{15}N and ^{13}C chemical shift powder pattern spectra of $^{13}C-^{15}N$ double-labeled silk fibroin model peptides were observed and simulated to determine the Euler angles for transforming the PAS to molecular symmetry axis (MSA) system. The specific orientations of N−H, N−C_1, C_1═O, and C_1−N bonds for Ala and Gly residues in the oriented silk fibroin fiber were determined with a combination of these Euler angles. The conformational space for the Ala and Gly residues of the silk fibroin fibers was substantially reduced with these bond orientations and the known $C_{\alpha(i-1)}-C_{\alpha(i+1)}$ vector orientation from fiber diffraction studies. The best fit torsion angles (ϕ,ψ) within the reduced conformational space were determined as (−140°, 142°) and (−139°, 135°), respectively within experimental error (±5°). The distance of the unit cell length determined here results in excellent agreement with the fiber diffraction data.

Introduction

Solid-state NMR is an approach that is well suited for the study of structure and dynamics of macromolecules in anisotropic environments. Such studies can take advantage of highly oriented systems which are immobile about each of three orthogonal axes on NMR time scales, such as fibers[1−8] or immobile about one or two axes, such as proteins or peptides that are embedded in a lipid environment.[9−19] In these oriented systems, the orientation-dependent nuclear spin interaction tensors serve as probes from which the orientation of specific bond vectors can be determined with respect to both the laboratory frame of reference and a unique molecular axis system aligned to the laboratory frame. The relative orientation of structural elements separated by torsional degrees of freedom can be determined from such constraints. The specific site constraints are achieved through selective isotope labeling.

Silk fibroin from *B. mori* silkworm is a fibrous protein whose primary structure is dominated by a repeating sequence of six residues, (Gly-Ala-Gly-Ala-Gly-Ser).[20−22] Two crystalline forms, silk I and silk II, have been reported as the dimorphs of silk fibroin from *B. mori* in the solid state on the basis of X-ray fiber diffraction,[23−25] electron diffraction,[26] conformational

* Address correspondence to this author.
[†] Tokyo University of Agriculture and Technology.
[‡] Florida State University.

(1) Brandolini, A. J.; Dybowski, D. *High-Resolution NMR of Synthetic Polymers in Bulk*, Komoroski, R. A., Ed.; VCH Publishers: Deerfield Beach, FL, 1986; pp 283−306.
(2) Schmidt-Rohr, K.; Spiess, H. W. *Multidimensional Solid-State NMR and Polymers*; Academic Press: London, 1994.
(3) Burwell, D. A.; Valentine, K. G.; Timmermans, J. H.; Thompson, M. E. *J. Am. Chem. Soc.* **1992**, *114*, 4144−4150.
(4) Asakura, T.; Konakazawa, T.; Demura, M.; Ito, T.; Maruhashi, Y. *Polymer* **1996**, *37*, 1965−1973.
(5) Yeo, J.-H.; Demura, M.; Asakura, T.; Fujito, T.; Imanari, M.; Nicholson, L. K.; Cross, T. A. *Solid State NMR* **1994**, *3*, 209−218.
(6) Asakura, T.; Yeo, J.-H.; Ando, I. *Polym. J.* **1994**, *26*, 229−233.
(7) Yeo, J.-H. Yeo; Asakura, T.; Shimazaki, H. *Makromol. Chem. Phys.* **1994**, *195*, 1423−1431.
(8) Schaefer, D. J.; Schadt, R. J.; Gardner, K. H.; Gabara, V.; Allen, S. R.; English, A. D. *Macromolecules* **1995**, *28*, 1152−1158.
(9) Sanders, C. R.; Landis, G. C. *J. Am. Chem. Soc.* **1994**, *116*, 6470−6471.
(10) Salvatore, B. A.; Ghose, R. Prestegard, J. H. *J. Am. Chem. Soc.* **1996**, *118*, 4001−4008.
(11) Tycko, R., Ed. *Nuclear Magnetic Resonance Probes of Molecular Dynamics*; Kluwer Academic Publishers: Dordrecht, 1994.
(12) Smith, R.; Separovic, F.; Milne, T. J.; Whittaker, A.; Bennett, F. M.; Cornell, B. A.; Makriyannis, A. *J. Mol. Biol.* **1994**, *241*, 456−466.
(13) North, C. L.; Barranger-Mathys, M.; Cafiso, D. S. *Biophys. J.* **1995**, *69*, 2392−2397.
(14) Ulrich, A. S.; Watts, A.; Wallat, I. Heyn, M. P. *Biochemistry* **1994**, *33*, 5370−5375.
(15) Cross, T. A. *Annu. Rep. NMR Spectrosc.* **1994**, *29*, 123−167.
(16) Ketchem, R. R.; Hu, W.; Cross, T. A. *Science* **1993**, *261*, 1457−1460.
(17) Teng, Q.; Nicholson, L. K.; Cross, T. A. *J. Mol. Biol.* **1991**, *218*, 607−619.
(18) Mai, W.; Hu, W.; Wang, C.; Cross, T. A. *Protein Sci.* **1993**, *2*, 532−542.
(19) Hu, W.; Lee, K.-C.; Cross, T. A. *Biochemistry* **1993**, *32*, 7035−7047.
(20) Asakura, T.; Watanabe, Y.; Itoh, T. *Macromolecules* **1984**, *17*, 2421−2426.
(21) Asakura, T.; Yoshimizu, H.; Yoshizawa, Y. *Macromolecules* **1988**, *21*, 2038−2041.

S0002-7863(97)02069-6 CCC: $15.00
Published on Web 01/27/1998

energy calculations,[27,28] infrared,[24,29] and ^{13}C and ^{15}N cross-polarization magic angle spinning (CPMAS) NMR spectroscopy.[30−34] The protein exists primarily in either the silk I or random coil forms in the silk glands of the silkworm, and undergoes a conformational transition to the silk II form during the spinning process.[35] The silk II conformation has been characterized by several X-ray fiber diffraction studies as a regular array of antiparallel β-sheets: In 1955, Marsh et al.[23] reported a fiber diffraction study of native *B. mori* silk fibroin based on the quantitative intensities of six equatorial reflections. The agreement between the observed and calculated structure factors generated by an *R* factor of 37%. Fraser et al.,[24] supported the general features of this earlier model for the silk II structure. Takahashi et al.[25] reported a more detailed X-ray fiber diffraction analysis of the silk II structure of *B. mori* silk fibroin fiber with 35 quantitative intensities. Although this and previous studies have yielded a general pattern of backbone folding by model fitting the observed cell dimensions to a regular β-sheet structure, a high-resolution structure of the silk II conformation has not been determined. Solid-state NMR is suited for the direct fiber structural characterization by observation of specific amino acid residues in the oriented silk fibroin fiber.[36,37] The repeating sequence that abounds in silk fibroin simplifies the interaction of the NMR data obtained for this protein, since a majority of the observed intensity from isotopically enriched Gly and Ala residues will arise from chemically and structurally equivalent sites.[39,41]

Fujiwara et al.[40] applied this solid-state NMR approach to oriented silk fibroin from *B. mori* labeled with [$^{13}C_1$]Ala, in which ^{13}C chemical shift line shapes obtained from fibers placed parallel and perpendicular to the magnetic field were simulated to yield the Euler angles relating the ^{13}C principal axis system to the macroscopic fiber axis. They found that the Euler angles that best simulated the observed line shapes placed the normal to the peptide plane containing the Ala carbonyl group at a smaller angle (60°) to the fiber axis than would typically be expected for a typical β-pleated sheet, raising questions concerning the accuracy of the current models for silk II structure.

The solid-state ^{15}N NMR spectra of similar *B. mori* silk fiber orientations have been reported and analyzed to yield orientations of specific molecular bonds.[41] The NMR-defined and fiber diffraction-modeled angles for the NH and NC bond vectors with respect to the fiber axis agree well for the ^{15}N Gly site. The orientation-dependent $^{13}C-^{15}N$ dipolar coupling was also observed for oriented silk fibers with the double-labeled [$^{13}C_1$]-Gly-[^{15}N]Ala peptide bond.[38] For the preparation of this sample the middle silk glands from the silkworm were cultivated in amino acid medium with both [$^{13}C_1$]Gly and [^{15}N]Ala amino acids. The dipolar splitting indicated that the angle between the $^{15}N-^{13}C_1$ peptide bond and the macroscopic fiber axis was 39°, 141°, 76°, or 104°. Among these angles, 141° is in agreement with the X-ray diffraction models ($\theta_{NC} = 139.5°$ by Marsh et al.,[23] and 140° by Takahashi et al.[25]).

In this paper, the ϕ, ψ torsion angles of Ala and Gly residues of *B. mori* silk fibroin will be determined with bond orientations obtained from ^{15}N and ^{13}C solid-state NMR spectroscopy and constraints from fiber diffraction unit cell parameters. For this purpose, [^{15}N]Ala, [^{15}N]Gly, [$^{13}C_1$]Ala, and [$^{13}C_1$]Gly sites in silk fibroin were isotopically labeled to high enrichment by oral administration of labeled amino acids to silkworm larvae or by the cultivation of the silk glands in a medium containing labeled amino acids. The specific NH and NC_1 bond orientations relative to the fiber axis were determined for both Ala and Gly residues from ^{15}N spectra of the aligned labeled silk fibers. Similarly, the C_1O and C_1N bond orientations were determined for both Ala and Gly residues from ^{13}C spectra of labeled silk fibers. The backbone torsion angles were calculated by the combination of these bond orientations for both Ala and Gly residues along with the most reliable constraints from X-ray fiber diffraction, the orientation of the major axis of the unit cell relative to the fiber axis.

Theory

The chemical shift anisotropy (CSA) interaction for ^{15}N and $^{13}C_1$ nuclei in a peptide plane can be interpreted with chemical shift tensor transformation as shown in Figure 1.[41] The ^{15}N and $^{13}C_1$ CSA principal axis systems (PAS) are frames in which the CSA tensors are diagonal, with principal components $\sigma_{11} \leq \sigma_{22} \leq \sigma_{33}$ (taking into account the negative gyromagnetic ratio of ^{15}N). The Euler angles that relate the PAS and MSA frames are denoted as α_D and β_D, where α_{DNC} and β_{DNC} are the Euler angles relating to the ^{15}N PAS and α_{DCN} and β_{DCN} relate to the $^{13}C_1$ PAS. Inherent in the orientational characterization presented in this effort is the assumption that the ^{15}N σ_{33} and $^{13}C_1$ σ_{11} tensor elements lie in the peptide plane resulting in $\gamma_D = 0°$.[49−52] The fiber axis system (FAS) is a reference frame fixed in the aligned sample and is defined such that the

(22) Asakura, T.; Kaplan, D. L. *Encyclopedia of Agricultural Science*; Arutzen, C. J., Ed.; Academic Press: New York, 1994; Vol. 4, pp 1−11.
(23) Marsh, R. E.; Corey, R. B.; Pauling, L. *Biochim. Biophys. Acta* **1955**, *16*, 1−34.
(24) Fraser, R. D. B.; MacRae, T. P. *Conformations of Fibrous Proteins and Related Synthetic Polypeptides*; Academic Press: New York, 1973.
(25) Takahashi, Y.; Gehoh, M.; Yuzuriha, K. *J. Polym. Sci., Polym. Phys. Ed.* **1991**, *29*, 889−891.
(26) Lotz, B.; Keith, H. D. *J. Mol. Biol.* **1971**, *61*, 201−215.
(27) Lotz, B.; Cesari, F. C. *Biochimie* **1979**, *61*, 205−214.
(28) Fossey, S. A.; Nemethy, G.; Gibson, K. D.; Scheraga, H. A. *Biopolymers* **1991**, *31*, 1529−1541.
(29) Asakura, T.; Kuzuhara, A.; Tabeta, R.; Saito, H. *Macromolecules* **1985**, *18*, 1841−1845.
(30) Saito, H.; Iwanaga, Y.; Tabeta., R.; Narita, M.; Asakura, T. *Chem. Lett.* **1983**, 427−430.
(31) Saito, H.; Tabeta, R.; Asakura, T.; Iwanaga, Y.; Shoji, A.; Ozaki, T.; Ando, I. *Macromolecules* **1984**, *17*, 1405−1412.
(32) Saito, H.; Ishida, M.; Yokoi, M.; Asakura, T. *Macromolecules* **1990**, *23*, 83−87.
(33) Ishida, M.; Asakura, T.; Yokoi, M.; Saito, H. *Macromolecules* **1990**, *23*, 88−94.
(34) Asakura, T.; Demura, M.; Date, T.; Miyashita, M.; Ogawa, K.; Williamson, M. P. *Biopolymers* **1997**, *41*, 193−203.
(35) Kerkam, K.; Viney, C.; Kaplan, D.; Lombardi, S. *Nature (London)* **1991**, *349*, 596−598.
(36) Asakura, T.; Yeo, J.-H.; Demura, M.; Itoh, T.; Fujito, T.; Imanari, M.; Nicholson, L. K.; Cross, T. A. *Macromolecules* **1993**, *26*, 6660−6663.
(37) Asakura, T.; Minami, M.; Shimada, R.; Demura, M.; Osanai, M.; Fujito, T.; Imanari, M.; Ulrich, A. S. *Macromolecules* **1997**, *30*, 2429−2435.
(38) Asakura, T.; Demura, M.; Hiraishi, Y.; Ogawa, K.; Uyama, A. *Chem. Lett.* **1994**, 2249−2252.
(39) Asakura, T.; Demura, M.; Uyama, A.; Ogawa, K.; Nicholson, L. K.; Cross, T. A. *Silks: Biology, Structure, Properties, Gentics*; ACS Symposium Series 544; American Chemical Society: Washington, DC, 1994; pp 148−154.
(40) Fujiwara, T.; Kobayashi, Y.; Kyogoku, Y.; Kataoka, K. *J. Mol. Biol.* **1986**, *187*, 137−140.
(41) Nicholson, L. K.; Asakura, T.; Demura, M.; Cross; T. A. *Biopolymers* **1993**, *33*, 847−861.
(42) Asakura, T.; Suzuki, H.; Watanabe, Y. *Macromolecules* **1983**, *16*, 1024−1026.
(43) Asakura, T.; Kawaguchi, Y.; Demura, M.; Osanai, M. *Insect Biochem.* **1988**, *18*, 531−538.
(44) Asakura, T.; Demura, M.; Nagashima, M.; Sakaguchi, R.; Osanai, M.; Ogawa, K. *Insect Biochem*. **1991**, *21*, 743−748.
(45) Asakura, T.; Sakaguchi, R.; Demura, M.; Manabe, T.; Uyama, A.; Ogawa, K.; Osanai, M. *Biotechnol. Bioeng*. **1993**, *41*, 245−252.
(46) Asakura, T.; Yamada, H.; Demura, M.; Osanai, M. *Insect Biochem*. **1990**, *20*, 261−266.

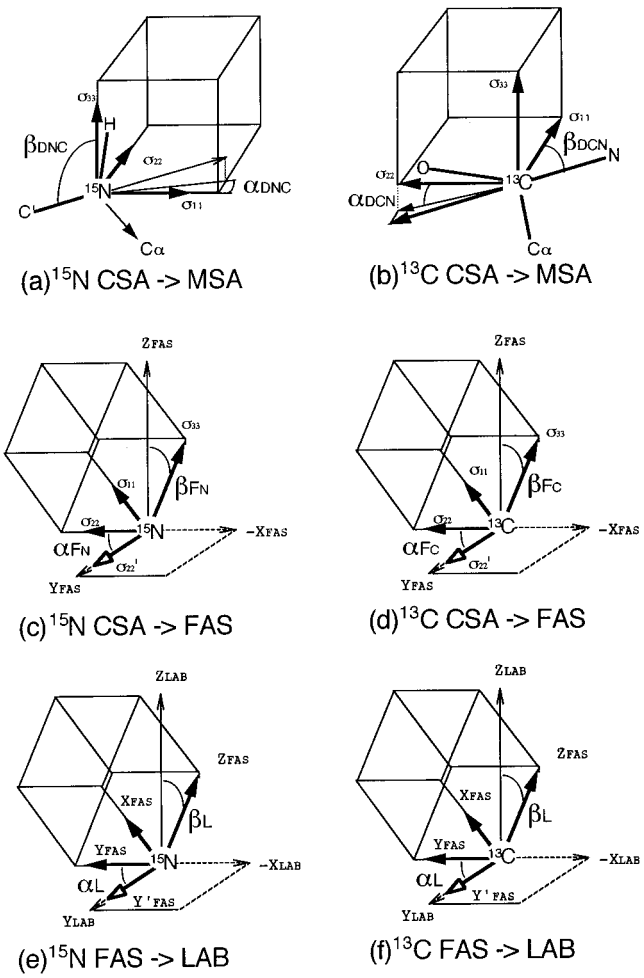

Figure 1. Transformations from the principal axis system (PAS) to the molecular symmetry axis (MSA) system of (a) ^{15}N and (b) $^{13}C_1$ sites in a peptide plane; from the PAS to the fiber axis system (FAS) of (c) ^{15}N and (d) $^{13}C_1$ sites; and from the FAS to the LAB frame of reference for ^{15}N (e) and ^{13}C (f) nuclei. The PAS is related to the FAS by the Euler angles α_{FX} and β_{FX}. This definition corresponds to setting $\gamma_{FX} = 0$.[41] Both ^{15}N and $^{13}C_1$ tensors have two tensor elements in the peptide plane as a result of the Euler angles α_{DNC} and α_{DCN} being 0°. Consequently, σ_{22} and σ_{33} are the unique elements for the ^{15}N and $^{13}C_1$ tensors that do not lie in the peptide plane, respectively.

macroscopic fiber axis lies in the Z_{FAS} direction. The PAS orientation relative to the FAS, is given by the Euler angles α_F and β_F. The locations of X_{FAS} and Y_{FAS} within the plane perpendicular to the fiber axis are defined arbitrarily such that σ_{33} lies in the XZ_{FAS} plane. These positions also yield the convenient result that $\gamma_{FN} = \gamma_{FC} = 0°$. The NMR spectra are observed in the laboratory frame of reference (LAB), in which the applied magnetic field ($\mathbf{B}_0$) is in the Z_{LAB} direction. The angles α_L and β_L are the Euler angles that transform the FAS into the LAB frame of reference. Only two angles are required for this transformation because the NMR experiment is sensitive

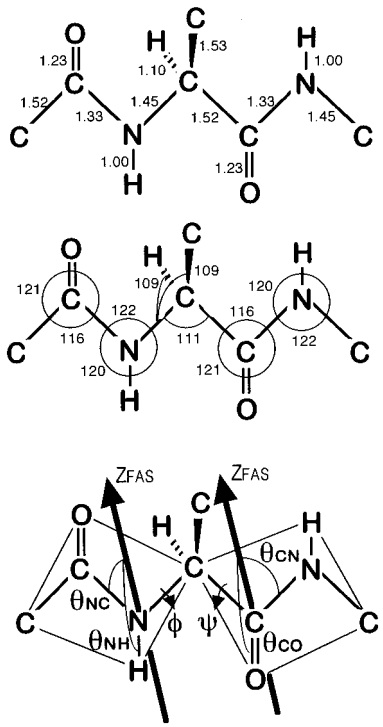

Figure 2. The bond lengths, bond angles, and definition of bond orientations, θ_{NH}, θ_{NC}, θ_{CO}, and θ_{CN} in a peptide plane with respect to fiber axis, Z_{FAS}.

only to the component of the tensor parallel to B_0. The angles α_L and β_L can be experimentally defined by the fiber orientation.

When the fiber axis is parallel to $\mathbf{B}_0$, $\alpha_L = 0$ and $\beta_L = 0$, and the observed chemical shift is given by

$$\sigma_{par} = \sigma_{11} \sin^2 \beta_F \cos^2\alpha_F + \sigma_{22} \sin^2\beta_F \sin^2\alpha_F + \sigma_{33} \cos^2\beta_F \quad (1)$$

When the fiber axis is perpendicular to $\mathbf{B}_0$, $\beta_L = 90°$, and $0° < \alpha_L < 360°$ (all values of α_L are equally represented in the spectrum). The observed chemical shift is given by[41]

$$\sigma_{per}(\alpha_F,\beta_F,\alpha_L) = F_{11} \cos^2 \alpha_L + 2F_{12} \cos \alpha_L \sin \alpha_L + F_{22} \sin^2 \alpha_L \quad (2)$$

where

$$F_{11} = \sigma_{11} \cos^2 \beta_F \cos^2 \alpha_F + \sigma_{22} \cos^2 \beta_F \sin^2 \alpha_F + \sigma_{33} \sin^2 \beta_F \quad (3A)$$

$$F_{12} = (\sigma_{22} - \sigma_{11}) \cos \beta_F \cos \alpha_F \sin \alpha_F \quad (3B)$$

$$F_{22} = \sigma_{11} \sin^2 \alpha_F + \sigma_{22} \cos^2 \alpha_F \quad (3C)$$

The F_{ij} terms are the i, j components of the CSA tensor expressed in the FAS reference frame. Equation 2 represents a family of curves that are manifested in the spectral line shape. Through spectral simulations using these equations, solutions to the α_F and β_F angles can be found.

The definition of the bond orientations, θ_{NH}, θ_{NC}, θ_{CO}, and θ_{CN} relative to the fiber axis is shown in Figure 2. To obtain the ^{15}N-based Gly and Ala bond orientation angles, θ_{NH} and θ_{NC}, the following equation is used:[41]

(47) International Union of Pure and Applied Chemistry-International Union of Biochemistry Commission on Biochemical Nomenclature, *J. Mol. Biol.* **1970**, *52*, 1.
(48) Teng, Q.; Cross, T. A. *J. Magn. Reson.* **1989**, *85*, 439−447.
(49) Harbison, G. S.; Jelinski, L. W.; Stark, R. E.; Torchia, D. A.; Herzfeld, J.; Griffin, R. G. *J. Magn. Reson.* **1984**, *60*, 79−82.
(50) Oas, T. G.; Hartzell, C. J.; Dahlquist, F. W.; Drobny, G. P. *J. Am. Chem. Soc.* **1987**, *109*, 5962−5965.
(51) Hartzell, C. J.; Whitfield, M.; Oas, T. G.; Drobny, G. P. *J. Am. Chem. Soc.* **1987**, *109*, 5966−5969.
(52) Hartzell, C. J.; Pratum, T. K.; Drobny, G. P. *J. Chem. Phys.* **1987**, *87*, 4324−4331.

$$\cos\theta_{NX} = \cos\beta_{FN}\cos\beta_{DNX} + \sin\beta_{FN}\cos\alpha_{FN}\cos\alpha_{DNX}\sin\beta_{DNX} + \sin\beta_{FN}\sin\alpha_{FN}\sin\alpha_{DNX}\sin\beta_{DNX} \quad (4)$$

where α_{DNX} and β_{DNX} (X is H or C_1) are the Euler angles as defined in Figure 1 and experimentally determined from $^{13}C_1-^{15}N$ double-labeled peptides. The α_{FN} and β_{FN} values are obtained as described above. The ^{13}C-based Gly and Ala bond orientation angles for θ_{CO} and θ_{CN} relative to the fiber axis are similarly calculated with the $^{13}C_1$ PAS to MSA system Euler angles α_{DCX} and β_{DCX} (X is O or N) as well as the ^{13}C PAS to FAS angles, α_{FC} and β_{FC}.

Experimental Section

Biosynthetic Labeling. Isotopic labeling of *B. mori* silk fibroin was achieved biosynthetically through the use of an artificial diet supplemented with [$^{13}C_1$]Gly and [$^{13}C_1$]Ala.[20,41−44] Isotopes were incorporated into cocoons of *B. mori* by feeding 20 mg of [$^{13}C_1$]-labeled amino acid ([$^{13}C_1$]Gly, 99% ICON, New York; [$^{13}C_1$]Ala, 99% ICON, New York) along with 27 g of an artificial diet (silk Mate 2M, Nippon Nosan Kogyo Co., Tokyo) to five, 6- or 7-day old silkworm larvae in the fifth instar stage. In vitro production of [^{15}N]Gly and [^{15}N]Ala enriched *B. mori* silk fibroin was performed with posterior silk glands isolated from 5-day-old silkworm larvae in the fifth instar stage.[45,46] Preparation of the culture medium containing these isotope labeled amino acids, the culture procedure, and isolation of the silk fibroin from posterior silkglands has all been described previously.[45,46]

Powder Sample Preparation. To prepare silk II powder samples, the cocoons of *B. mori* were first degummed twice with 0.5 w/v % Marseilles soap solution at 100 °C for 30 min and then washed with distilled water. The silk sericin, another silk protein, was removed from the surface of silk fibers by this degumming treatment. The silk fibroin was then dissolved in 9 M LiBr at 40 °C. After dialysis against distilled water for 4 days, the solution was clarified by centrifugation at 10 000 rpm for 20−30 min. The supernatant was collected and gently evaporated by forced air flow to a concentration of 3 w/v %, and the aqueous solution of silk fibroin was freeze-dried. Then, this powder sample was immersed in 80 v/v% methanol aqueous solution for 10 min and dried to prepare the silk II structure.[30−33]

Oriented Sample Preparation. Preparation of *B. mori* silk fibroin fiber samples was achieved by first loosening the threads on the cocoons by placing them in 100 °C water for 5 min. The ends of the cocoon fibers were located, bundled together, and wound onto a glass tube. To remove sericin, the silk fibers were treated with 0.5 w/v % Marseilles soap solution at 100 °C for 30 min while on the glass tube and then washed with distilled water. To prepare blocks of oriented fibers, the silk fibers were removed from the glass tube as a bundle and again placed in hot water (about 85 °C) for approximately 15 min to enable easy separation of the fibers from one another. The silk fibers were then rewound onto a form which produces sheets of highly oriented fibers. These sheets were fixed with quick-setting epoxy and then cut into 4.5 mm × 10 mm pieces, stacked together, and fixed with epoxy to form an 4.5 mm × 4.5 mm × 10 mm block that fit within the radio frequency coil of the NMR probe.

NMR Spectroscopy. ^{15}N solution NMR spectra were observed with a JEOL FX-90Q NMR spectrometer operating at 25 °C. The static solid-state ^{15}N and ^{13}C NMR experiments were performed at 25 °C on a JEOL GX400 spectrometer operating at 40.4 and 100.4 MHz, respectively. The probe was equipped with a 7-mm inner diameter coil. Cross polarization was employed with high-power ^{1}H decoupling during the signal acquisition period. Typical NMR parameters were 6 μs 90° pulse with a 3 ms mixing time and 7 s repetition delay for ^{15}N NMR and 5 μs 90° pulse with a 3 ms mixing time and 7 s repetition delay for ^{13}C NMR. The ^{15}N chemical shifts were referenced to $^{15}NH_4NO_3$ by setting the signal of solid $^{15}NH_4Cl$ to 18.0 ppm. The ^{13}C chemical shifts were referenced to tetramethylsilane by setting the ^{13}C signal of solid hexamehylbenzene to 17.3 ppm (CH_3).

Calculations. The bond lengths and bond angles are fixed as shown in Figure 2 (IUPAC−IUB[47]), and the ω angles of peptide planes were assumed to be 180°. The α_D and β_D angles (Figure 1) for the $^{13}C_1-^{15}N$ double-labeled model compounds were determined from the simulation of $^{13}C-^{15}N$ dipolar modulated powder pattern using the approach described by Teng and Cross.[48] The ^{15}N and ^{13}C solid-state NMR spectral line shapes as a function of β_L were simulated according to eqs 1 and 2 by varying α_F and β_F. A Gaussian probability distribution of fiber axis orientations was employed to account for the spectral broadening observed in both the parallel and perpendicular orientations.[41] An error analysis was performed by summing the difference between the experimental and calculated points in the spectral line shape to obtain the best fit of the α_F and β_F pair.

The orientation of the $C_{\alpha(i-1)}-C_1-^{15}N-C_{\alpha(i)}$ peptide plane with respect to the fiber axis was calculated using NH and NC_1 bond orientations. Similarly, the $C_{\alpha(i)}-^{13}C_1-N-C_{\alpha(i+1)}$ plane orientation was determined from the C_1O and C_1N bond orientations. An error analysis was performed by summing the difference between the experimental and calculated bond orientations. The ϕ,ψ torsion angles were determined from the relative orientation of adjacent peptide planes within the experimental error of the defined bond orientations. The minimum of this sum indicates the best fit of (ϕ,ψ) values. All calculations were performed with a Mips Fortran 77 compiler on a Unix workstation (Silicon Graphics, Inc., Indy R4400, IRIX5.3).

Results

Characterization of Euler Angles. To determine the Euler angles that relate the $^{13}C_1$ and ^{15}N PAS to the MSA system for the [$^{13}C_1$]Gly and [^{15}N]Ala sites, two peptides were synthesized by liquid-phase synthesis as described elsewhere,[53,54] as model compounds for *B.mori* silk fibroin: Boc-Gly-Ala-[$^{13}C_1$]Gly-[^{15}N]Ala-Gly-Ala-OPac and Boc-Ala-Gly-[$^{13}C_1$]Ala-[^{15}N]Gly-Ala-Gly-OPac.

In this way nearly 100% isotopic enrichment is achieved in the doubly labeled [$^{13}C_1-^{15}N$] peptide bond. The conformation of the crystalline peptides was confirmed to be an antiparallel β sheet from the isotropic ^{13}C chemical shifts of the Ala C_α and C_β sites in CPMAS spectra.[30,34] *B. mori* silk fibroin in crystalline form shows very similar chemical shifts characteristic of silk II and distinct from silk I.[29,30] Therefore, these peptides represent good models for characterizing the Euler angles that relate the PAS and MAS system in silk II. Figure 3 shows ^{13}C and ^{15}N chemical shift powder pattern spectra of the [$^{13}C_1$]Gly−[^{15}N]Ala-labeled model compound under conditions where both spectra are modified by the $^{13}C-^{15}N$ dipolar interaction. From the ^{15}N powder pattern simulation,[48,50−52] the angles that relate the ^{15}N PAS to the MSA system, α_{DNC} and β_{DNC} were determined to be 0 ± 5° and 109 ± 5°, respectively. These angles indicate that the $N-C_1$ bond is in the $\sigma_{11}-\sigma_{33}$ plane of the ^{15}N PAS. From the simulation of the ^{13}C powder pattern spectrum the angles that relate the $^{13}C_1$ PAS to the MAS system, α_{DCN} and β_{DCN} were determined to be 0 ± 5° and 35 ± 5°, respectively. These angles indicate that the $N-C_1$ bond is in the $\sigma_{11}-\sigma_{22}$ plane of the $^{13}C_1$ PAS. For the [$^{13}C_1$]Ala−[^{15}N]-Gly peptide bond similar spectra were obtained and the Euler angles were characterized. The results, summarized in Table 1, are, as expected, slightly different for the two peptide planes.

Figure 4 shows the ^{15}N solution NMR spectra of [^{15}N]Ala and [^{15}N]Gly *B. mori* silk fibroins in aqueous solution along with the spectrum of a natural abundance sample. The ^{15}N-labeled samples were obtained by the cultivation of the posterior silk glands in a medium containing [^{15}N]Gly or [^{15}N]Ala amino acids. The isotopic enrichment in the labeled samples is more than 15 times higher than the natural abundance, thereby

(53) Rich, D. H.; Sing, J. *The peptide analysis, Synthesis, Biology*; Gross, E., Meienhofer, J., Eds.; Academic Press: New York, 1979; Vol. 1, pp 241−261.

(54) Asakura, T.; Yamazaki, Y.; Seng, K. W.; Demura, M. *J. Mol. Struct.*, in press.

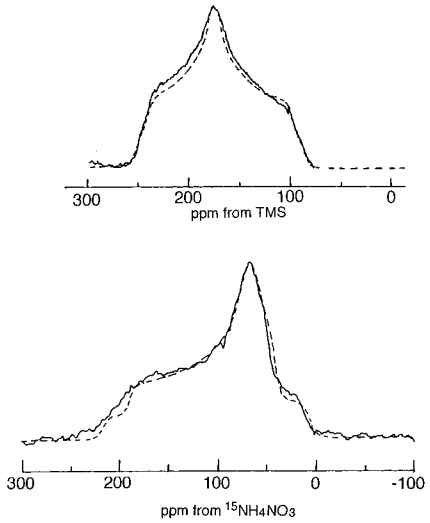

Figure 3. Solid-state ^{13}C (upper) and ^{15}N (lower) powder pattern spectra of Boc-Gly-Ala-[$^{13}C_1$]Gly-[^{15}N]Ala-Gly-Ala-OPac. The best fit theoretical line shapes are superimposed on the experimental spectra. The α_{DCN} and β_{DCN} values for the [$^{13}C_1$]Gly were determined to be $\alpha_{DCN} = 0°$, $\beta_{DCN} = 35°$. The ^{15}N powder pattern spectrum revealed $\alpha_{DNC} = 0°$ and $\beta_{DNC} = 109°$.

Table 1. Euler Angles That Relate the $^{13}C_1$ and ^{15}N CSA PAS to the MSA System for [$^{13}C_1$]Gly and [$^{13}C_1$]Ala, and [^{15}N]Gly[a] and [^{15}N]Ala Sites Determined from the ^{13}C and ^{15}N Powder Patterns of Model Peptides[b]

angle	Gly (deg)	Ala (deg)
	^{15}N CSA	
α_{DNC}	0	0
β_{DNC}	104	109
α_{DNH}	0	0
β_{DNH}	18	13
	$^{13}C_1$ CSA	
α_{DCN}	0	0
β_{DCN}	35	33
α_{DCO}	0	0
β_{DCO}	89	91

[a] Reference 41. [b] Experimental error of each Euler angle was ±5°.

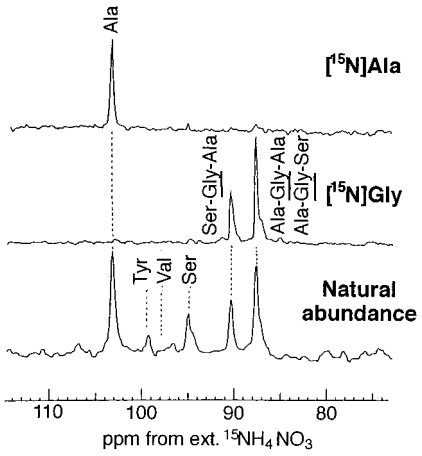

Figure 4. ^{15}N solution NMR spectra of [^{15}N]Ala and [^{15}N]Gly *B. mori* silk fibroins in aqueous solution. The natural abundance spectrum is shown for a comparison.

permitting the solid-state NMR experiments that follow. Moreover, high isotopic enrichment of [$^{13}C_1$]Gly- and [$^{13}C_1$]Ala-labeled silk fibroin was obtained by administering isotope-labeled amino acids to the silkworms as reported previously.[20]

The ^{15}N chemical shift tensor for the [^{15}N]Gly residues of silk fibroin was essentially the same as reported previously,[41] and therefore, only the [^{15}N]Ala tensor is described here. Figure 5 shows the solid-state ^{15}N NMR powder spectrum of [^{15}N]-Ala silk fibroin in the silk II form. By simulation, the ^{15}N chemical shift tensor elements (CSA) for the [^{15}N]Ala residues were determined to be $\sigma_{11} = 33$, $\sigma_{22} = 56$, and $\sigma_{33} = 195$ ppm. The ^{15}N chemical shift tensors including experimental error are summarized in Table 2.

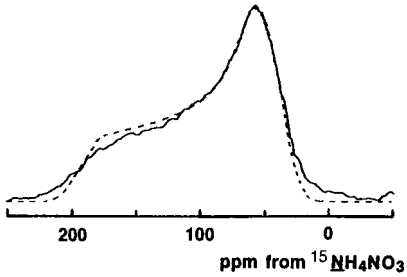

Figure 5. Solid-state ^{15}N NMR powder pattern of [^{15}N]Ala silk fibroin in the silk II form, prepared by immersing the isotope-labeled silk powder in methanol. The calculated powder pattern is also shown. The best fit chemical shift tensor element values were determined to be $\sigma_{11} = 33$ ppm, $\sigma_{22} = 56$ ppm, and $\sigma_{33} = 195$ ppm.

Table 2. ^{13}C and ^{15}N Chemical Shift Tensors for Gly and Ala Residues

	[^{15}N]Gly[a,c]	[^{15}N]Ala[a]	[$^{13}C_1$]Gly[b]	[$^{13}C_1$]Ala[b]
σ_{11}[d]	22	33	245	242
σ_{22}[d]	54	56	179	186
σ_{33}[d]	186	195	99	96

[a] Parts per million (ppm) from external $^{15}NH_4NO_3$. [b] Parts per million (ppm) from external TMS. [c] Reference 41. [d] Experimental error of each tensor element is ±3 ppm for the ^{15}N tensors and ±5 ppm for the ^{13}C tensors.

The ^{13}C solid-state NMR powder patterns of [$^{13}C_1$]Gly and [$^{13}C_1$]Ala silk fibroins are shown in Figure 6, parts a and b, respectively. The chemical shift tensors were determined by simulation to be $\sigma_{33} = 99$, $\sigma_{22} = 179$, and $\sigma_{11} = 245$ ppm for [$^{13}C_1$]Gly silk fibroin and $\sigma_{33} = 96$, $\sigma_{22} = 186$, and $\sigma_{11} = 242$ ppm for the [$^{13}C_1$]Ala silk fibroin. These simulations did not take into account the $^{14}N-^{13}C$ dipolar interaction which can be seen to be present. Consequently, the error is somewhat greater than for the ^{15}N simulations. The chemical shift tensors and estimates of experimental error are summarized in Table 2.

Figure 7 shows the ^{15}N solid-state NMR spectra of [^{15}N]Ala silk fibroin fiber block aligned with the fiber axis set parallel (a) and perpendicular (b) to the magnetic field direction, $\mathbf{B}_o$. On the basis of a 20% noncrystalline fraction in silk fibroin fibers, 20% of the spectral intensity has been subtracted as powder pattern from these spectra.[41] By simulation of the resulting spectra including the distribution of fiber axis orientations, p, the set of Euler angle solutions that relate the ^{15}N PAS to the fiber axis (Z_{FAS}), α_{FN} and β_{FN} were determined as shown in Table 3. The residual orientational distribution of fiber axes is primarily attributed to local orientational heterogeneity of the polypeptide backbone relative to the macroscopic fiber axis.

Figure 8 shows the ^{13}C solid-state NMR spectra of [$^{13}C_1$]-Gly silk fibroin block fibers as a function of fiber axis orientation (β_L) with respect to $\mathbf{B}_o$. The reliability of the structural parameters, α_{FC} and β_{FC}, given in Table 4 is high resulting from the consistent simulation of the full data set. Similar parameters are also obtained for the oriented block of [$^{13}C_1$]Ala silk fibroin as given in Table 4.

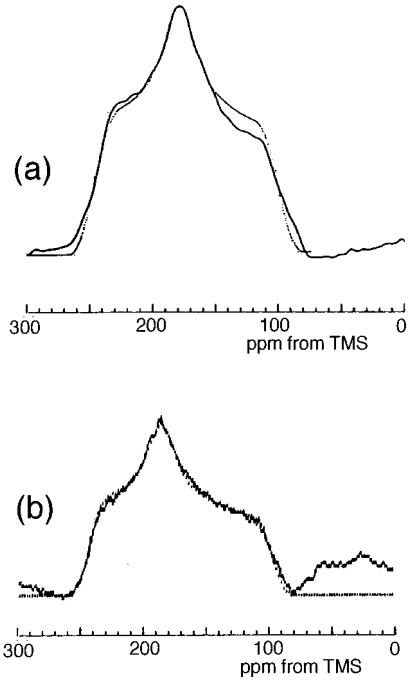

Figure 6. Solid-state ^{13}C NMR powder patterns of (a) [$^{13}C_1$]Gly and (b) [$^{13}C_1$]Ala silk fibroins. The calculated powder patterns are also shown. The chemical shift tensor elements were determined to be (a) $\sigma_{11} = 245$, $\sigma_{22} = 179$, $\sigma_{33} = 99$ ppm, and (b) $\sigma_{11} = 242$, $\sigma_{22} = 186$, $\sigma_{33} = 96$ ppm, respectively.

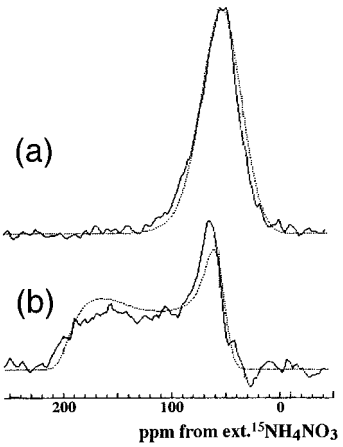

Figure 7. Solid-state ^{15}N NMR spectra of [^{15}N]Ala *B. mori* silk fibroin aligned with the fiber axis (a) parallel and (b) perpendicular to the magnetic field, $\mathbf{B}_o$. The best fit theoretical line shapes are superimposed on the experimental spectra. The distribution of fiber axis orientations, p, was determined to be $\pm 10°$. The resulting Euler angle solution set is given in Table 3.

Discussion

Unique Bond Orientations. The Euler angles that relate the $^{13}C_1$ and ^{15}N PAS to both the FAS and MSA system were determined experimentally from ^{15}N- and [$^{13}C_1$]Gly-labeled sites and ^{15}N- and [$^{13}C_1$]Ala-labeled sites in silk fibroin. The bond orientations with respect to the fiber axis, θ_{NH} and θ_{NC}, were calculated according to eq 4 with the experimentally defined Euler angles. For example, the θ_{NH} value was calculated with α_{FN}, β_{FN}, α_{DNH}, and β_{DNH}. The θ_{NH} and θ_{NC} values of [^{15}N]-Gly and [^{15}N]Ala sites are listed in Table 3. Similarly, the bond orientations θ_{CO} and θ_{CN} were calculated from experimentally defined Euler angles and listed in Table 4. Previously, it was found that eight possible orientations (eight sets of α_{FN} and β_{FN}) can be defined from spectra of fibers aligned parallel and perpendicular to $\mathbf{B}_o$. For instance, eight possible orientations of the NH bond (θ_{NH}) for the [^{15}N]Gly-labeled site relative to

Table 3. The Euler Angles, α_{FN} and β_{FN}, and Bond Orientation Data,[a] θ_{NH} and θ_{NC}, for [^{15}N]Gly[b] and [^{15}N]Ala Sites within the Silk II Structure[c]

α_{FN} (deg)	β_{FN} (deg)	θ_{NH}(deg)	θ_{NC}(deg)
	[^{15}N]Gly CSA		
25	72	88	40*
25	108	124	24
155	72	56	156
155	108	92	140
205	72	56	156
205	108	92	140
335	72	88	40*
335	108	124	24
	[^{15}N]Ala CSA		
2	70	83	39*
2	110	123	2
178	70	57	178
178	110	97	141
182	70	57	178
182	110	97	141
358	70	83	39*
358	110	123	2

[a] Calculated with α_{DNX} and β_{DNX} values listed in Table 1 using eq 4. The asterisk marks indicates the more probable solutions for each [^{15}N] site. [b] Reference 41. [c] Experimental error of each Euler angle was $\pm$ 5°.

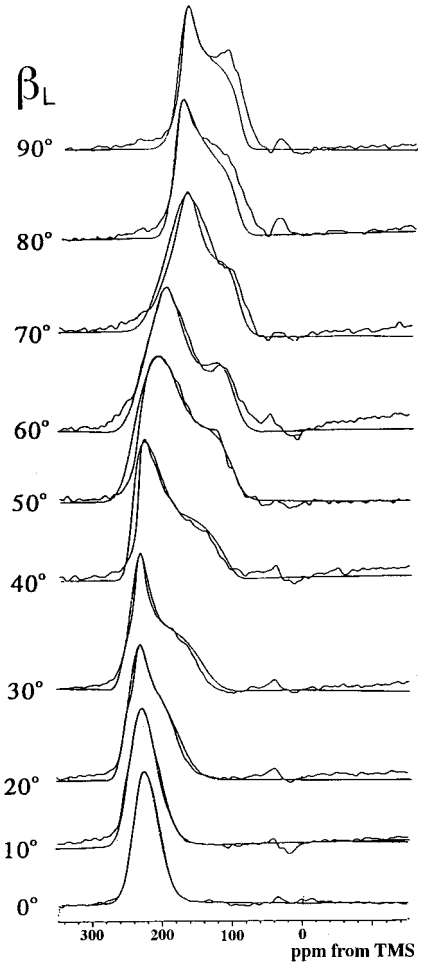

Figure 8. Solid-state ^{13}C NMR spectra of an oriented block of [$^{13}C_1$]-Gly fibroin fibers as a function of the β_L, the angle between the fiber axis and $\mathbf{B}_o$. The best fit theoretical line shapes obtained with the parameters, $\alpha_{FC} = 90°$, $\beta_{FC} = 22°$ and $p = 11°$, are superimposed on the experimental spectra.

Table 4. The Euler Angles, α_{FC} and β_{FC}, and Bond Orientation Data,[a] θ_{CO} and θ_{CN}, for [$^{13}C_1$]Gly and [$^{13}C_1$]Ala Sites within the Silk II Structure[b]

α_{FC} (deg)	β_{FC} (deg)	θ_{CO} (deg)	θ_{CN} (deg)
	[$^{13}C_1$]Gly CSA		
90	22	91	41
90	158	89	139*
270	22	91	41
270	158	89	139*
	[$^{13}C_1$]Ala CSA		
84	15	88	34
84	165	89	143*
96	15	91	37
96	165	93	146*
264	15	91	37
264	165	93	146*
276	15	88	34
276	165	89	143*

[a] Calculated with α_{DCX} and β_{DCX}. The asterisk marks indicate the more probable solutions for each [^{13}C] site. The orientations of [$^{13}C_1$]Gly CSA are eliminated to four sets because of same solutions. [b] Experimental error of each Euler angle was ±5°.

the fiber axis orientation are obtained and shown in Table 3. Previously, a combination of ^{15}N chemical shift and ^{15}N−^{1}H dipolar data were used to eliminate four of the eight orientations. Two of the remaining four were simply related to one another by a 180° rotation and therefore did not represent a different structure, but only a different orientation of the structure with respect to $\mathbf{B}_o$. On the basis of these data alone, it has not been possible to reduce the NMR-derived bond orientations to a unique solution.[41]

In this paper, we reduce the possible bond orientation angles for the silk II structure by using the experimental observation from X-ray fiber diffraction that the $C_{\alpha(i-1)}-C_{\alpha(i+1)}$ axis is parallel to the fiber axis.[25] This is based on the good assumption that the fiber diffraction unit cell is a two residue repeat unit along the peptide chain direction corresponding to a β-sheet type of structure. The bond orientations for the two peptide planes defined by $C_{\alpha(i-1)}-C_{1(i-1)}-N_{(i)}-C_{\alpha(i)}$ and $C_{\alpha(i)}-C_{1(i)}-N_{(i+1)}-C_{\alpha(i+1)}$ with respect to the $C_{\alpha(i-1)}-C_{\alpha(i+1)}$ axis can be calculated as a function of the torsion angles (ϕ,ψ) as shown in Figure 9. Within the range of ϕ,ψ angles broadly corresponding to a β-sheet type of structure the range of θ angles is limited. For the Gly residue defined by the $Ala_{(i-1)}-Gly_{(i)}-Ala_{(i+1)}$ repeat unit the only combination of θ_{NH} and θ_{NC} from Table 3 that is consistent with the range of possible bond orientations presented in Figure 9 is $\theta_{NH} = 88°$ and $\theta_{NC} = 40°$ denoted by an asterisk in Table 3. Similarly, one combination of θ_{CO} and θ_{CN} from Table 4 is consistent with the bond orientation range from Figure 9, $\theta_{CO} = 89°$ and $\theta_{CN} = 139°$. For the Ala residue defined by the $Gly_{(i-1)}-Ala_{(i)}-Gly_{(i+1)}$ repeat unit a unique θ_{NH} (83°) and θ_{NC} (39°) pair of bond orientations is defined, while an ambiguity for the θ_{CO} (89° or 92°) and θ_{CN} (142° or 145°) remains. However, this latter pair of solutions is within the experimental error (±5°) of each solution. These results are consistent with the $^{13}C_1-^{15}N$ dipolar data previously reported[38] that defined two possible orientations for the θ_{NC} of the Ala residue, 39° and 141°. In addition, this dipolar data also confirms the θ_{NC} value (141°) for the Gly residue.

Unique Torsion Angles. Each unique bond orientation constrains the ϕ,ψ torsional solution set. In Figure 10 regions of the ϕ,ψ torsional space consistent with specific bond orientations and their error range are shown. The overlap region from the four bond orientations for each residue is expanded in Figure 11 showing contours that represent the sum of observed and predicted bond orientation differences. The best fit solution

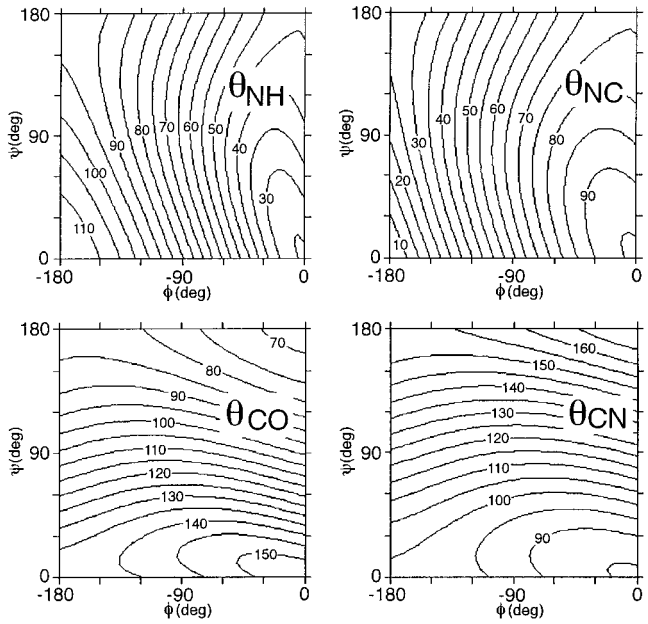

Figure 9. Variation in bond orientations, θ_{NH}, θ_{NC}, θ_{CO}, and θ_{CN}, as a function of torsion angles (ϕ, ψ) when the $C_{\alpha(i-1)}-C_{\alpha(i+1)}$ axis is constrained to be parallel to the FAS. The ϕ,ψ range shown is restricted to the generalized β-sheet region of the Rhamachandran diagram.

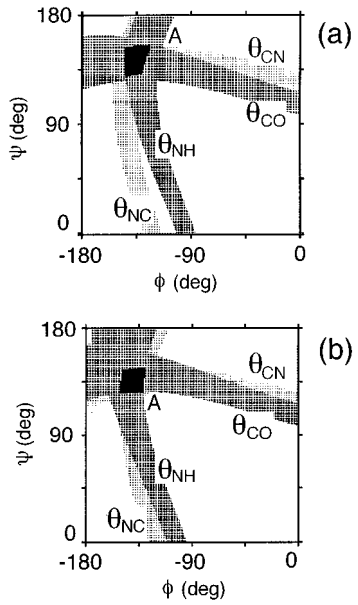

Figure 10. Variation in bond orientations constrained by the $C_{\alpha(i-1)}-C_{\alpha(i+1)}$ axis being parallel to the FAS and by the NMR orientational constraints for the (a) Ala and (b) Gly residues. The conformational space is restricted to an experimental error of ±5° for each bond orientation. The region of overlap is expanded in Figure 11.

based on minimum bond orientation difference for Ala is $\phi = -140°$ and $\psi = 142°$ (Figure 12). Similarly, for glycine the best fit is $\phi = -139°$ and $\psi = 135°$. The torsion angles for Ala and Gly residues in this structural determination and previous silk II structural models are summarized in Table 5. For the structural characterization here and for the other models except for Takahashi's, the ω torsion angles have been assumed to be 180°. In addition to the crystallographic c axis being characterized as parallel to the fiber axis (Z_{FAS}), the unit cell length has been rather precisely described by the fiber diffraction reports as 6.94 Å[23] and 6.98 Å[25] while energy-based modeling has resulted in a unit cell length of 7.06 Å. The dipeptide repeat distance determined here is 6.98 Å based on the Ala results

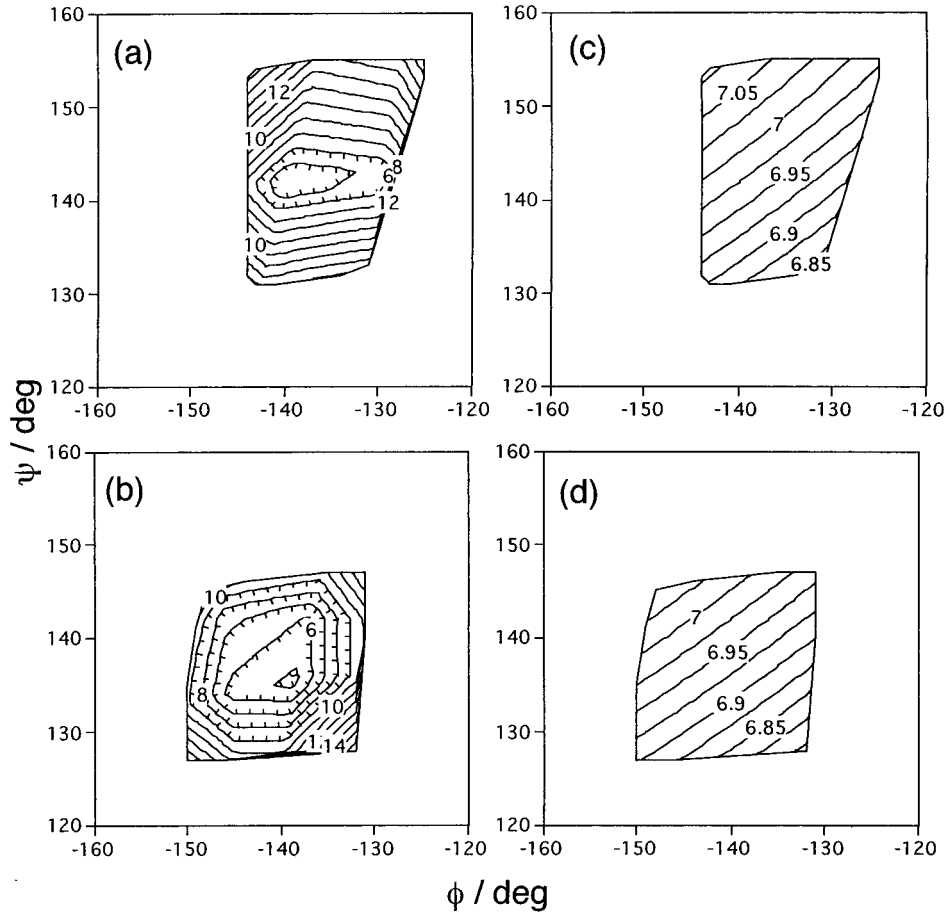

Figure 11. Comparison of the restricted ϕ,ψ conformational space for the (a and c) Ala and (b and d) Gly residues. Panels a and b indicate sum of four kinds of errors between observed and calculated bond orientations, θ_{NH}, θ_{NC}, θ_{CO}, and θ_{CN}. Panels c and d indicate the $C_{\alpha(i-1)}-C_{\alpha(i+1)}$ distance.

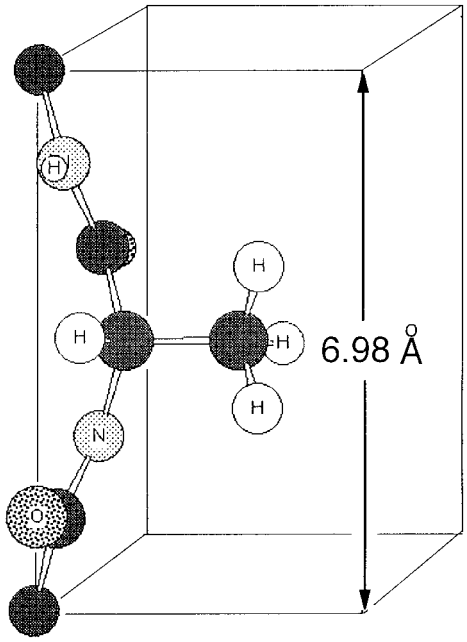

Figure 12. The best fit structure for the Ala residue in the silk II structure. ϕ and ψ values are (−140°, 142°).

and 6.92 Å based on the Gly results, in excellent agreement with the fiber diffraction data.

The data reported here is consistent with our previously published orientational constraints from dipolar interactions.[41] In addition, Fujiwara et al.[40] has observed oriented fibers of *B. mori* silk fibroin labeled with [$^{13}C_1$]Ala. These low-field, low sensitivity results were interpreted to define a somewhat distorted β-sheet structure in which the peptide planes were tipped by 30° with respect to the fiber axis. Here it is shown that the peptide plane with the Ala carbonyl carbon is tipped by only 20°, thereby avoiding the β-sheet distortions. Despite these compromising issues for the Fujiwara data, the ^{13}C constraints are in qualitative agreement with the constraints

Table 5. Torsion Angles (ϕ, ψ) of Silk II Structure of *B. mori* Silk Fibroin Determined from Solid-State NMR and Other Methods

method	ϕ (deg)	ψ (deg)	$C_{\alpha(i-1)}-C_{\alpha(i+1)}$ distance (Å)	sum of θ_{err}[d] (deg)
NMR and Fiber Diffraction				
this work				
Ala	−140	142	6.98	4.0
Gly	−139	135	6.92	4.8
X-ray Fiber Diffraction				
Marsh's model[a]			6.94[e]	
Ala	−139	140		
Gly	−139	140		
Takahashi's model[b]			6.98[e]	
Ala	−134	154		
Gly	−150	142		
Energy calculation				
Fossey's model[c]			7.06	
Ala	−149	148		
Gly	−150	146		

[a] Reference 23. [b] Reference 25. [c] Reference 28. [d] Sum of θ_{err} is the sum of four minimum errors between observed and calculated θ values for θ_{NH}, θ_{NC}, θ_{CO}, and θ_{CN} at the best fit torsion angle. [e] Unit cell length.

reported here. They did not develop their constraints into ϕ,ψ torsional angles. Similarly, the orientational constraints observed in Nicholson et al.[41] were primarily used to support extant structural models rather than to define ϕ,ψ angles.

Conclusion

For four decades the best structure of silk II has been a model β-sheet that fits the fiber diffraction data well. However, a broad range of ϕ,ψ torsion angles are consistent with the unit cell parameters from these diffraction studies. In this unique study the best of the fiber diffraction structural constraints are used to resolve the structural ambiguities from solid-state NMR-derived orientational constraints. These latter constraints can lead to high-resolution structural characterization without resorting to model fitting. Here unique ϕ,ψ torsion angles were determined for the Gly and Ala residues in the silk II structure. Many details of the β-sheet packing and noncrystalline domains have yet to be resolved, but solid-state NMR has generated new structural constraints for this long-standing challenge in structural biology.

Acknowledgment. We thank Professor L. K. Nicholson (Cornell University) for helpful discussions. T.A. acknowledges supports from Bio-oriented Technology Research Advancement Institution, Yazaki Foundation, Japan, Ciba-Geigy Foundation, Japan, Iwatani Foundation, Japan, The International Human Frontier Science Program and Ryoka System Inc., Japan. T.A.C. acknowledges support of the U.S. National Science Foundation, DMB 96-03935.

JA972069L