Macromolecules

Article

pubs.acs.org/Macromolecules

Local Structure and Dynamics of Serine in the Heterogeneous Structure of the Crystalline Domain of *Bombyx mori* Silk Fibroin in Silk II Form Studied by 2D $^{13}C-^{13}C$ Homonuclear Correlation NMR and Relaxation Time Observation

Keiko Okushita,[†,‡] Atsushi Asano,[‡] Michael P. Williamson,[§] and Tetsuo Asakura*[,†,∥]

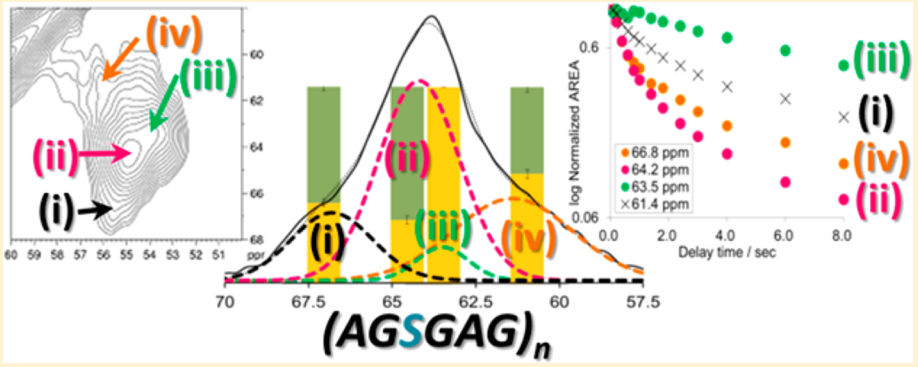
[†]Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan

[‡]Department of Applied Chemistry, National Defense Academy, Yokosuka 239-8686, Japan

[§]Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, U.K.

[∥]Institute for Molecular Science, 38 Nishigo-Naka, Myodaiji, Okazaki 444-8585, Japan

Supporting Information



ABSTRACT: The crystalline fraction (Cp fraction) of silk fibroin in silk II form from the silkworm *Bombyx mori* is the classic example of antiparallel β-sheet and consists mainly of Ala, Ser, and Gly. In the ^{13}C CP/MAS NMR spectrum, the Ala Cβ, Ser Cα, and Ser Cβ peaks are asymmetric, showing the heterogeneous nature of the structure, which is shown to consist of two different packing geometries, denoted domains A and B. In this work, these peaks were resolved and assigned using 2D $^{13}C-^{13}C$ homonuclear correlation NMR spectra collected with different dipolar- assisted rotational resonance (DARR) mixing times. The local structure and dynamics of serine in the Cp fraction are discussed in detail. Model peptides of the Cp fraction, [3-^{13}C]Ser-(AGSGAG)$_5$ with different [3-^{13}C]Ser labeling positions, were prepared to clarify that the individual peaks of the Ser Cβ carbons come from different domains, not from different positions within a chain. The dynamics of these individual peaks were studied by measuring ^{13}C T_1 values, which provide information on serine side chain dynamics and thus on the formation of intermolecular hydrogen bonding through the Ser OH group. Moreover, cross peaks between domains A and B were observed in the Ala Cβ DARR spectrum with a long mixing time of 400 ms, indicating a close contact between the two geometries. We conclude that the crystalline region is heterogeneous, comprising two closely associated packing geometries that form separate but small domains.

INTRODUCTION

Silks continue to attract the attention of researchers in biology, biochemistry, biophysics, analytical chemistry, polymer technology, textile technology, and tissue engineering because of their supreme physical and medical properties.[1] *Bombyx mori* (*B. mori*) silk fibroin fiber is spun from an aqueous solution by silkworms at room temperature, yet is strong and tough. The overall composition of the silk fibroin in mol % consists of glycine (42.9%), alanine (30.0%), serine (12.2%), tyrosine (4.8%), and valine (2.5%).[1] The silk fibroin is comprised 55% of repeated AGSGAG sequence, which is thought to form the crystalline fraction (Cp).[2] This Cp fraction can be prepared from *B. mori* silk fibroin by proteolytic digestion.

Two crystalline structures have been reported for *B. mori* silk fibroin. One is silk I and represents the structure of the silk fibroin stored in the silk gland of *B. mori* before spinning. The other is silk II and is the structure of the silk fiber after spinning. We have previously proposed a structural model for silk I as a "repeated β-turn type II structure" on the basis of several solid state NMR analyses.[3−5] The X-ray diffraction pattern of native silk fibroin fiber shows that silk II is an antiparallel β-sheet. A model was proposed by Marsh et al. in

Received: May 2, 2014
Revised: June 8, 2014
Published: June 17, 2014

1955[6] and remains the accepted structure. However, no high-resolution crystal structure has been reported, and experimental data consistently indicate heterogeneity. Lotz et al.[7] demonstrated structural disorder of the silk II $(AG)_n$ sequence from their X-ray and electron diffraction analyses, with a Gly-Gly intersheet spacing greater than that seen in polyglycine and an Ala-Ala spacing less than that seen in polyalanine. Fraser et al.[8] have also reported that the antiparallel β-sheet has a slightly greater intersheet spacing for $(AGSGAG)_n$ as compared to that for $(AG)_n$. In our previous papers,[9,10] using ^{13}C CP/MAS NMR, we found that the silk II structure is heterogeneous: the Cp fraction comprises 18% distorted β-turns, 25% β-sheet (domain A), and 13% β-sheet (domain B), while the remaining 45% amorphous fraction is composed almost equally of distorted β-turn and β-sheet. This information was obtained from a detailed analysis of the Ala-Cβ carbon peaks in the ^{13}C CP/MAS NMR of the Cp fraction and appropriate model peptides.

In this paper, we clarify the heterogeneous nature of the structure in the crystalline domain of *B. mori* silk fibroin in silk II form using two-dimensional (2D) $^{13}C-^{13}C$ DARR (dipolar-assisted rotational resonance)[11] spectra of the ^{13}C uniformly labeled Cp fraction of *B. mori* silk fibroin. The observed 2D DARR cross peaks with different mixing times reveal short- and long-range through-space dipolar contacts between ^{13}C spins. Especially, a detailed assignment of the Ser-Cβ peaks can be attained by 2D DARR measurements, which give assignments for the asymmetric and broad peak correlating the Ser-Cβ and Ala-Cβ carbons. Namely, the two Ala-Cβ peaks showed distinct cross peaks with the Ser-Cβ peak. The ^{13}C chemical shifts of the Ser-Cβ obtained from the 2D DARR cross peak revealed that the Ser-Cβ has different local environments in the two β-sheet chains. Serine is a key determinant of the silk I or silk II structure through hydrogen bonding, involving the $O^{\gamma}H$ group of the Ser side chain as reported previously.[12−16] Namely, the Ser side chain is capable of forming intramolecular hydrogen bonds involving the $O^{\gamma}H$ group and the backbone carbonyl group of the Gly residue, stabilizing the silk I form.[15] A large difference in the side chain mobility between hydrogen-bonded and non-hydrogen-bonded formations through the Ser OH group has been reported from ^{13}C spin−lattice relaxation time (T_1^C) observation of the Cp fraction and model peptides.[12−14] On the other hand, the ^2H NMR powder pattern spectra of [3,3-^{2}H]Ser-*B. mori* silk fiber with silk II form have been analyzed by line shape simulation.[16] Two types of motion could be characterized quantitatively: one component undergoing a rapid three-site jump and a second component representing a slow exchange between sites with unequal occupancies and with a small amplitude of libration. In addition, ^2H solid state NMR studies of uniaxially aligned [3,3-^{2}H]Ser-*B. mori* silk fiber with silk II form have been performed to determine the side-chain conformation and the orientational distribution of the Ser residues in the slow motional component: the dominant conformer of the Ser side chain is *gauche*$^+$ around the N−Cα−Cβ−O bond.[16] Therefore, we also prepared [3-^{13}C]Ser-$(AGSGAG)_5$[14] with different selective [3-^{13}C]Ser labeling positions as model peptides of the Cp fraction in order to obtain structural and dynamical information more selectively.

Furthermore, we detected a cross peak between the two peaks of the Ala-Cβ (domains A and B) at a long mixing time of 400 ms, while there was no cross peak with short mixing time. These observations suggested that there are two locally different β-sheet geometries and that the two geometries are in close proximity to each other. The small domain size between the two β-sheets was also suggested by ^1H spin−lattice relaxation measurements in the laboratory frame (T_1^H) and in the rotating frame $(T_{1\rho}^H)$.

EXPERIMENTAL SECTION

Materials. *Crystalline Fraction of [U-^{13}C] B. mori Silk Fibroin.* *B. mori* larvae were reared in our laboratory. Uniformly ^{13}C-labeled *B. mori* silk fibroin, [U-^{13}C] *B. mori* silk fibroin, was achieved biosynthetically by oral administration of an artificial diet with ^{13}C-uniformly labeled glucose to larvae of the fifth instar.[17,18] The supplement was mixed with 2 g of an artificial diet per day. The amount of [U-^{13}C] glucose was 10 mg each on the fourth and fifth day of the fifth larval stage. The total amount of [U-^{13}C] glucose administered was thus 20 mg per silkworm. The [U-^{13}C] Cp fraction of *B. mori* silk fibroin was prepared from regenerated [U-^{13}C] *B. mori* silk fibroin.[19] Chymotrypsin (40 mg) dissolved in a few milliliters of water was added to an aqueous solution of about 4 g of fibroin buffered with $Na_2HPO_4{\cdot}12H_2O$ and $NaH_2PO_4{\cdot}2H_2O$ at pH 7.8. The solution (300 mL) was incubated at 37 °C for 24 h, and the precipitate was separated by centrifuging at 10 000 rpm followed by washing with 0.03 N HCl to inactivate the enzyme reaction. The precipitate was further washed several times with distilled water, ethyl alcohol, and diethyl ether and freeze-dried, yielding a crystalline fraction corresponding to 55% of the original fibroin. The obtained precipitate forms the silk II structure and is referred to as the Cp fraction.

Synthesis of Peptides [3-^{13}C]Ser^9-, [3-^{13}C]Ser^{15}-, [3-^{13}C]Ser^{21}-$(AGSGAG)_5$, and $(AGSGAG)_5$. H-[3-^{13}C]Ser OH purchased from Cambridge Isotope Lab., Inc., was converted to Fmoc-[3-^{13}C]Ser(*t*-Bu)-OH (yield: 60%).[20] The details of the synthesis are described in the Supporting Information. The Fmoc solid-phase method using a fully automated Pioneer Peptide Synthesis System (Applied Biosystems Ltd. Japan) was used to prepare [3-^{13}C]Ser^9-, [3-^{13}C]-Ser^{15}-, and [3-^{13}C]Ser^{21}-$(AGSGAG)_5$ peptides together with unlabeled peptide as reported previously.[21] These peptides were dissolved in 9 M LiBr and then dialyzed against distilled water. The obtained precipitate forms the silk II structure.

^{13}C CP/MAS NMR and T_1^C Experiments. The ^{13}C CP/MAS NMR measurements were conducted on a Bruker AVANCE-400 spectrometer operating at 100 MHz for the ^{13}C nucleus. Cross-polarization (CP) was employed for sensitivity enhancement with high-power ^1H decoupling during the signal acquisition. A ^1H 90° pulse width of 5 μs duration was used with 1 ms contact time and 3 s repetition time. The number of scans for uniformly ^{13}C-labeled *B. mori* silk fibroin samples was 64. Approximately 15K FIDs were added to generate the spectra of synthesized peptides. The ^{13}C chemical shifts were calibrated indirectly through the adamantane methylene peak observed at 28.8 ppm relative to TMS (tetramethylsilane) at 0 ppm. The ^{13}C spin−lattice relaxation times, T_1^C, were observed using a JEOL ECX 400 spectrometer operating at 100 MHz for the ^{13}C nucleus. We used a pulse sequence for T_1^C determination developed by Torchia.[22] The MAS spinning speed was 8 kHz, and the ^1H 90° pulse, ^{13}C 90° pulse, contact, and repetition times were 3.9 μs, 5 μs, 1.2 ms, and 5 s, respectively. The T_1^C components were fitted to two exponentials according to the following equation:

$$M_{CP}(\tau) = M_{CP}(0)\left\{I \exp\left(-\frac{\tau}{T_{1\ \mathrm{short}}^C}\right) + (1-I)\exp\left(-\frac{\tau}{T_{1\ \mathrm{short}}^C}\right)\right\}$$

where $M_{CP}(\tau)$ and $M_{CP}(0)$ stand for the ^{13}C NMR peak intensities at times τ and 0, respectively. Spectra were accumulated 512 times with arrayed variable delay range of 0.05−8.0 s at room temperature.

^{13}C−^{13}C Dipolar Assisted Rotational Resonance Experiment. The ^{13}C DARR measurements[11] were conducted on a Bruker DMX-500 spectrometer operating at 125 MHz for the ^{13}C nucleus. The experiment was performed at a magic-angle spinning (MAS) frequency of 11 kHz with a 4 mm o.d. rotor. The CP contact time and the recycle delay were 1.5 ms and 3 s, respectively. The mixing time was varied from 10 to 400 ms, and 64 or 72 scans were collected at each time

dx.doi.org/10.1021/ma500908m | *Macromolecules* 2014, 47, 4308−4316

increment of the indirect dimension. The ^{1}H power for the recoupling was adjusted to the same as that of the MAS frequency in Hz.

T_1^H and $T_{1\rho}^H$ Experiments. The T_1^H and $T_{1\rho}^H$ experiments[23,24] were conducted on a Varian 400WB NMR system operating at 100 MHz for the ^{13}C nucleus. The experiments were performed at a MAS frequency of 9 kHz with a 3.2 mm o.d. rotor. The CP contact time and the recycle delay were 200 μs and 5 s, respectively. Both T_1^H and $T_{1\rho}^H$ measurements were performed at room temperature. A ^{1}H spin-locking frequency of 64 kHz was used for the $T_{1\rho}^H$ measurement.

RESULTS AND DISCUSSION

1. Deconvolution of Ser-Cα and Cβ Peaks and Assignment to the β-Sheets A and B. Figure 1 shows ^{13}C CP/MAS spectra of [U-^{13}C] *B. mori* silk fiber (a) and [U-^{13}C] Cp fraction in silk II form (b). *B. mori* silk fiber mostly consists of two repeated sequences: AGSGAG and $(GX)_mGY$ (X = A or V).[2] The Cp fraction mainly consists of repeated AGSGAG sequences and occupies about 55 mol % of the silk fibroin. The AGSGAG sequence is thus the main part of the silk II structure. The spectrum of the [U-^{13}C] Cp fraction shows similar line shapes and the same chemical shifts as those of the [U-^{13}C] *B. mori* silk fiber. This observation indicates that the main structure in *B. mori* silk fiber is the same as that of the Cp fraction in silk II structure. However, the existence of other sequences like $(GX)_mGY$[2] in the silk fiber makes analysis difficult and impedes clear investigation of the silk II structure. Not surprisingly, the peak signals from Ser-Cβ and Ala-Cβ in *B. mori* fiber become relatively wider than those of the Cp fraction. This broadening is due to the existence of amorphous domains consisting of $(GX)_mGY$ sequence and N-terminal and C-terminal regions.[2] In contrast, the Cp fraction has no such domains, so the observed ^{13}C NMR spectrum reflects the silk II structure arising almost entirely from AGSGAG. Therefore, NMR investigation of the Cp fraction gives important information on the silk II structure of *B. mori* silk fiber in detail.

In the Ala-Cβ region of the Cp fraction, three peaks are clearly seen at 16.2, 19.6, and 21.7 ppm. It has been reported that the characteristic ^{13}C NMR line shape in the Ala-Cβ region appears also in the silk II structure. The three Ala-Cβ peaks were assigned to a distorted β-turn and two different packings of β-sheets, respectively.[9,10] The β-sheet peaks at 19.6 and 21.7 ppm have an intensity ratio close to 2:1 (Figure 1b), suggesting the presence of two locally different β-sheets, as proposed previously by Takahashi et al.,[25] which we designate as A and B, respectively. We have already proposed a candidate structure for the two β-sheets, that is, antiparallel β-sheet structures with a different intersheet arrangement.[9,10] The chemical shifts of both β-sheets A and B are consistent with the values predicted from the typical torsion angle (ϕ, ψ) of β-sheet structure on the Ala-Cβ chemical shift contour plot.[26] The typical torsion angle of a β-sheet structure is (−140°, 140°),[27] and the ^{13}C chemical shift is approximately 19.5 ppm (Figure 2a, orange region).

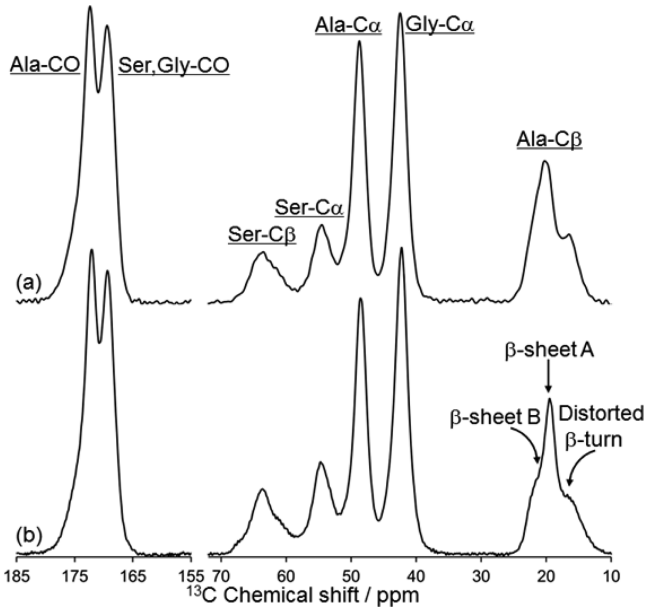


Figure 1. ^{13}C CP/MAS spectra of (a) [U-^{13}C] *B. mori* fiber and (b) [U-^{13}C] Cp fraction (silk II) observed using 400 MHz NMR at MAS = 9 kHz.

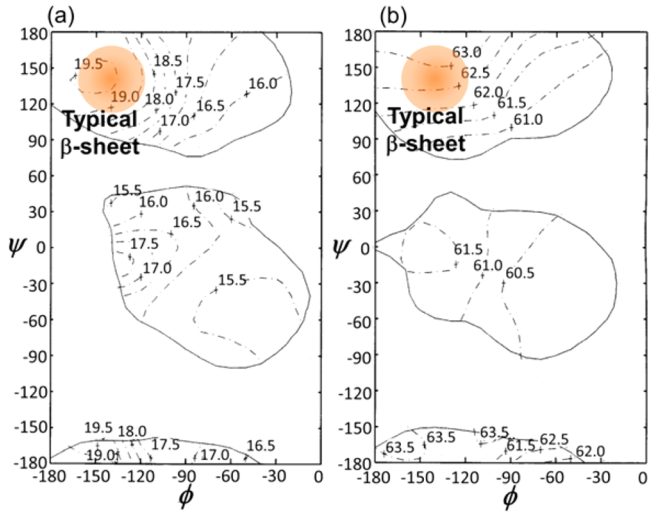


Figure 2. ^{13}C chemical shift contour plots of Cβ carbons of Ala (a) and Ser (b) residues in proteins published previously.[26] The (ϕ, ψ) angle regions colored in orange are typical β-sheet regions of (ϕ, ψ) = (−140°, 140°).

Thus, the β-sheet A at 19.6 ppm probably has the typical torsion angle of a β-sheet. In contrast, the β-sheet B is observed at the relatively lower field of 21.7 ppm. This observation implies the existence of an intersheet deshielding packing effect in the β-sheet B region. Therefore, both peaks at 19.6 and 21.7 ppm are attributed to β-sheets with different intersheet packing.

A broad peak was observed for the Ser-Cβ at around 63.7 ppm (Figure 1b). This value is slightly larger than the 62.8 ppm that is predicted from the typical β-sheet geometry (the orange region of Figure 2b).[26] Additionally, the peak shape indicates the existence of several Ser residues in different environments. To resolve individual peaks from the broad and complicated Ser-Cβ peak, we employed 2D ^{13}C−^{13}C DARR experiments. Figure 3a shows a 2D ^{13}C−^{13}C DARR spectrum of the [U-^{13}C] Cp fraction at a mixing time of 10 ms. The cross peaks between Ser-Cα and Ser-Cβ shown in the expanded box in Figure 3a must be intraresidue because spin transfer across long distance cannot occur with such a short mixing time. Four such cross peaks are clearly observed: hereafter we abbreviate them as (Cα 56.0, Cβ 61.4), (Cα 54.0, Cβ 63.5), (Cα 54.9, Cβ 64.2), and (Cα 55.6, Cβ 66.8). Three of these peaks have Cβ chemical shifts larger than that of the typical β-sheet chemical shift of 62.8 ppm. The large chemical shift value is ascribed to an intersheet packing effect as stated above for Ala-Cβ. In contrast,

dx.doi.org/10.1021/ma500908m | Macromolecules 2014, 47, 4308−4316

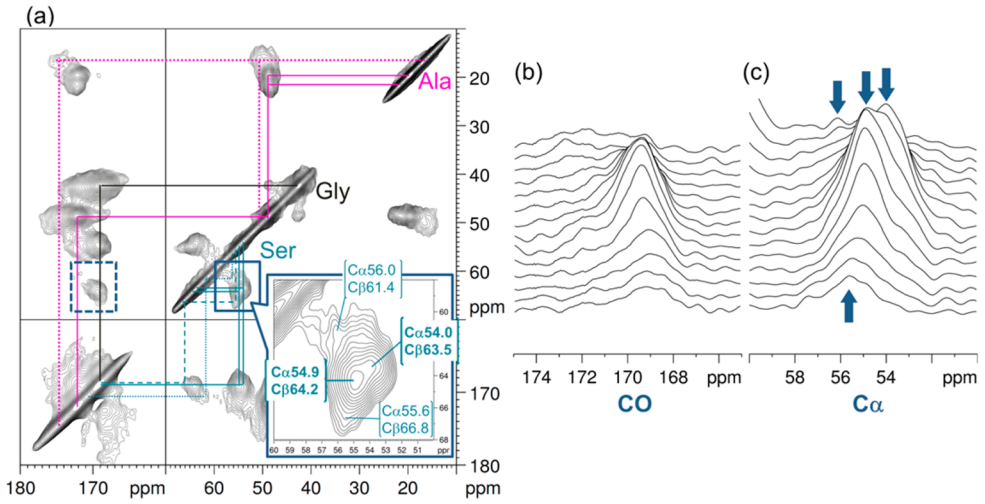


Figure 3. 2D $^{13}C-^{13}C$ DARR spectrum of [U-^{13}C] Cp fraction (silk II) observed with a mixing time of 10 ms. (a) Intraresidue assignment of the detected peaks and expansion of the Ser Cα-Cβ correlation. Stack plots of Ser-CO ((b), broken square region in (a)) and Cα ((c), expanded in (a)) regions.

the cross peak (Cα 56.0, Cβ 61.4) is attributable to a random coil because both chemical shifts are in agreement with the ^{13}C solution NMR chemical shifts of Ser-Cβ in proteins with random coil conformation in aqueous solution.[28]

To compare the amount of the three Ser-Cβ β-sheet peaks and to assign them to individual β-sheet structures, a peak deconvolution using four Gaussian functions was performed in the Ser-Cβ peak region for both [U-^{13}C] *B. mori* fiber and the [U-^{13}C] Cp fraction (Figure 4). The four chemical shift values for the fitting were obtained from the 2D spectra in the expanded region of Figure 3a. Moreover, we limited the number of variables by using the same line widths to fit the 3 β-sheet peaks (peaks i–iii in Figure 4). Peak iv (random coil fraction) was fitted best using a larger peak width, which seems reasonable because the random coil or disordered region has a larger structural distribution than the ordered region in polymers. The Ser-Cβ peak from the *B. mori* fiber spectrum (Figure 4a) includes signal from the soluble and/or low-molecular-weight fraction[2] because the fiber is not cleaved at Tyr by α-chymotrypsin. Therefore, Ser-Cβ peaks in the fiber spectrum show a large structural distribution, resulting in a broad line width. In support of this argument, it has been shown that the relative amount of structural heterogeneity or randomness of the Ser-Cβ peak was increased on the replacement of Ser by Tyr, in the sequence AGYGAG.[29] The replacement induced a broadening of the Ser-Cβ peak around 67 ppm. The area of the peak at 64.2 ppm (peak ii in Figure 4) is much larger than that at 63.5 ppm (peak iii). The relative ratio of β-sheets A and B has been determined from the Ala-Cβ peaks to be 2:1.[10] Considering the Ala-Cβ conclusion, the more intense Ser-Cβ peak at 64.2 ppm may thus be assigned to the β-sheet A structure from the intensity. But there is a discrepancy between the relative peak areas fitted in the Ser Cβ region (37:12) and the Ala Cβ region (2:1), which is presumably due to the difficulty of fitting the highly overlapping peaks in the Ser Cβ region. Peak assignments and their relative intensities are tabulated in Table 1. For more detailed assignment, we analyze the correlation between Ala-Cβ and Ser- Cβ peaks in the next section. Moreover, we evaluate the change in the mobility through the hydrogen-bonding formation of the Ser side chain OH group.

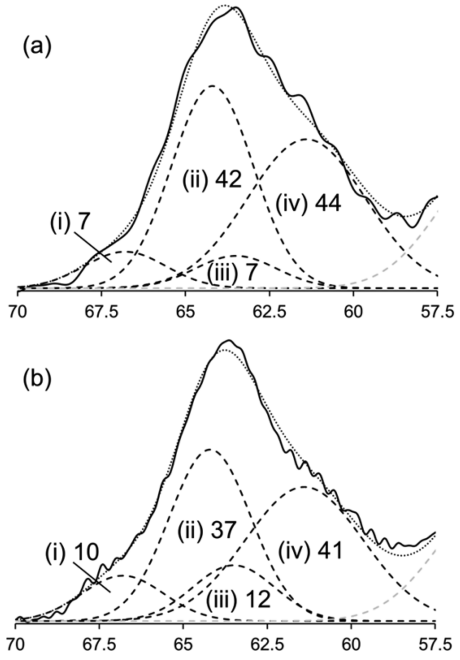


Figure 4. Gaussian line shape fitting result of Ser-Cβ peaks in ^{13}C CP/MAS spectra of (a) [U-^{13}C] *B. mori* silk fiber and (b) [U-^{13}C] Cp fraction (silk II). The solid lines show the observed spectra. The broken lines show individual Gaussian functions (- - -) and the sum of all functions (···). The fractions are assigned to (i) disordered A, (ii) A, (iii) B, and (iv) random coil.

2. Local Intermolecular Structure around Ser in Each β-Sheet Structure. Figure 5a shows the correlations between Ala-Cβ and Ser in the DARR spectrum at the long mixing time of 400 ms. Several correlations between Ser-Cβ around 64 ppm and Ala-Cβ at 19.6/21.7 ppm were detected. In standard antiparallel β-sheet models of silk II such as the Marsh model (Figure 5b),[6] the intermolecular intercarbon distance between Ser-Cα/Cβ and Ala-Cβ in the adjacent sheet is 0.48/0.41 nm while the intrachain distance is much longer at 0.71/0.70 nm. In a DARR spectrum, shorter internuclear distances contribute more effectively to cross peak intensity. Therefore, one expects

dx.doi.org/10.1021/ma500908m | *Macromolecules* 2014, 47, 4308–4316

Table 1. ^{13}C Chemical Shift Assignment of Cp Fraction (Silk II)

residue	domain	Cβ (ppm)	Cα (ppm)	CO (ppm)	fraction (%) in each residue
Ala	random coil/ distorted β-turn	16.2	50.5	173.4−175.1	32[b]
	β-sheet A	19.6	48.5	172.0	45[b]
	β-sheet B	21.7			23[b]
Gly			42.4	169.0	
Ser	random coil	61.4	56.0	170.6	41
	β-sheet B	63.5	54.0	169.1	12
	β-sheet A	64.2	54.9	169.1	37
	β-sheet A[a]	66.8	55.6	169.1	10

[a]Distorted structure. [b]Reference 10.

that the cross peak between Ser-Cα/Cβ and Ala-Cβ will be dominated by the intersheet rather than the intramolecular contribution. In Figure 5a, the Ser-Cα peaks at 54.9 and 54.0 ppm (corresponding to geometries A and B, respectively) show correlations with the Ala-Cβ peaks at 19.6 and 21.7 ppm, also corresponding to geometries A and B, respectively. This observation supports the presence of separate A and B domains because any model that has alternating or statistically random A and B strands would have A−B interactions as the shortest intersheet distance. The Ser-Cβ peak at 64.2 ppm (geometry A) also has a strong cross peak with Ala-Cβ peak A, while the Ser-Cβ B peak has correlations that are too weak to interpret. However, correlations are also seen between geometries A and B, for Ser-Cβ A to Ala-Cβ B and between Ser-Cα and Ala-Cβ. If the two geometries A and B formed large separate domains, interdomain cross peaks could not be observed in the 2D

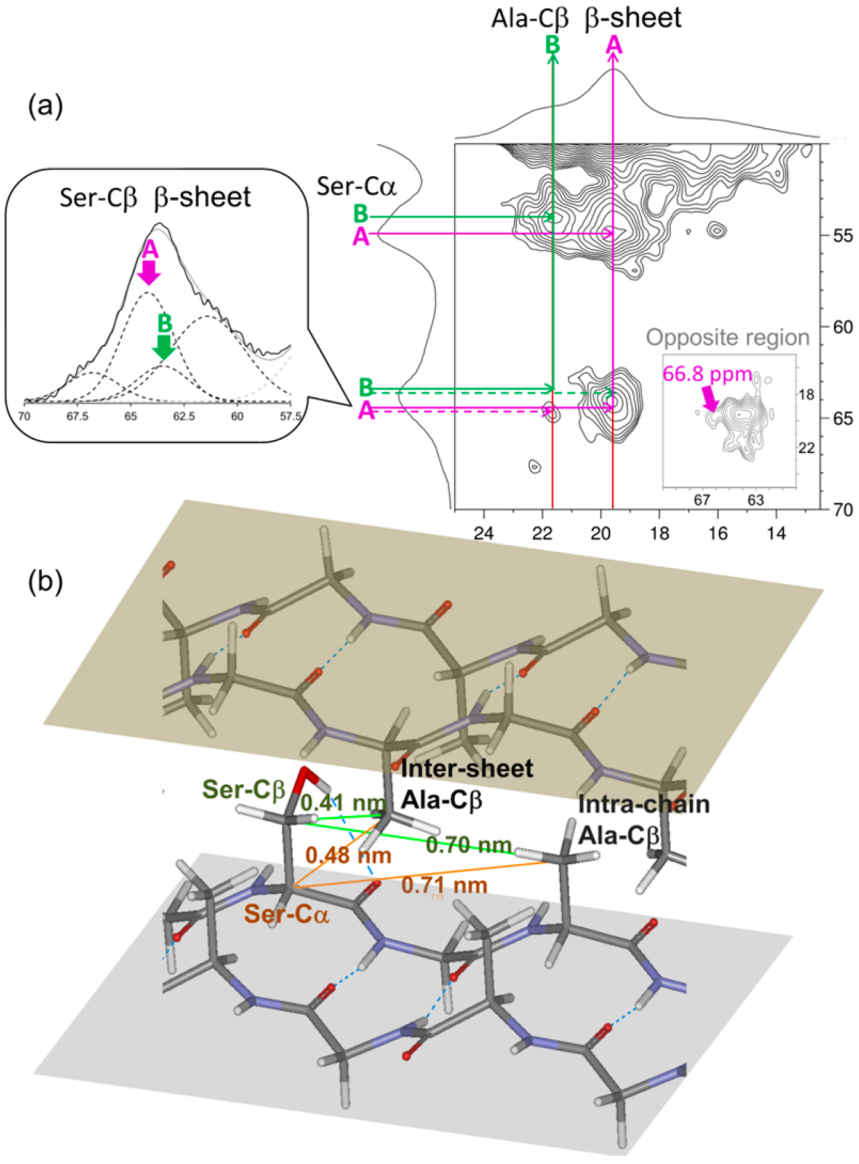


Figure 5. (a) (Ala-Cβ)−(Ser-Cα/Cβ) correlations in the 2D ^{13}C−^{13}C DARR spectrum of [U-^{13}C] Cp fraction (silk II) at a mixing time of 400 ms. Solid lines show the correlations within each β-sheet structure A and B, which have been assigned in Figures 3 and 4. The spectrum on the left is the ^{13}C CP/MAS spectrum of the same sample observed at 500 MHz. (b) Intercarbon distances in the Marsh model, which is the generally accepted antiparallel silk II model. Even in this model, the shortest intercarbon distances between Ser-Cα/Cβ and Ala-Cβ are intermolecular rather than intramolecular.

dx.doi.org/10.1021/ma500908m | Macromolecules 2014, 47, 4308−4316

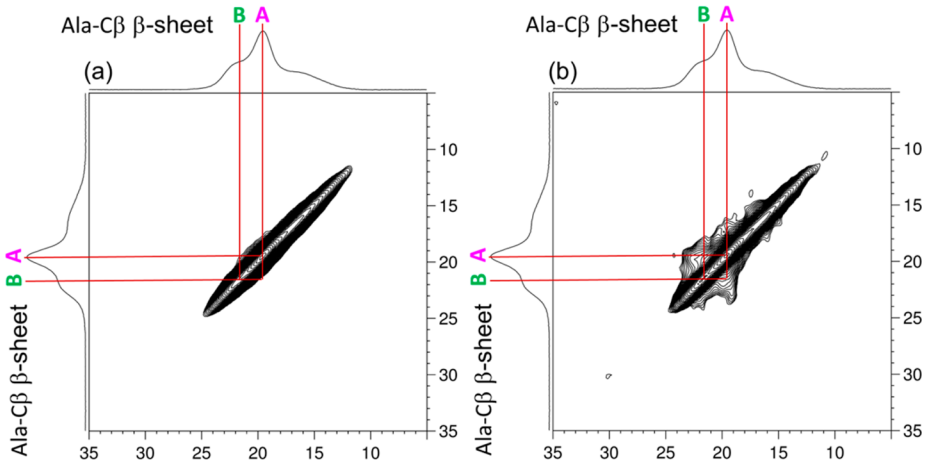


Figure 6. Correlation between the two β-sheet structures A and B in the 2D $^{13}C-^{13}C$ DARR spectra of [U-^{13}C] Cp-fraction (silk II) at mixing times of 10 ms (a) and 400 ms (b).

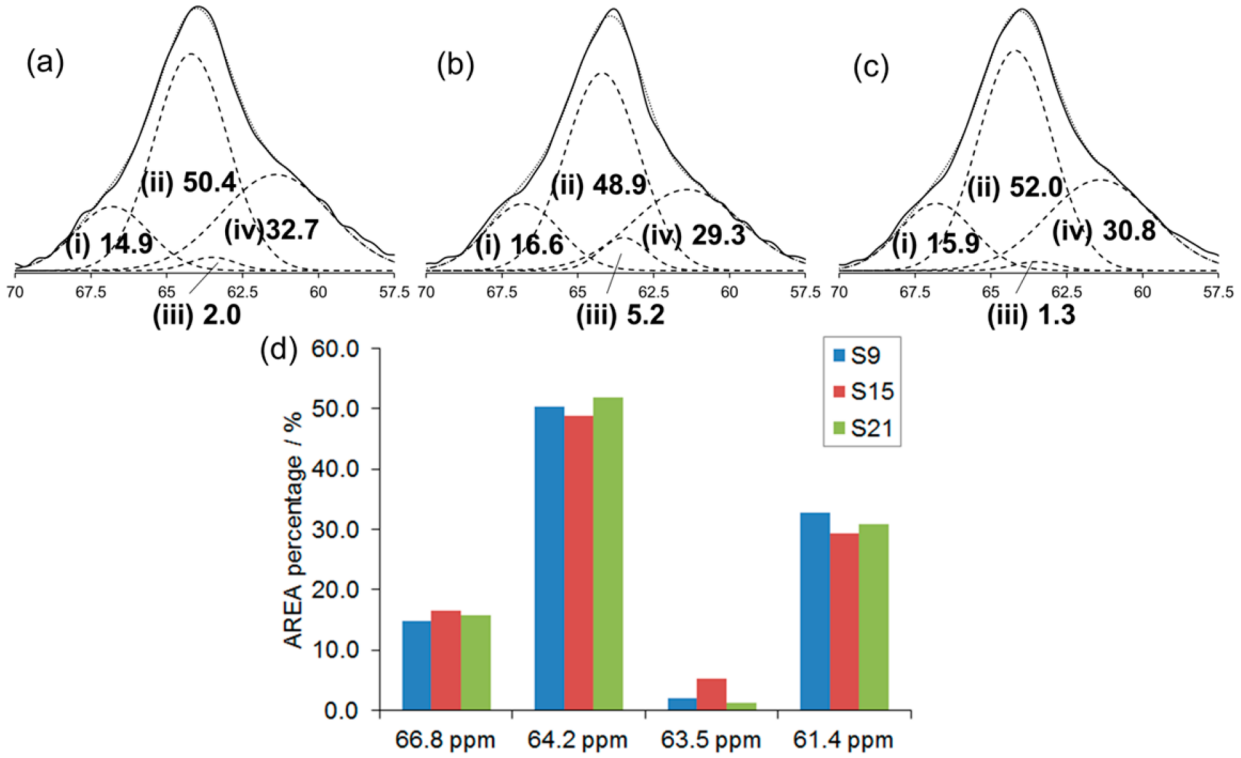


Figure 7. Gaussian line shape fitting results of Ser-Cβ peaks in ^{13}C CP/MAS spectra of (a) [3-^{13}C] Ser^{9}(AGSGAG)$_5$, (b) [3-^{13}C] Ser^{15}(AGSGAG)$_5$, and (c) [3-^{13}C] Ser^{21}(AGSGAG)$_5$. The peak intensities from natural abundance components were subtracted from the original intensities in their CP/MAS spectra. The broken lines show individual Gaussian functions (- - -) and the sum of all functions (···). (d) A column chart of the percentage area at each Ser-Cβ peak of all ^{13}C-labeled (AGSGAG)$_5$ peptides. The fractions at 66.8, 64.2, 63.5, and 61.4 ppm are assigned respectively to (i) disordered A, (ii) A, (iii) B, and (iv) random coil.

DARR spectrum. Hence, it is concluded that the β-sheets A and B form small domains with extensive interdomain contact.

In the 2D DARR spectrum with a 400 ms mixing time, a cross peak can be observed between Ala-Cβ A and Ala-Cβ B (Figure 6), indicating that the two geometries A and B must be very close together. In a preliminary DARR measurement, a very weak correlation was detected from an intercarbon distance of 0.65 nm in *a* > 99% ^{13}C triple-labeled model compound. A previous paper[30] has also estimated 0.70 nm to be the upper limit for DARR-detectable distances. These two results therefore imply that Ala-Cβ A and Ala-Cβ B must be within 0.7 nm, which is short compared to the intersheet increment of *B. mori* silk fibroin which has been shown to be 0.94 nm by the previous X-ray study.[6]

For the Ser-Cβ peak at 66.8 ppm (Figure 4, component i), a correlation with Ala-Cβ geometry A was found on the other side of the diagonal (Figure 5a, "opposite region"), demonstrating that the Ser-Cβ at 66.8 ppm is close to the β-sheet A structure. This peak tends to increase in intensity when the β-sheet structure is perturbed by the replacement of Ser to Tyr in the sequence.[29] We therefore assign the peak at 66.8 ppm to a distorted β-sheet A structure in the Cp fraction. This peak has a chemical shift 2.6 ppm higher than the normal domain A resonance, and the shift difference is larger for Ser

dx.doi.org/10.1021/ma500908m | *Macromolecules* 2014, 47, 4308−4316

Cβ than Ser Cα (Table 1). A change of this magnitude is not compatible with a change in χ_1 angle or backbone hydrogen bonding,[31,32] and we therefore suggest that the distortion may involve hydrogen bonding to the Ser OH.

3. Dynamics of Ser Residues and Intermolecular Hydrogen Bonding Formation through the Ser OH Side Chain. Expansions of the Ser Cβ peaks in the ^{13}C CP/MAS NMR spectra of [3-^{13}C]Ser^{9}-(AGSGAG)$_5$, [3-^{13}C]Ser^{15}-(AGSGAG)$_5$, and [3-^{13}C]Ser^{21}-(AGSGAG)$_5$ are shown in parts a, b, and c of Figure 7, respectively. The peak intensities from natural abundance components were subtracted from the original intensities in the CP/MAS spectra. The whole regions of the original CP/MAS spectra are also shown in the Supporting Information. The number of components and the chemical shift values used for the fitting were again obtained from Figure 3a. However, the fitted line widths were slightly narrower than those of the Cp fraction (Figure 4b) because the synthesized model peptides are chemically defined and therefore more uniform. Selective ^{13}C labeling of Ser Cβ carbons is useful for analysis of the distribution of serines between the two domains and of the dynamics of the domains using T_1^C measurements.

First, comparing the distribution of peak area for Ser-Cβ peaks at each chemical shift, it can be seen that there is no significant difference within experimental error among all three ^{13}C-labeled peptides. This implies that the chemical shift or the local conformation is independent of the position of the Ser Cβ carbon in a chain. Therefore, it is suggested that the separate A and B domains discussed above are formed independent of the location in the (AGSGAG)$_n$ sequence: in other words, the individual peaks of the Ser Cβ carbons come from different domains, not from different positions within a chain.

To analyze the molecular mobility of each component in the Ser-Cβ carbon region, we measured T_1^C for each individual component, at 66.8, 64.2, 63.5, and 61.4 ppm, in the Ser^{15}-Cβ peak of [3-^{13}C]Ser^{15}-(AGSGAG)$_5$, which was selected as a representative ^{13}C-labeled sample. The T_1^C plots are shown in Figure 8, and the T_1^C values are summarized in Table 2. From the log plots of peak intensity versus delay time (Figure 8a), it is noted that three of the four peaks have multiple T_1^C components. The plots were fitted to two exponential decays as described above, and as a result, the ratio and T_1^C values of the shorter and longer relaxation time components ($T_1^C{}_{short}$ and $T_1^C{}_{long}$) were obtained. The results are summarized in Table 2 and Figure 8b. Only the peak at 63.5 ppm could be fitted by a single-exponential function. The relative ratios of β-sheet with two components for the main peak at 64.2 ppm are very similar to the previous T_1^C results from nonlabeled (AGSGAG)$_5$ peptide.[14]

From the previous results of T_1^C measurement with increasing temperature, it is known that the Ser-Cβ peaks have a mobility in the strong collision limit; that is, the components with longer T_1^C values show lower molecular mobility.[16] The peak at 63.5 ppm is a single component with the longest T_1^C value among the Ser-Cβ peaks (Table 2 and Figure 8b), which means that this peak has a rigid conformation possibly arising from strong hydrogen bonds through the Ser side chain OH group. In our previous paper on the ^{2}H solid-state NMR analysis of [3,3-^{2}H$_2$] Ser-labeled *B. mori* silk fiber,[5] the dominant side-chain χ_1 dihedral angle was found to be the *gauche*$^+$ conformation. Moreover, Fraser et al. have reported that the intersheet spacing became larger in the case of (AGSGAG)$_n$ than that in (AG)$_n$.[8] This is reasonable, since the

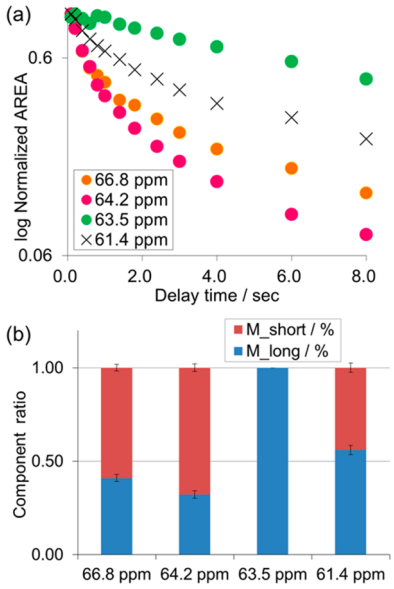


Figure 8. (a) Log plot of normalized area of each Ser-Cβ peak observed by T_1^C measurement of [3-^{13}C] Ser^{15}(AGSGAG)$_5$ via Torchia's sequence[22] versus delay time. (b) Stacked column chart of components' ratios of shorter and longer T_1^C values for each Ser-Cβ peak.

Table 2. T_1^C Relaxation Times of [3-^{13}C] Ser$_{15}$(AGSGAG)$_5$ Measured at 400 MHz

	66.8 ppm	64.2 ppm	63.5 ppm	61.4 ppm
I_{short}	0.59	0.68		0.44
I_{long}	0.41	0.32	1.00	0.56
$T_1^C{}_{short}$/s	0.32	0.45		0.73
$T_1^C{}_{long}$/s	4.92	4.43	10.75	8.43

side chain of Ser is bulkier than Ala, and so the replacement of Ala to Ser in the sequence is expected to distort the β-sheet packing. Moreover, in the previous section, the Ser-Cβ peak at 63.5 ppm assigned to β-sheet B domain showed only a weak correlation with the Ala-Cβ peak of β-sheet B in the 2D DARR spectrum at the long mixing time (Figure 5a), suggesting a relatively large distance between the two kinds of Cβ carbons in domain B. Taking these points together, we conclude that serine in domain B (the Ser-Cβ at 63.5 ppm) forms intermolecular hydrogen bonds most easily, thereby holding adjacent strands relatively rigidly and far apart.

On the other hand, in the main β-sheet peak at 64.2 ppm (the more highly populated β-sheet domain A), the longer T_1^C component was minor and the relaxation value was shortest (64.2 ppm in Table 2 and Figure 8b). In the 2D DARR spectrum at 400 ms mixing time (Figure 5a), the Ser-Cβ peak at 64.2 ppm showed a strong correlation with the Ala-Cβ peak of β-sheet A. These results suggest that Ser residues in domain A are relatively stable but do not form stable hydrogen bonds with other chains. Interestingly, the random coil peak showed more percentage of the longer T_1^C component. This shows that in the random coil structure the percentage of hydrogen-bonding conformation is larger than that of the main β-sheet A. It is worth noting that in a liquid NMR spectrum random coil structure usually implies an isotropically mobile component, whereas in solid-state NMR spectra of materials with high hydrogen-bonding networks, random coil structure does not always mean a mobile component.

dx.doi.org/10.1021/ma500908m | *Macromolecules* 2014, 47, 4308−4316

4. Domain Size of A and B β-Sheet Geometries. To place limits on the domain size of the β-sheets A and B, we measured ^{1}H spin−lattice relaxation times in the rotating frame ($T_{1\rho}^{H}$) and in the laboratory frame (T_1^{H}).[23,24,33,34] Spin diffusion is an efficient mechanism for cross-relaxation in the solid over distances of several nanometers. If domains are within a few nanometers of each other, spin diffusion will cause spin−lattice relaxation times to be averaged, whereas if domains are larger than this, different relaxation times can be expected for different domains. Each observed relaxation curve was observed as a single-exponential decay, and the obtained $T_{1\rho}^{H}$ and T_1^{H} values are tabulated in Table 3. All T_1^{H} values were 0.9 ± 0.1 s; there is no clear difference among the respective peaks. Similarly, the $T_{1\rho}^{H}$ of the [U-^{13}C] Cp-fraction was 19 ± 1 ms for the β-sheet peaks, whereas the random coil signal has a relatively short $T_{1\rho}^{H}$ of 12 ms. This result indicates that the random coil region is spatially distinct from the crystalline region but suggests that the domain size of A and B geometries must be no more than 5 nm. The ^{13}C CP/MAS spectrum of a short model peptide G(AG)$_3$ with ca. 2.1 nm length in a complete β-sheet shows only β-sheet peaks in the Ala-Cβ region.[10] However, the longer peptide (AG)$_9$ with ca. 6.0 nm length shows a distorted β-turn/random coil peak at the higher field side of the main Ala-Cβ peak.[10] These observations suggest that the estimated domain size of 5 nm is reasonable because the Cp fraction consists of both complete and distorted β-sheets. We have conducted detailed calculations of cross peak intensities in DARR spectra for different domain sizes. Calculated intensities turn out to be relatively insensitive to the domain size because relayed transfer of magnetization is efficient. Therefore, an upper limit of 5 nm is the most precise limit imposed by the data.

Table 3. T_1^{H} Relaxation Times Measured at 400 MHz and $T_{1\rho}^{H}$ Relaxation Times Measured with ^{1}H Spin-Locking Frequency of 64 kHz

			T_1^{H} (s) rt	$T_{1\rho}^{H}$ (ms) rt
Ala-Cβ	16.2 ppm	random coil/distorted β-turn	0.89	12.0
	19.6 ppm	β-sheet A	0.92	18.9
	21.7 ppm	β-sheet B	0.93	18.1
Ser-Cβ	61.4 ppm	random coil	0.92	11.7
	63.5 ppm	β-sheet B	0.95	18.5
	64.2 ppm	β-sheet A	0.91	19.5
	66.8 ppm	distorted β-sheet A	0.95	18.8

CONCLUSIONS

We have assigned the heterogeneous Ser peaks observed in the ^{13}C solid-state CP/MAS NMR spectrum of [U-^{13}C] Cp-fraction in silk II structure, which was obtained from [U-^{13}C] *B. mori* silk fiber. The Ser peaks comprise one random coil and three β-sheet components: A, B, and a perturbed A structure at around 67 ppm. Moreover, the 2D DARR spectrum of the [U-^{13}C] Cp-fraction with long mixing time suggested the close proximity of the two β-sheet domains within a few nanometers. This result was also supported from the ^{1}H spin−lattice relaxation measurements. From T_1^{C} measurement, it was suggested that the preferred conformation of the Ser side chain has comparatively little hydrogen bonding with other chains. Consequently, it is necessary to revise the standard model of *Bombyx mori* silk fibroin from the uniform antiparallel β-sheet proposed more than half a century ago. It is now clear that the crystalline region contains a heterogeneous mixture of two antiparallel geometries.

ASSOCIATED CONTENT

Supporting Information

Synthetic route from the commercial Ser-OH to Ser-OFmoc; the whole ^{13}C CP/MAS spectra of nonlabeled, [3-^{13}C]Ser^9, [3-^{13}C]Ser^{15}, and [3-^{13}C]Ser^{21} labeled (AGSGAG)$_5$. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Fax 81-42-383-7733; e-mail asakura@cc.tuat.ac.jp (T.A.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

T.A. acknowledges support from Grant-in-Aid for Scientific Research from Ministry of Education, Science, Culture and Supports of Japan (23245045, 25620169, 26248050) and Ministry of Agriculture, Forestry and Fisheries of Japan (Agri-Health Translational Research Project).

REFERENCES

(1) Yukuhiro, K.; Sezutsu, H.; Yonemura, N. Evolutionary Divergence of Lepidopteran and Trichopteran Fibroins. *Biotechnology of Silk*; Asakura, T., Miller, T., Eds.; Springer: Heidelberg, 2014; pp 25−47.
(2) Zhou, C.; Confalonieri, F.; Jacquet, M.; Perasso, R.; Li, Z.; Janin, J. *Proteins: Struct. Funct., Genet.* **2001**, *44*, 119.
(3) Asakura, T.; Ohgo, K.; Komatsu, K.; Kanenari, M.; Okuyama, K. *Macromolecules* **2005**, *38*, 7397.
(4) Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komatsu, K. *J. Mol. Biol.* **2001**, *306*, 7397.
(5) Asakura, T.; Suzuki, Y.; Yazawa, K.; Aoki, A.; Nishiyama, Y.; Nishimura, K.; Suzuki, F.; Kaji, H. *Macromolecules* **2013**, *46*, 8046.
(6) Marsh, R.; Corey, R. B.; Pauling, L. *Biochim. Biophys. Acta* **1955**, *16*, 134.
(7) Lotz, B.; Keith, H. D. *J. Mol. Biol.* **1971**, *61*, 201.
(8) Fraser, B.; MacRae, T. P. *Conformation of Fibrous Proteins and Related Synthetic Polypeptides*; Academic Press: New York, 1973; pp 293−343.
(9) Asakura, T.; Yao, J.; Yamane, T.; Umemura, K.; Ulrich, A. *J. Am. Chem. Soc.* **2002**, *124*, 8794.
(10) Asakura, T.; Yao, J. *Protein Sci.* **2002**, *11*, 2706.
(11) Takegoshi, K.; Nakamura, S.; Terao, T. *Chem. Phys. Lett.* **2001**, *344*, 631.
(12) Saito, H.; Tabeta, R.; Kuzuhara, A.; Asakura, T. *Bull. Chem. Soc. Jpn.* **1986**, *59*, 3383.
(13) Saito, H.; Ishida, M.; Yokoi, M.; Asakura, T. *Macromolecules* **1990**, *23*, 8387.
(14) Suzuki, Y.; Aoki, A.; Nakazawa, Y.; Knight, D. P.; Asakura, T. *Macromolecules* **2010**, *43*, 9434.
(15) Yamane, T.; Umemura, K.; Asakura, T. *Macromolecules.* **2002**, *35*, 8831.
(16) Kameda, T.; Ohkawa, Y.; Yoshizawa, K.; Naito, J.; Ulrich, A. S.; Asakura, T. *Macromolecules* **1999**, *32*, 7166.
(17) Zhao, C.; Asakura, T. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 301.
(18) Asakura, T.; Suzuki, Y.; Nakazawa, Y.; Yazawa, K.; Holland, G. P.; Yarger, J. L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *69*, 2368.
(19) Asakura, T.; Kuzuhara, A.; Tabeta, R.; Saito, H. *Macromolecules* **1985**, *18*, 1841.

(20) Wang, J.; Okada, Y.; Li, W.; Yokoi, T.; Zhu, J. *J. Chem. Soc., Perkin Trans.* **1997**, *1*, 621.
(21) Asakura, T.; Yao, J.; Ohgo, K.; Sugino, R.; Raghuvansh, K. *Biomacromolecules* **2004**, *5*, 1763.
(22) Torchia, J. *J. Magn. Reson.* **1978**, *30*, 1969.
(23) Stejskal, E. O.; Schaefer, J.; Sefcik, M. D.; McKay, R. A. *Macromolecules* **1981**, *14*, 275.
(24) Asano, A.; Takegoshi, K. Polymer Blends and Miscibility. In *Solid State NMR of Polymers*; Ando, I., Asakura, T., Eds.; Elsevier Science B.V.: Amsterdam, 1988; pp 351−414.
(25) Takahashi, Y.; Gehoh, M.; Yuzuriha, K. *Int. J. Biol. Macromol.* **1999**, *24*, 127.
(26) Asakura, T.; Iwadate, M.; Demura, M.; Williamson, M. P. *Int. J. Biol. Macromol.* **1999**, *4*, 167.
(27) Demura, M.; Minami, M.; Asakura, T.; Cross, T. A. *J. Am. Chem. Soc.* **1998**, *120*, 1300.
(28) Asakura, T.; Watanabe, Y.; Uchida, A.; Minagawa, H. *Macromolecules* **1984**, *17*, 1075.
(29) Asakura, T.; Ohgo, K.; Ishida, T.; Taddei, P.; Monti, P.; Kishore, R. *Biomacromolecules* **2005**, *6*, 468.
(30) Egawa, A.; Fujiwara, T.; Mizuguchi, T.; Kakitani, Y.; Koyama, Y.; Akutsu, H. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 790.
(31) Iwadate, M.; Asakura, T.; Williamson, M. P. *J. Biomol. NMR* **1999**, *13*, 199.
(32) Shen, Y.; Bax, A. *J. Biomol. NMR* **2013**, *56*, 227.
(33) VanderHart, D. L.; McFadden, G. B. *Solid State Nucl. Magn. Reson.* **1996**, *7*, 4566.
(34) Demco, D. E.; Johansson, A.; Tegenfeldt, J. *Solid State Nucl. Magn. Reson.* **1995**, *4*, 1338.