DOI: 10.1021/ma1018878

Macromolecules
ARTICLE

Structural Analysis of the Synthetic Peptide (Ala-Gly-Ser-Gly-Ala-Gly)$_5$, a Model for the Crystalline Domain of *Bombyx mori* Silk Fibroin, Studied with ^{13}C CP/MAS NMR, REDOR, and Statistical Mechanical Calculations

Yu Suzuki,[†] Akihiro Aoki,[†] Yasumoto Nakazawa,[†] David P. Knight,[‡] and Tetsuo Asakura*,[†]

[†]*Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan, and* [‡]*Oxford Biomaterials Ltd., Department of Zoology, Oxford University, South Parks Road, Oxford OX1 3PS, U.K.*

Received August 17, 2010; Revised Manuscript Received October 11, 2010

ABSTRACT: In our previous study, we have proposed a lamellar structure for Ala-Gly repeated copolymeric peptide, a model for crystalline region of *Bombyx mori* silk fibroin. Here, we propose the structure of (AGSGAG)$_5$ with silk II form which is a more mimic of the crystalline region of *B. mori* silk fibroin than (AG)$_{15}$. The local structure for each Ala residue was determined from ^{13}C CP/MAS NMR spectra of 10 different [3-^{13}C]Ala-(AGSGAG)$_5$ peptides differing in their ^{13}C labeling positions. The highest field peak for the Ala Cβ carbon (16.7 ppm) assigned to a distorted β-turn structure and/or random coil changes significantly depending on the ^{13}C labeling position. In addition, the fractions of the random coil and/or distorted β-turn component of each Ser residue were determined by REDOR experiments from the ^{13}C−^{15}N atomic distances of five versions of the above peptide with different [1-^{13}C]Gly-Ser-[^{15}N]Gly positions. By combining the structural information of Ala and Ser residues from solid state NMR, with statistical mechanical calculation previously used for (AG)$_{15}$, the probable lamellar structures of (AGSGAG)$_5$ in the solid state are proposed. The models of two turns in the central part of the sequence of (AGSGAG)$_5$ consist of approximately 8−12 amino acids. The effect of the introduction of Ser residue on the local structure of Ala-Gly copolymeric peptides is also discussed on the basis of the evidence from ^{13}C solid state spin−lattice relaxation experiments.

Introduction

Silks are fibrous proteins with properties that intrigue scientists ranging from polymer chemists to biomedical researchers.[1] *Bombyx mori* (*B. mori*) produces silk fibroin filaments with outstanding mechanical properties including useful strength and toughness, despite being spun from the aqueous solution at ambient temperature and pressure. Silk fibroin has good biocompatibility, resistance to thermal and enzymatic degradation, and mechanical resilience, and it can be formed into a wide range of forms including nano- and microparticles, films, foams, sponges, and composites. This portfolio of properties suggests considerable potential in medicine and surgery.[2]

The silk fibroin has two distinct structures in the solid state, silk I (before spinning) and silk II (after spinning). The conformation of silk I has been shown to have a repeated type II β-turn structure using solid state NMR approaches such as 2D spin diffusion NMR under off magic angle spinning, rotational echo double resonance (REDOR), and conformation-dependent ^{13}C chemical shift data as well as X-ray diffraction analysis.[3,4] Concerning the structure of silk II, Marsh et al.[5] were the first to propose an antiparallel β-sheet model based on a fiber diffraction study of native *B. mori* silk fiber. Even though the local protein conformation is still the basic β-sheet as proposed by Marsh et al., the refined silk II model incorporates stacking of the β-sheet planes in two different arrangements.[6] Synthetic peptides (AG)$_n$ have been used in spectroscopic studies as a model system for analysis of the crystalline domain of *B. mori* silk fibroin. A detailed chain-folded lamellar structure has been proposed for (AG)$_{64}$ in the silk II form on the basis of SAX and WAX scattering for a dried hydrogel of regenerated silk fibroin slowly converted to a silk II structure.[7]

Solid state NMR has also been used to clarify the silk II structure of (AG)$_{15}$.[8−10] Broad and asymmetric Ala Cβ peaks were observed in the CP/MAS NMR spectra, indicating a heterogeneous structure in contrast to the single ordered secondary structure of silk I. The relative proportions of the various heterogeneous components were determined from their relative peak intensities of the deconvoluted line shape. For example, the deconvoluted Ala Cβ peaks yielded 27% *distorted* β-turn (16.7 ppm), 46% β-sheet (alternating Ala residues) (19.1 ppm), and 27% β-sheet (parallel Ala residues) (22.4 pppm).[9] Moreover, ^{13}C solid-state NMR on selectively labeled peptides labeled singly with ^{13}C at different Ala methyl carbons was used to examine the silk II structure of (AG)$_{15}$ indicating a lamellar structure containing β-turns.[11] We also used 2D spin diffusion NMR technique under off magic angle spinning condition, REDOR, and ^{13}C chemical shift contour plots of the Cβ carbon of Ala residue to study the lamellar structure in detail, as well as statistical mechanical calculations.[12] Thus, it has been shown that high resolution solid state NMR study coupled with selective stable isotope labeling of the peptide samples can provide a great deal of information on the detailed structures of silk fibroins. The lamellar structure of (AG)$_{15}$ revealed in this way is in good agreement with that of (AG)$_{64}$ elucidated by X-ray diffraction.[7]

The primary structure of the heavy chain of *B. mori* silk fibroin contains multiple repetitions of (G-A-G-A-G-S) which make up

*To whom correspondence should be addressed. E-mail: asakura@cc.tuat.ac.jp.

 Published on Web 10/21/2010

Table 1. Stable-Isotope Labeled Peptides of $(AGSGAG)_5$ for ^{13}C CP/MAS and REDOR Experiments

sample no.	peptide	method
Ala1	[3-^{13}C]AGSGAG AGSGAG AGSGAG AGSGAG AGSGAG	^{13}C CP/MAS
Ala5	AGSG[3-^{13}C]AG AGSGAG AGSGAG AGSGAG AGSGAG	^{13}C CP/MAS
Ala7	AGSGAG [3-^{13}C]AGSGAG AGSGAG AGSGAG AGSGAG	^{13}C CP/MAS
Ala11	AGSGAG AGSG[3-^{13}C]AG AGSGAG AGSGAG AGSGAG	^{13}C CP/MAS
Ala13	AGSGAG AGSGAG [3-^{13}C]AGSGAG AGSGAG AGSGAG	^{13}C CP/MAS
Ala17	AGSGAG AGSGAG AGSG[3-^{13}C]AG AGSGAG AGSGAG	^{13}C CP/MAS
Ala19	AGSGAG AGSGAG AGSGAG [3-^{13}C]AGSGAG AGSGAG	^{13}C CP/MAS
Ala23	AGSGAG AGSGAG AGSGAG AGSG[3-^{13}C]AG AGSGAG	^{13}C CP/MAS
Ala25	AGSGAG AGSGAG AGSGAG AGSGAG [3-^{13}C]AGSGAG	^{13}C CP/MAS
Ala29	AGSGAG AGSGAG AGSGAG AGSGAG AGSG[3-^{13}C]AG	^{13}C CP/MAS
Ser3	A[1-^{13}C]GS[^{15}N]GAG AGSGAG AGSGAG AGSGAG AGSGAG	REDOR
Ser9	AGSGAG A[1-^{13}C]GS[^{15}N]GAG AGSGAG AGSGAG AGSGAG	REDOR
Ser15	AGSGAG AGSGAG A[1-^{13}C]GS[^{15}N]GAG AGSGAG AGSGAG	REDOR
Ser21	AGSGAG AGSGAG AGSGAG A[1-^{13}C]GS[^{15}N]GAG AGSGAG	REDOR
Ser27	AGSGAG AGSGAG AGSGAG AGSGAG A[1-^{13}C]GS[^{15}N]GAG	REDOR

approximately half of the total silk fibroin.[13] Then, the aim of the present paper is to describe the solid-state structure of the sequential peptide model, $(AGSGAG)_5$, a better model for the crystalline region of *B. mori* silk fibroin than $(AG)_{15}$. The effect of introduction of Ser residues on the lamella structure of an Ala-Gly copolymer will be emphasized. The fractions of the random coil and/or distorted β-turn component of each Ala residue were determined using ^{13}C CP/MAS NMR spectra of versions of [3-^{13}C]Ala-$(AGSGAG)_5$ with 10 different ^{13}C labeling positions. In addition, the fractions of the random coil and/or distorted β-turn component of each Ser residue were determined from the $^{13}C-^{15}N$ atomic distances of five versions of the above peptide with different [1-^{13}C]Gly-Ser-[^{15}N]Gly positions determined by REDOR experiments. The probable structures of the silk II form of $(AGSGAG)_5$ in the solid state were proposed by combining the structural information of Ala and Ser residues from solid state NMR, and statistical mechanical calculations. The ^{13}C solid state spin−lattice relaxation times were observed to obtain information on the effect on the local structure and chain mobility of introducting Ser residue into an Ala-Gly copolymer.

Experimental Section

Sample Preparations. The 10 [3-^{13}C]Ala-sequential peptides, $(AGSGAG)_5$, and five peptides with different [1-^{13}C]Gly-Ser-[^{15}N]Gly positions together with natural abundance peptide as listed in Table 1 were synthesized by the solid-phase method as described elsewhere.[4] After synthesis, the samples were dissolved in 9 M LiBr aqueous solution and then dialyzed against distilled water for three days at 4 °C. The precipitate was then collected and dried for solid-state NMR observation.

^{13}C CP/MAS NMR Measurements. ^{13}C CP/MAS NMR experiments[14] were performed on a Chemagnetics Infinity 400 MHz spectrometer with an operating frequency of 100.0 MHz for ^{13}C at a sample spinning rate of 7 kHz using a 4 mm diameter Zr rotor. We used the following: 10240 scans for nonlabeled sample and 1,024 scans for [3-^{13}C]Ala labeled samples, 1 ms CP contact time and 5 s recycle delay. 50 kHz radio frequency field strength was used for $^1H-^{13}C$ decoupling with the acquisition period of 12.8 ms. ^{13}C chemical shifts were calibrated indirectly through the adamantane methylene peak observed at 28.8 ppm relative to TMS (tetramethylsilane) at 0 ppm. For performing the deconvolution of Ala Cβ peak, we initially assumed the chemical shift and line width of the three components to be fixed. But thereafter we changed the value of these parameters slightly to obtain a closer simulation of the observed spectra. Gaussian line shape was assumed in all simulations.

REDOR Measurements. ^{13}C-detected $^{13}C-^{15}N$ REDOR experiments[15] were performed on a Bruker DSX-400 AVANCE spectrometer operating at 400.1 MHz for ^{1}H, 100.6 MHz for ^{13}C and 40.5 MHz for ^{15}N using a BL 4 mm Bruker MAS probe. The MAS spinning speed was 7 kHz and the recycle delay was set to 5 s. π pulses for ^{13}C and ^{15}N channels were 16.0 and 12.8 μs, respectively and a total of 512 acquisitions were collected. REDOR evolution times ranged up to 24.3 ms. The REDOR results were corrected for the natural abundance background by calculation. Values of $\Delta S/S_0 = 1 - S/S_0$, where S_0 and S represent the ^{13}C signals obtained respectively in the absence and presence of dipolar dephasing, were computed as the ratios of peak intensities in the REDOR spectra.

PDB Analysis. X-ray crystallographic data from the Protein Data Bank at the Research Collaboratory for Structural Bioinformatics (RCSB) were used to search the torsion angles of Gly-Ser-Gly sequences in natural proteins.[16] Only structures determined at 2.0 Å resolution or better and R factor $\leq$20% were used. A subset of 310 occurrences was obtained from the database after excluding multiple entries of proteins with a similarity greater than 50%. Then, the atomic distance between the carbonyl carbon of the first Gly residue and the amide nitrogen of the third Gly residue in the Gly-Ser-Gly sequence was calculated.[17]

^{13}C Spin−Lattice Relaxation Time (T_1) Measurements. ^{13}C spin−lattice relaxation times were observed using a Bruker DSX-400 AVANCE spectrometer. We used a pulse sequence for ^{13}C T_1 determination developed by Torchia.[18] The MAS spinning speed was 8 kHz. The ^{1}H 90° pulse, ^{13}C 90° pulse, contact and repetition times were 3.9 μs, 5 μs, 1.2 ms, and 5 s, respectively. Spectra were accumulated 5120 times with arrayed variable delay range of 0.05−8.0 s at 24, 40, 60, and 80 °C.

Results

^{13}C CP/MAS NMR Spectra of $(AGSGAG)_5$. Figure 1 shows the ^{13}C CP/MAS NMR spectra of (a) [3-^{13}C]Ala^{17th}-$(AGSGAG)_5$, (b) nonlabeled $(AGSGAG)_5$ and (c) difference spectrum (a − b) as an example. In general, the Cα and Cβ carbons in proteins and peptides exhibit conformation-dependent ^{13}C chemical shifts.[19−21] These chemical shifts have been used for structural analyses of silk fibroins by us.[22−25] The Ala Cβ peak of Ala^{17th}-$(AGSGAG)_5$ obtained from samples precipitated by dialysis of 9 M LiBr solution of the peptide, was asymmetric and broad as shown in Figure 1. This spectral pattern indicates a silk II structure although the sample, $(AG)_{15}$ prepared by the same treatment has been reported to be a silk I.[3,4] Thus the introduction of repeated Ser residues in the sequence $(AG)_{15}$ to give $(AGSGAG)_5$ is evidently responsible for its structural change from silk I to silk II during precipitation. The Ala Cβ peak can be deconvoluted into three components reflecting the local conformation and intermolecular arrangements of the chains as mentioned in the introduction. The highest field broad peak at around 16.7 ppm was assigned to a distorted β-turn structure, which is characterized by a large distribution in torsion angles around an average conformation of a type II

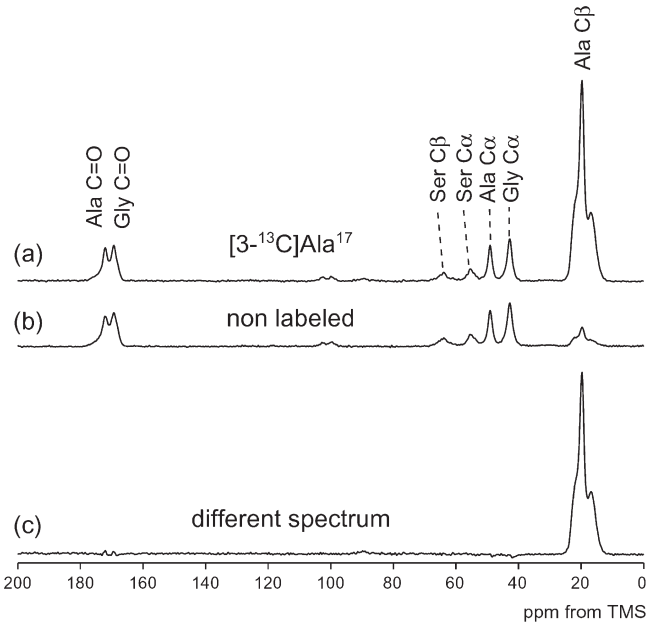

Figure 1. ^{13}C CP/MAS NMR spectra of (a) $(AGSGAG)_2$Ala-Gly-Ser-Gly-[3-^{13}C]Ala^{17}-Gly$(AG\text{-}SGAG)_2$ and (b) natural abundance $(AGSGAG)_5$ and (c) the difference spectrum, spectrum a − spectrum b.

β-turn. The chemical shift of this broad peak is also in agreement with that of random coil peak of the Cβ carbon of Ala residues of proteins in aqueous solution as discussed previously.[9,10] In order to determine quantitatively the fraction of conformation assigned to the 16.7 ppm peak, we obtained the difference spectrum c by subtracting natural abundance spectrum b from the ^{13}C-labeled peptide spectrum a. This subtraction (a − b) provides local structural information exclusively for selectively ^{13}C labeled Ala residues.

Determination of the Fraction of Distorted Type II β-Turn and/or Random Coil for Ala Residues from the Relative Intensity of 16.7 ppm Peak in [3-^{13}C]Ala-$(AGSGAG)_5$ with Different ^{13}C Labeling Positions. Figure 2 shows the difference Ala Cβ spectra of 10 [3-^{13}C]Ala-$(AGSGAG)_5$ peptides, with different labeling position together with the deconvolutions assuming Gaussian distributions. Previously we reported ^{13}C T_1 values of three components of Ala $^{13}CH_3$ peak. The values were 550 ms (21.9 ppm), 270 ms (19.6 ppm) and 430 ms (16.5 ppm).[8] Thus, the highest field peak which was assigned to random coil or distorted β-turn had a T_1 value between that of the other two components. In addition, we observed ^{13}C CPMAS and ^{13}C DDMAS Ala Cβ spectra, noting that the relative intensities of these three components are the same between two kinds of spectra within experimental error. We therefore, used ^{13}C CPMAS spectra to obtain the fraction of β-sheet and random coil.

The fraction corresponding to the 16.7 ppm component was determined to be 75% for the Cβ carbon of Ala^1 in the peptide indicating that the Ala residue at the start of the sequence has a high probability of forming random coil. The fraction was markedly lower (26%) at Ala^5 but rose to a second maximum (44%) at Ala^{19}. The fractions for the Ala residues located centrally in the peptide (Ala^{11}, Ala^{13}, Ala^{17}, and Ala^{19}) were higher than those closer to the ends (Ala^5, Ala^7, Ala^{23}, and Ala^{25}). It was also noted that in each of the four cases, the two Ala residues located between two Ser residues in the sequence, −SGAGAGS− in the peptide showed similar fractions of the 16.7 ppm component. Ala^{29} adjacent to the C-terminal Gly residue showed a higher fraction (33%) compared with the two preceding Ala residues

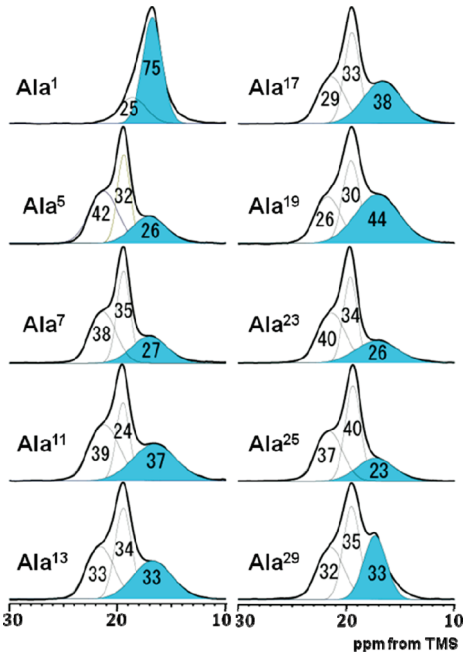

Figure 2. Fraction (%) of the 16.7 ppm peak calculated from the deconvolution of Ala Cβ peaks in the ^{13}C CP/MAS NMR spectra of 10 [3-^{13}C]$(AGSGAG)_5$ with different [3-^{13}C]Ala labeling sites as a function of the ^{13}C labeled position.

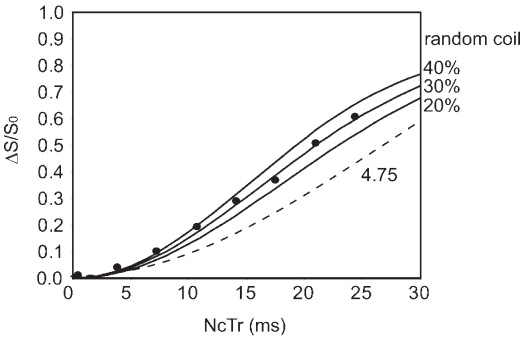

Figure 3. Observed and calculated REDOR plots of $(AGSGAG)_2$Ala-[1-^{13}C]Gly-Ser-[^{15}N]Gly- AlaGly$(AGSGAG)_2$ named Ser15. The solid lines represent the REDOR curves calculated by changing the fraction of random coil (3.80 Å) and β-sheet (4.75 Å) forms and the fractions of random coil are labeled. The broken line corresponds to β-sheet (4.75 Å).

Ala^{23} (26%) and Ala^{25} (23%). This probably results from a higher probability of the residue adjacent to the C-terminus to form random coil.

Determination of the Fraction of Distorted Type II β-Turn and/or Random Coil for the Ser Residues by the REDOR Method. To investigate the whole conformation of $(AGSGAG)_5$ in silk II form, it is necessary to determine the local conformation of Ser residues as well as Ala residues. In this study, the information was obtained from the inter nuclear atomic distances between [1-^{13}C]Gly and [^{15}N]Gly nuclei in the sequence, [1-^{13}C]Gly-Ser-[^{15}N]Gly, indirectly because the distance reflect the local conformation of the Ser residue between two Gly residues. REDOR was used for the purpose. The experimental $\Delta S/S_0$ values were plotted against dipolar evolution time for $(AGSGAG)_2$Ala[1-^{13}C]GlySer-[^{15}N]Gly-AlaGly$(AGSGAG)_2$ (see Figure 3 for an example) and some of the plots have already been reported previously.[26] As a starting point, the theoretical REDOR curve was calculated by assuming the distance, between [1-^{13}C]Gly and [^{15}N]Gly nuclei in the sequence, [1-^{13}C]Gly-Ser-[^{15}N]Gly

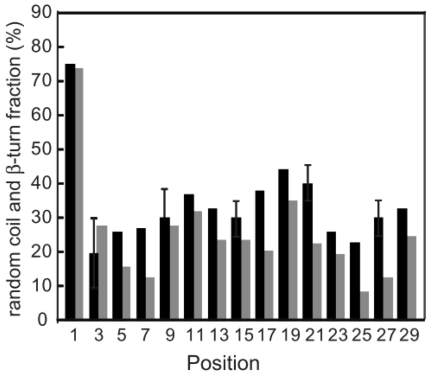

Figure 4. Observed relative intensities (black) of distorted β-turn and/or random coil structure component of $(AGSGAG)_5$ obtained from the peak at 16.7 ppm of Ala Cβ and REDOR bimodal fitting for Ser residues were shown with different position. The observed relative intensities (gray) of distorted β-turn and/or random coil structure component of $(AG)_{15}$ were also shown for the comparison purpose.

of 4.75 Å as in the antiparallel β-sheet structure obtained from crystal structure of *B. mori* silk fibroin determined by Takahashi et al.[6] However, the curve could not reproduce the observed REDOR curve, indicating that the conformation of the Ser residue was a mixture of random coil and antiparallel β-sheet structures as expected. It was therefore necessary to obtain the information on the internuclear atomic distance between [1-^{13}C]Gly and [^{15}N]Gly nuclei in the sequence, [1-^{13}C]Gly-Ser-[^{15}N]Gly when the Ser residue takes random coil conformation. As described in Experimental Section (PDB analysis), 310 torsion angles of the Ser residues in natural proteins selected from PDB database, and the corresponding distances between the carbonyl carbon of the first Gly residue and the amide nitrogen of the third Gly residue in the Gly-Ser-Gly sequences were used. The atomic distance for random coil conformation was calculated to be 3.80 Å by averaging these distances. By assuming the fraction of antiparallel β-sheets and random coil, the REDOR curve was calculated. The theoretical REDOR curve of the fraction, 30% random coil and 70% antiparallel β-sheet gave the best-fit as shown in Figure 3. Thus, the fraction of random coil for Ser^{15th} residue was determined to be $30 \pm 5\%$ and the same fitting process was performed for the remaining four peptides.

Change in the Fractions of the Distorted β-Turn and/or Random Coil Components with Residue Position. The fractions of the distorted β-turn and/or random coil components were determined from the relative intensities of [3-^{13}C]Ala Cβ peaks at 16.7 ppm for Ala residues and for Ser residues from the REDOR curves of the [1-^{13}C]Gly-Ser-[^{15}N]Gly labeled versions. Figure 4 shows these data (black) plotted against the residue position within $(AGSGAG)_5$ and corresponding plot (gray) for $(AG)_{15}$ reported previously.[11,12] Comparison of the plots for $(AG)_{15}$ and $(AGSGAG)_5$ show that on the whole, the fractions of the distorted β-turn and/or random coil component were higher in $(AGSGAG)_5$ except for the two Ala residues at the N-terminal sites. The way the fractions varied with position was slightly different in the two peptides.

In the case of $(AG)_{15}$ the fraction decreased gradually toward the inner part of the chain. The fraction then increased at both the positions 9 and 11, indicating the appearance of the folded lamellar structure with a β-turn at these positions. After the 11th position, the fraction decreased slightly. Moving from the 13th to the 17th position, the fraction of the peak at 16.7 ppm was constant before reaching a new maximum at the 19th position. Then the fraction decreased again. Toward the C-terminal, the fraction increased again. Thus, we proposed lamella structure with β-sheet structure and distorted β-turn at the 11th and 19th Ala residues for $(AG)_{15}$.[11,12] In the case of $(AGSGAG)_5$, the general features of the plot were similar to that of $(AG)_{15}$: More than 70% non-β-sheet component at the first Ala residue and two maxima at around the 11th and 19th Ala residues. However, the fractions of the distorted β-turn and/or random coil component were generally higher throughout for $(AGSGAG)_5$ except for Ala^3 close to the N-terminal site. The plot of $(AGSGAG)_5$ suggests the appearance of the folded lamellar structure with two β-turns, one at about residue 11 and the other at about residue 19 like $(AG)_{15}$ although the position of the β-turns may be less sharply defined in $(AG)_{15}$. Thus, the presence of Ser residues in the peptide chain may result in a slight but significant perturbation of the lamella structure seen in $(AG)_{15}$.

Statistical Mechanical Calculation. We used statistical mechanical calculation[27] to investigate possible lamella structures in $(AGSGAG)_5$. This approach has been successfully applied to $(AG)_{15}$ for the same purpose.[12] The calculations were initially based on the following assumptions:

(1) After β-turn formation along the chain there is at least one pair of the intramolecular hydrogen-bonded strands forming a β-sheet structure. Therefore, of the Ala residues, the first residue which can form the β-turn is the third, while the 29th Ala is unable to form the β-turn. (2) The direction of the β-turn formation along the chain is always from the N-terminal to the C-terminal. (3) There are one or two turns in the peptide. Under these assumptions, 147 possible structures were produced with one or two β-turn positions in one chain of $(AGSGAG)_5$. The occurrence probability, $p(i)$, of the ith structure for $(AGSGAG)_5$ was calculated as

$$p(i) = \exp(-\Delta E(i)/kT)$$

where $\Delta E(i)$ is the potential energy of the ith structure, k is the Boltzmann's constant, and T is the absolute temperature. The relative potential energy of the ith structure, $\Delta E(i)$ is estimated with the stabilization energy, ΔE_H, due to the formation of inter-residue hydrogen bonds for each residue with β-sheet structure ΔE_S, due to Ser residues incorporated in β-sheet and the destabilization energy, and ΔE_T, due to formation of turn structure is calculated as

$$\Delta E(i) = n_H(i)\Delta E_H + n_S(i)\Delta E_S + n_T(i)\Delta E_T$$

where $n_H(i)$, $n_S(i)$, and $n_T(i)$ are the numbers of hydrogen bonds, Ser residues incorporated in β-sheet and turns in the ith structure, respectively.

Then the occurrence probability, $P(j)$, where the Ala and Ser residue j in the $(AGSGAG)_5$ molecules do not contribute to the intermolecular hydrogen-bonding formation within a β-sheet structure is calculated as

$$P(j) = \Sigma\delta(ij)\exp(-\Delta E(i)/kT)/\Sigma\exp(-\Delta E(i)/kT)$$

where $\delta(ij) = 0$ when the Ala and Ser residue j in the ith structure is involved in the inter-residue hydrogen bonds with β-sheet structure, while $\delta(ij) = 1$ when the same Ala and Ser is not involved in the inter-residue hydrogen bonding formation with β-sheet structure. The values of ΔE_H, ΔE_S, and ΔE_T were optimized with the minimum root-mean-square deviation of observed and calculated $P(j)$.

The fraction of non-β-sheet conformations calculated with optimized ΔE_H, ΔE_S, and ΔE_T was plotted against the residue position (Figure 5a). However, agreement between the

observed and calculated plots was not satisfied compared to the same analysis for $(AG)_{15}$. Therefore, the calculations were performed again by changing the assumption slightly as follows: The residues incorporated into β-sheet structure next to the β-turn positions form one hydrogen bond instead of two. Other assumptions are the same as those described above. Under these assumptions, 147 possible structures with one or two β-turn positions in one chain were taken into account. Figure 6 shows the contour plots to obtain ΔE_H, ΔE_S and ΔE_T from the minimum root-mean-square deviation of observed and calculated $P(j)$. The values for the energies, $\Delta E_H = -0.35$, $\Delta E_S = -0.10$, and $\Delta E_T = 0.90$ kcal/mol (1 cal = 4.183 J) were obtained from the optimization. With optimized ΔE_H, ΔE_S, and ΔE_T, the fraction of non-β-sheet conformations calculated was plotted against the residue position (Figure 5b). A considerably better agreement between the observed and calculated plots was obtained compared with the first calculation.

^{13}C Solid State Spin−Lattice Relaxation Time Observation. In order to obtain information on the dynamics of Ser residue introduced in the peptide $(AG)_n$ experimentally, ^{13}C spin−lattice relaxation time of Ser Cβ carbons was observed for $(AGSGAG)_5$. A plot of a series of partially relaxed spectra of natural abundance $(AGSGAG)_5$ sample is shown with delay times in Figure 7. The experimental decay curve, represented by solid circles, is indicated not to be a single exponential and it has been resolved into plural exponentials with different T_1 values by the least-squares method. Since the two components agreed well with the observed data and also two components have been previously reported for the relaxation data for the Ser Cβ carbons of *B. mori* silk fibroin,[28] it was concluded that there exist two components with different T_1 values. The two components are given as

$$M(\tau)/M(0) = \zeta \exp(-\tau/T_1{}^{L}) + (1-\zeta)\exp(-\tau/T_1{}^{S})$$

where $T_1{}^{L}$ and $T_1{}^{S}$ denote that longer and shorter T_1, respectively, and ζ denotes the fraction of Ser Cβ with $T_1{}^{L}$ and $(1-\zeta)$ denotes the fraction of Ser Cβ with $T_1{}^{S}$. The effect of temperature on the behaviors of the two Ser Cβ T_1 components can be used to discover whether both lie within the slow mobility region on the plot of T_1 intensity versus correlation time.[29] Our results are summarized in Table 2. Over the temperature range 24−80 °C, the fraction of $T_1{}^{L}$ relative to $T_1{}^{S}$, denoted by ζ, were mostly constant at approximately 60%. Both T_1 values decreased with a rise in temperature indicating that both T_1 components of Ser Cβ carbons lay within the slow mobility region on the plot of T_1 intensity versus correlation time and that $T_1{}^{L}$ component has a relatively slower mobility than the $T_1{}^{S}$ component. The existence of two different T_1 components clearly shows that Ser residues exist in two states with different molecular dynamics. Curve fitting suggests 40% of the Ser residues in $(AGSGAG)_5$ are in the $T_1{}^{L}$ less mobile state. The lower mobility of the $T_1{}^{L}$ component compared with the $T_1{}^{S}$ is thought to result from a combination of intermolecular hydrogen bonding through the Ser-OH groups and high steric hindrance of the Ser side chains that are located between two β-sheet planes while the $T_1{}^{S}$ state is that of the more mobile non-hydrogen bonded Ser side chains.

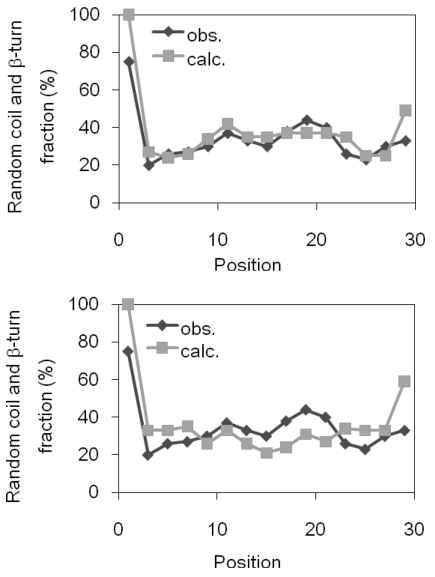

Figure 5. Observed (black) and calculated relative intensities (gray) of the β-turn and/or random coil structure component with different position. (a) The first calculation was performed according to the similar method applied to $(AG)_{15}$ reported previously.[12,13] The agreement between the observed and calculated plots was not satisfied compared to the same analysis for $(AG)_{15}$. (b) The revised second calculation was performed by changing the assumption slightly as follows: The residues incorporated into β-sheet structure next to the β-turn positions form one hydrogen bond instead of two. The agreement between the observed and calculated plots was satisfied.

Discussion

We show that $(AGSGAG)_5$ when precipitated by dialysis of a 9 M LiBr solution of the peptide, took a silk II structure while $(AG)_{15}$ treated in the same way has been shown to be a silk I. However introducing a single serine residue into the central position of $(AG)_{15}$ to give $(AG)_7SG(AG)_7$ does not result in a silk II structure after the same treatment; it stays in silk I.[10] Thus the introduction of several regularly arranged serines into $(AG)_{15}$ is required to obtain a silk II structure indicating that this arrangement is required to destabilize the silk I form.

The fractions of the distorted β-turn and/or random coil components were determined from the relative intensities of

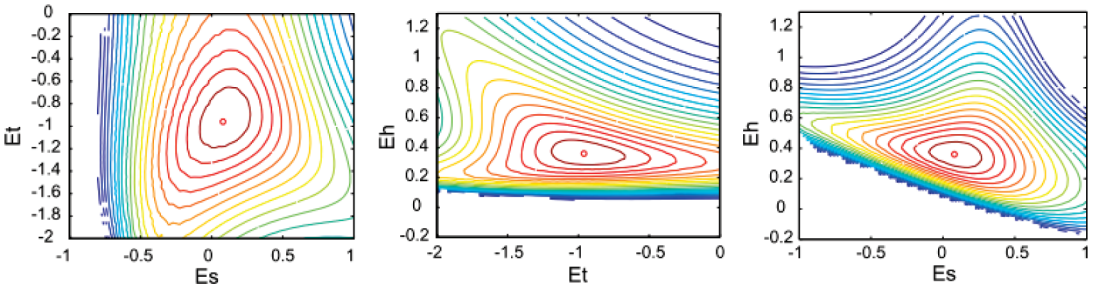

Figure 6. Contour plots used to obtain ΔE_H, ΔE_S, and ΔE_T from the minimum root-mean-square deviation in the calculations of observed and calculated the occurrence probability, $P(j)$. The values for the energies, $\Delta E_H = -0.35$, $\Delta E_S = -0.10$, and $\Delta E_T = 0.90$ kcal/mol (1 cal = 4.183 J) were obtained from the optimization. Details are described in the text.

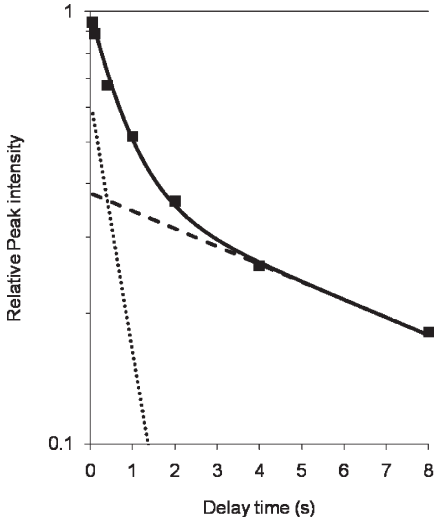

Figure 7. Observed plot (solid squares) of a series of partially relaxed ^{13}C peak intensities for Ser Cβ in natural abundance(AGSGAG)$_5$ vs delay time. A double exponential function gave a good fit to the experimental results ($R^2 = 0.98$). After fitting, the longer T_1 (dashed line), shorter T_1 (dotted line), and the fraction of Ser Cβ with each T_1 were determined.

Table 2. Shorter and Longer ^{13}C-T_1 Relaxation Times and Fraction (%) of each T_1 Component of Ser Cβ Carbon at Variable Temperatures

	24 °C	40 °C	60 °C	80 °C
shorter T_1	0.74 (60%)	0.72 (60%)	0.25 (70%)	0.16 (60%)
longer T_1	10.60 (40%)	8.16 (40%)	6.90 (30%)	2.12 (40%)

[3-^{13}C]Ala Cβ peaks at 16.7 ppm for Ala residues and from the REDOR curves of [1-^{13}C]Gly-Ser-[^{15}N]Gly parts for Ser residues. Comparison of the plots for (AG)$_{15}$ and (AGSGAG)$_5$ show that on the whole, the fractions of the distorted β-turn and/or random coil component were higher in (AGSGAG)$_5$ except for the two Ala residues at the N-terminal sites. Close examination of our previous studies[10] on ^{13}C CP/MAS NMR of the Ala Cβ region of nonisotope labeled (AG)$_n$ and (AGSGAG)$_m$ shows that the fraction of random coil/distorted β-turn region in Ala Cβ peaks was slightly higher in (AGSGAG)$_m$ than in (AG)$_n$. These observations taken together indicate that the presence of repeated Ser residues in the peptide chain may result in some disturbance to the β-sheet-distorted β-turn lamella structure. In this connection it is interesting to note that Fraser et al.[30,31] reported that unit cell dimensions (a, b, c) of (AG)$_n$ and (AGAGSG)$_m$ are 9.42, 6.95, 8.87 Å and 9.39, 6.85, 9.05 Å respectively. In the c-axis, the direction perpendicular to the plane of the pleated sheet the unit cell in (AGAGSG)$_m$ was 9.05 Å, markedly larger than the 8.87 Å in (AG)$_n$. This data strongly supports our hypothesis that side chains of Ser residues in (AGSGAG)$_5$ stick out perpendicular to the plane of the β-sheet producing a larger c-spacing compared to (AG)$_{15}$. We suggest that this larger separation allows looser, less regularly located turns in the lamella structure.

The ^{13}C spin−lattice relaxation time observations of Ser Cβ carbons of (AGSGAG)$_5$ indicate that there are two components in the Ser side chain dynamics. This has already been noted in *B. mori* silk fibroins using ^{13}C T_1 studies of Ser Cβ carbons[28] and solid state ^2H NMR observation of the [3,3-$^2H_2$]Ser residues.[32,33] The slower component was shown to have more restricted motion. Curve fitting for the spin−lattice relaxation data indicated that the slow component accounted for approximately 40% of the Ser residues. The slower mobility of this component is thought to result from intra- and intermolecular hydrogen bond formation through Ser OH groups and/or from steric hindrance of the Ser side chains as a result of their location between two β-sheet planes.

Table 3. Eleven Structures with Higher Occurrence Probabilities over 0.03 Selected from 147 Possible Structures Used for the Second Statistical Mechanical Calculation[a]

probability	turn position
0.133	15
0.062	11, 23
0.044	7, 15
0.044	9, 19
0.044	9, 21
0.044	11, 21
0.044	10, 20
0.041	17
0.035	13
0.035	14
0.035	16

[a] The turn position of each structure is also listed. Details are described in the text.

Among the 147 possible structures used for statistical mechanical calculation, the structures with occurrence probabilities over 0.03 after the second calculation (see above) were listed in Table 3. The sum of the fractions of the five structures with one turn was 0.279 and that of the six structures with two turns was 0.282. Recently Gong et al. proposed a tentative structural model for cross-β protofibrils prepared from regenerated *B. mori* silk fibroin solution on the basis of X-ray diffraction, TEM, and AFM images.[34] The nanofibrils were approximately 5 nm in width while the GS or GA at the C-end of a typical segment GAGAGSGAGAGS gave rise to a single turn between a β-strand dimer. The laminated β-sheet length proposed by them is reasonably consistent with our models which consists of turns and β-strand with approximately 8−12 amino acids at the central sequence of the proposed lamella structure with two turns for the peptide, (AGSGAG)$_5$.

In conclusion, our findings suggest a novel model for the structure of the crystalline regions in *B. mori* fibroin in which the serine side chains, bulkier and more hydrophilic compared with those of (AG)$_n$ increase the separation between the β-sheets creating space for the formation of looser, more distorted, and less regularly located turns. These differences may have implications for chain mobility and the size regularity of the crystallites in silk with important implications for the mechanical behavior. To obtain further information on the structure and dynamics of the Ser residue, we are now seeking to prepare (AGSGAG)$_5$ with stable isotope-labeled Ser residues.

Acknowledgment. T.A. acknowledges support from a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Culture and Sports of Japan (18105007), Promotion of Basic Research Activities for Innovative Biosciences, Japan, and SENTAN, JST, Japan.

References and Notes

(1) Asakura, T.; Kaplan, D., *Encylopedia of Agricultural Science*; Academic Press: New York, 1994; p 1.
(2) Hakimi, O.; Knight, D. P.; Vollrath, F.; Vadgama, P. *Composites, Part B: Eng*. **2007**, *38*, 324–337.
(3) Asakura, T.; Ohgo, K.; Komatsu, K.; Kanenari, M.; Okuyama, K. *Macromolecules* **2005**, *38*, 7397–7403.
(4) Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komatsu, K. *J. Mol. Biol*. **2001**, *306*, 291–305.
(5) Marsh, R. E.; Corey, R. B.; Pauling, L. *Biochim. Biophys. Acta* **1955**, *16*, 1–34.
(6) Takahashi, Y.; Gehoh, M.; Yuzuriha, K. *Int. J. Biol. Macromol.* **1999**, *24*, 127–38.
(7) 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.

(8) Asakura, T.; Yao, J.; Yamane, T.; Umemura, K.; Ulrich, A. S. *J. Am. Chem. Soc.* **2002**, *124*, 8794–5.
(9) Yao, J. M.; Ohgo, K.; Sugino, R.; Kishore, R.; Asakura, T. *Biomacromolecules* **2004**, *5*, 1763–1769.
(10) Asakura, T.; Yao, J. M. *Protein Sci.* **2002**, *11*, 2706–2713.
(11) Asakura, T.; Nakazawa, Y.; Ohnishi, E.; Moro, F. *Protein Sci.* **2005**, *14*, 2654–2657.
(12) Asakura, T.; Sato, H.; Moro, F.; Nakazawa, Y.; Aoki, A. *J. Am. Chem. Soc.* **2007**, *129*, 5703–5709.
(13) Zhou, C.; Confalonieri, F.; Jacquet, M.; Perasso, R.; Li, Z.; Janin, J. *Proteins: Struct., Funct., Genetics* **2001**, *44*, 119–122.
(14) Pines, A.; Gibby, M. G.; Waugh, J. S. *J. Chem. Phys.* **1972**, *56*, 1776–1777.
(15) Gullion, T.; Schaefer, J. *J. Magn. Reson.* **1989**, *81*, 196–200.
(16) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
(17) Ohgo, K.; Ashida, J.; Kumashiro, K. K.; Asakura, T. *Macromolecules* **2005**, *38*, 6038–6047.
(18) Torchia, D. A. *J. Magn. Reson.* **1978**, *30*, 613–616.
(19) Ando, I.; Asakura, T. Chemical Shift Map In *Modern Magnetic Resonance*; Asakura, T., Saito, H., Ando, I., Eds.; Application in Chemistry 1; Springer: Berlin, 2006; Vol. *1*.
(20) Iwadate, M.; Asakura, T.; Williamson, M. P. *J. Biomol. NMR* **1999**, *13*, 199–211.
(21) Spera, S.; Bax, A. *J. Am. Chem. Soc.* **1991**, *113*, 5490–5492.
(22) Saito, H.; Tabeta, R.; Asakura, T.; Iwanaga, Y.; Shoji, A.; Ozaki, T.; Ando, I. *Macromolecules* **1984**, *17*, 1405–1412.
(23) Asakura, T.; Kuzuhara, A.; Tabeta, R.; Saito, H. *Macromolecules* **1985**, *18*, 1841–1845.
(24) Asakura, T.; Iwadate, M.; Demura, M.; Williamson, M. P. *Int. J. Biol. Macromol.* **1999**, *24*, 167–171.
(25) Asakura, T.; Demura, M.; Date, T.; Miyashita, N.; Ogawa, K.; Williamson, M. P. *Biopolymers* **1998**, *41*, 193–203.
(26) Suzuki, Y.; Asakura, T. *Polym. J.* **2010**, *42*, 354–356.
(27) Flory, P. J. *Statistical Mechanics of Chain Molecules*; John Wiley & Sons, Inc.: New York, 1969.
(28) Saito, H.; Ishida, M.; Yokoi, M.; Asakura, T. *Macromolecules* **1990**, *23*, 83–88.
(29) Farrar, T.; Becker, E. *Pulse and Fourier Transform NMR*; Academic Press: New York, 1971.
(30) Fraser, R. D. B.; MacRae, T. P.; Stewart, F. H. C. *J. Mol. Biol.* **1966**, *19*, 580–582.
(31) Fraser, R. D. B.; MacRae, T. P.; Stewart, F. H. C.; Suzuki, E. *J. Mol. Biol.* **1965**, *11*, 706–712.
(32) Saito, H.; Tabeta, R.; Kuzuhara, A; Asakura, T. *Bull. Chem. Soc. Jpn.* **1986**, *59*, 3383–3387.
(33) Kameda, T.; Ohkawa, Y.; Yoshizawa, K.; Naito, J.; Ulrich, A. S.; Asakura, T. *Macromolecules* **1999**, *32*, 7166–7171.
(34) Gong, Z. G.; Huang, L.; Yang, Y. H.; Chen, X.; Shao, Z. Z. *Chem. Commun.* **2009**, *48*, 7506–7508.