J|A|C|S
COMMUNICATIONS

Published on Web 07/09/2002

Heterogeneous Structure of Silk Fibers from *Bombyx mori* Resolved by ^{13}C Solid-State NMR Spectroscopy

Tetsuo Asakura,*,§ Juming Yao,§ Tsutomu Yamane,§ Kosuke Umemura,§ and Anne S. Ulrich†

Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan, and Institut für Molekularbiologie, University of Jena, Winzerlaer Strasse 10, 07754 Jena, Germany

Received February 15, 2002; Revised Manuscript Received June 14, 2002

Silk fibroin from the silkworm *Bombyx mori* is an Ala/Gly-rich protein, which is spun from aqueous solution at room temperature to produce strong and elastic fibers.[1] Solid-state NMR was used to resolve the molecular conformation of silk fibroin before and after spinning, and model peptides helped to assign various structural features. The multicomponent ^{13}C-MAS signal of the Ala side chain could thus be deconvoluted to yield a quantitative description of up to five different structural motifs at the native silk fiber.

When unprocessed silk fibroin from the silk gland is allowed to dry, it assumes a structure known as silk I, which can also be regenerated from solubilized silk by dialysis from 9 M LiBr.[2] The silk I structure is fundamentally different from that of the native silk fiber after spinning, called silk II, which may be regenerated by treatment with formic acid.[3] Because of the quasi-crystalline nature of silk I and II, only limited information is available about the detailed conformation and packing of the protein chains. Using various solid-state ^{13}C- and ^{15}N-NMR labeling schemes we have recently proposed a comprehensive model for the structure of silk I.[4]

Silk II fibers consist of both quasi-crystalline and amorphous domains. The primary structure of *Bombyx mori* silk contains multiple repeats of $(Ala\text{-}Gly\text{-}Ser\text{-}Gly\text{-}Ala\text{-}Gly)_n$ which make up 55% of the total fiber and form the insoluble quasi-crystalline Cp-fraction after chymotrypsin cleavage. The remaining sequence of fibroin is rich in Tyr and constitutes the amorphous part of the fiber. Using X-ray fiber diffraction to pick up the crystalline regions, the structure of silk II was characterized first by Marsh et al.[5] as a regular array of antiparallel β-sheets. Later, Fraser et al.,[6] Lotz and Keith,[7] and Fossey et al.[8] supported the general features of this antiparallel β-sheet model, but some of them also noted an irregular structure to be present in the silk fibers.[6,7] More recently, Takahashi et al.[9] proposed that each crystal site is statistically occupied by two antiparallel β-sheet chains with different relative orientations. The two different kinds of intersheet stacking occur at a ratio of 1:2.

The structural features of silk fibers are conveniently studied using synthetic peptides, namely $(Ala\text{-}Gly)_{15}$ as a model for the Cp fraction, and $(AG)_3YG(AG)_3YG(AG)_3YG(AG)_3$ for the Tyr-rich amorphous fraction. In Figure 1 we demonstrate by magic angle spinning (MAS) solid-state ^{13}C NMR that both the synthetic model peptide $(Ala\text{-}Gly)_{15}$ as well as the Cp fraction of *B. mori* fibroin can be converted from the silk I form (1A and 1B) into silk II (1C and 1D) by formic acid treatment. The resulting chemical shifts in the Ala and Gly regions of these samples are very similar to the spectrum of the native silk fiber (1E). We also note that the lack of Ser in the model peptide $(Ala\text{-}Gly)_{15}$ does not make any difference when compared to the characteristic fibroin repeat $(Ala\text{-}Gly\text{-}Ser\text{-}Gly\text{-}Ala\text{-}Gly)_n$ of native silk.[4b] Interestingly, in Figure 1

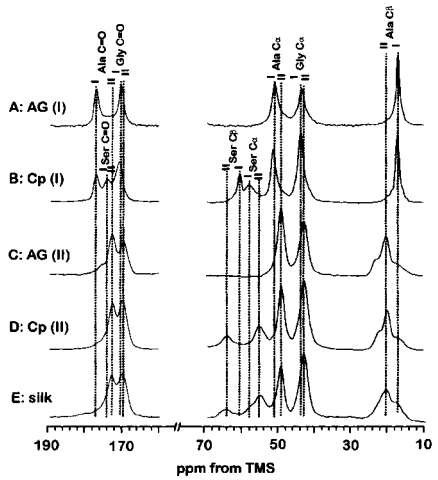

Figure 1. ^{13}C (100 MHz) CP/MAS NMR spectra of (A) the model peptide $(AG)_{15}$, and (B) the Cp fraction of fibroin, both of which were prepared from LiBr in the silk I form. When prepared from formic acid, (C) the model peptide $(AG)_{15}$, and (D) the Cp fraction assume the silk II form, like (E) the native *B. mori* silk fibroin fiber after elimination of silk sericin. The samples were lyophilized or dried under vacuum. The sample spinning rate was 5 kHz, and RF power 80 kHz.

the Ala Cβ peak becomes asymmetric and broad after conversion into silk II (1C−E), and in all three samples the signal can be deconvoluted into three components, as shown in the expanded spectra of Figure 2. This observation suggests that silk II possesses an intrinsically heterogeneous structure. The broad component at the highest field has essentially the same chemical shift as the sharp Ala Cβ peak at 16.7 ppm of silk I (Figure 1).[4a] However, the corresponding structure cannot be attributed to the well-defined β-turn [type II] structure of silk I, because the diagnostic sharp peaks at 50.8 ppm (Ala Cα) and 176.8 ppm (Ala carbonyl) are missing, and instead broad lines with low intensity are observed.[4a] Therefore, we assign the broad component at around 16.5 ppm in Figure 2B to a *distorted* β-turn structure, which is characterized by a large distribution in torsion angles around an average conformation of a type II β-turn.[2] Such features observed here by NMR have not been accessible from previous X-ray analysis.

When assigning the other two components of the multicomponent Ala Cβ signal with chemical shifts of 19.6 and 21.9 ppm (Figure 2B), it is clear that the overall structure of all three silk II samples in Figure 1C−E is consistent with an antiparallel β-sheet, given the diagnostic sharp peaks of the protein backbone at 48.9 ppm (Ala Cα) and 172.2 ppm (Ala carbonyl).[2,3a] Actually, the 2D spin-diffusion NMR spectrum indicates that the conformation of $(AG)_{15}$ in the silk II form is mainly an antiparallel β-sheet, and the torsion angles of the Ala residue are determined to be $\phi, \varphi = -150°, 150°$ ($\pm 10°$) (Figure 3). Since the Ala Cβ methyl groups are located

* To whom correspondence should be addressed. E-mail: asakura@cc.tuat.ac.jp.
§ Tokyo University of Agriculture and Technology.
† University of Jena.

10.1021/ja020244e CCC: $22.00

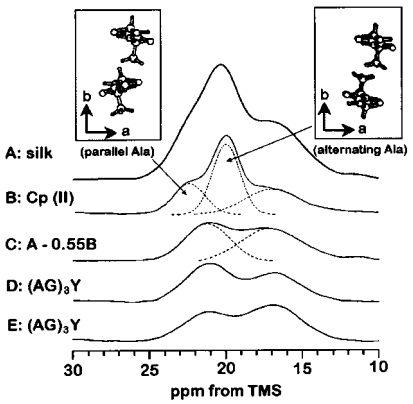

Figure 2. Expanded Ala Cβ peaks of (A) native *B. mori* silk fibers, and (B) the Cp fraction. The difference spectrum (C), representing the Tyr-rich amorphous part of the fiber, was generated by subtracting 55% (B) from (A). Shown as dotted lines underneath are the spectral deconvolutions with Gaussian peaks. The relative intensities of the three peaks in (B) were determined to be 23:45:32 from low field to high field (left to right). Likewise, the relative intensities of the doublet in (C) are 50:50. The Ala Cβ line shapes of the Tyr-rich model peptide $(AG)_3YG(AG)_3YG(AG)_3YG(AG)_3$ are compared (D) after formic acid treatment, and (E) when prepared by dialysis from LiBr.

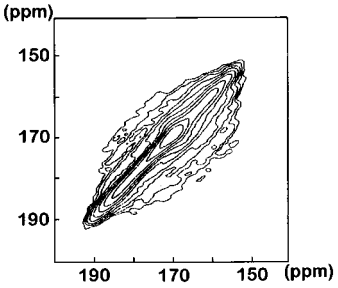

Figure 3. The 2D spin-diffusion NMR spectrum of $(AG)_6A[1\text{-}^{13}C]G[1\text{-}^{13}C]AG(AG)_7$ under off-MAS condition. By simulation[4a] the torsion angles of the 15th Ala residue were determined as ϕ, $\varphi = -150°$, 150° (±10°). The experimental conditions are the same as those in ref 4a.

outside of the protein backbone, the occurrence of two peaks suggests that there may exist differences in the mode of side-chain packing, while the same backbone torsion angles are maintained. To confirm this interpretation we used an *ab initio* molecular orbital method to calculate the ^{13}C shielding constants of the Ala Cβ carbons.[10] Two different kinds of stacks were generated from $(Ala\text{-}Gly)_n$ by packing the protein chains according to the two models proposed by Takahashi et al. as shown in Figure 2.[9] Both of these models possess an antiparallel β-sheet structure; however, the methyl groups are oriented differently between the adjacent sheets. In the first case, all Ala methyl groups point in the same direction, which is perpendicular to the sheet planes. In the second case they alternately point in opposite directions, being also perpendicular to the sheet planes. The initial coordinates were produced by Insight II and energy-minimized at the B3LYP/6-31G* levels. The ^{13}C chemical shielding of the Ala Cβ carbon was calculated by the GIAO method using the RHF/6-31G* levels.[10] The calculated shielding values of the Ala Cβ carbons in the former case (parallel Ala) predict a shift of 2.5 ppm toward lower field compared to the latter case (alternating Ala). This is consistent with the observed chemical shift difference of 2.3 ppm between the two deconvoluted components (Figure 2B). The peak at 21.9 ppm is therefore assigned to the Ala methyl groups aligned in parallel, while the peak at 19.6 ppm corresponds to those pointing in opposite directions. To confirm this assignment, the ^{13}C spin−lattice relaxation times, T_1, were measured. The values are 550 ms (21.9 ppm), 270 ms (19.6 ppm), and 430 ms (16.5 ppm). The shortest value, 270 ms, is indicative of steric hindrance of the Ala methyl groups. In fact, in the alternating Ala case, the intersheet distance between Ala methyl carbons is about 4 Å. The relative peak intensities at 21.9 and 19.6 ppm are approximately 1:2 (Figure 2B), which is in good agreement with the ratio of different packing modes suggested from X-ray diffraction.[9]

As seen in Figures 1 and 2, the line shape of the Ala Cβ carbon of the native *B. mori* silk fiber (1E, 2A) differs slightly from the spectra of the model peptide $(Ala\text{-}Gly)_n$ (1C) and of the Cp fraction (1D, 2B) prepared as silk II. Given that the Cp fraction constitutes only 55% of the total silk fiber,[2] we attribute this difference to the amorphous region that gets cleaved off by chymotrypsin. The corresponding difference spectrum (Figure 2C) reveals a doublet which closely resembles the Ala Cβ signal of the model peptide $(AG)_3YG(AG)_3YG(AG)_3YG(AG)_3$ (Figure 2D).[11] Notably, the spectrum of this peptide is very similar after formic acid treatment or LiBr dialysis (Figure 2D,E), only the relative intensities of the doublet line shape are slightly changed. Therefore, we conclude that the structural transition from silk I to silk II does not affect the Tyr-rich region and rather occurs exclusively within the repeated $(Ala\text{-}Gly\text{-}Ser\text{-}Gly\text{-}Ala\text{-}Gly)_n$ region of *B. mori* silk fibroin. Deconvolution of the Ala Cβ difference spectrum (Figure 2C) yields two peaks with equal intensities at around 16.5 and 21.1 ppm (Figure 2C). The broad peak with a chemical shift around 16.5 ppm is assigned to a distorted β-turn [typeII] conformation, as in the case of silk II. The component at lower field is assigned to a distorted β-sheet structure, being broad and lying between the two characteristic peaks of an undistorted β-sheet. Having interpreted each of the deconvoluted components of the ^{13}C NMR peak of Ala Cβ in Figure 2B,C, we can now estimate the relative proportions of the various heterogeneous components in the native silk fiber of *B. mori* from their relative peak intensities: The Cp fraction (constituting 55% of the total silk fiber) thus consists of 18% distorted β-turn, 13% β-sheet with parallel Ala, and 25% β-sheet with alternating Ala. Half of the local conformations in the remaining amorphous Tyr-rich region (45% of the total silk fiber) are distorted β-turns, the other half, distorted β-sheets.

Acknowledgment. T.A. acknowledges Professor D. Kaplan at Tufts University for stimulating discussions, and support from the Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan. A.S.U. thanks the DFG for support of TP B13 in SFB 197.

References

(1) Asakura, T.; Kaplan, D. L. In *Encyclopedia of Agricultural Science*; Arutzen, C. J., Ed.; Academic Press: New York, 1994; Vol. 4, pp 1−11.
(2) Asakura, T.; Kuzuhara, A.; Tabeta, R.; Saito, H. *Macromolecules* **1985**, *18*, 1841−1845.
(3) (a) Ishida, M.; Asakura, T.; Yokoi, M.; Saito, H. *Macromolecules* **1990**, *23*, 88−94. (b) Panitch, A.; Matsuki, K.; Cantor, E. J.; Cooper, S. J.; Atkins, E. D. T.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *Macromolecules* **1997**, *30*, 42−49.
(4) (a) Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komatsu, K. *J. Mol. Biol.* **2001**, *306*, 291−305. (b) Asakura, T.; Yamane, T.; Nakazawa, Y.; Kameda, T.; Ando, K. *Biopolymers* **2001**, *58*, 521−525.
(5) Marsh, R.; Corey, R. B.; Pauling, L. *Biochim. Biophys. Acta* **1955**, *16*, 1−34
(6) Fraser, B.; MacRae, T. P. *Conformations of Fibrous Proteins and Related Synthetic Polypeptides*; Academic Press: New York, 1973; pp 293−343.
(7) Lotz, B.; Keith, H. D. *J. Mol. Biol.* **1971**, *61*, 201−215.
(8) Fossey, S. A.; Nemethy, G.; Gibson, K. D.; Scheraga, H. A. *Biopolymers* **1991**, *31*, 1529−1541.
(9) Takahashi, Y.; Gehoh, M.; Yuzuriha, K. *Int. J. Biol. Macromolecules* **1999**, *24*, 127−138.
(10) Wolinski, K.; Hilton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.
(11) Zhou, C.-Z.; Confalonieri, F.; Medina, N.; Zivanovic, Y.; Esnault, C.; Yang, T.; Jacquet, M.; Janin, J.; Duguet, M.; Perasso, R.; Li, Z.-G. *Nucleic Acids Res.* **2000**, *28*, 2413−2419.

JA020244E