DOI: 10.1021/ma901949x

Macromolecules
ARTICLE

Structural Characterization of Silk-Based Water-Soluble Peptides $(Glu)_n(Ala\text{-}Gly\text{-}Ser\text{-}Gly\text{-}Ala\text{-}Gly)_4$ ($n = 4-8$) as a Mimic of *Bombyx mori* Silk Fibroin by ^{13}C Solid-State NMR

Aya Nagano,[‡,†] Yuka Kikuchi,[†] Hirohiko Sato,[†,§] Yasumoto Nakazawa,[∥] and Tetsuo Asakura*[,†]

[†]*Department of Biotechnology, Tokyo University of Agriculture, Technology, Koganei, Tokyo, 184-8588, Japan,* [‡]*Research Department, Japan Medical Materials Corporation, Osaka 532-0003, Japan,* [§]*Analysis Research Department, Chemical Laboratories, Nissan Chemical Industries Ltd., Funabashi, Chiba 274-8507, Japan, and* [∥]*Nature and Science Museum, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan*

Received September 2, 2009; Revised Manuscript Received October 18, 2009

ABSTRACT: Peptides with a combination of hydrophilic and hydrophobic sequences mimicking the primary structure of *Bombyx mori* silk fibroin were synthesized and studied in the solid state by NMR using ^{13}C selective labeling coupled with ^{13}C conformation-dependent chemical shifts and 2D solid-state spin-diffusion NMR. The hydrophilic sequence was poly(L-glutamic acid) $(E)_n$, and the hydrophobic one was the consensus sequence of the crystalline fraction of *B. mori* silk fibroin, $(AGSGAG)_4$. The balance of hydrophilic and hydrophobic characters of the peptide was controlled by changing the relative length, n, of $(E)_n$ from 4 to 8. When $n = 4$ and 5, the structure of the hydrophobic sequence is basically Silk I (the structure of *B. mori* silk fibroin before spinning in the solid state), and the polyglutamate sequences are random coil. On the other hand, when $n = 6-8$, the structure of the polyglutamate sequence changes gradually from random coil to β-sheet, and the hydrophobic sequence adopts a mixture of β-sheet and random coil/distorted β-turn forms, although the fraction of the latter form decreases gradually by increasing the number n from 6 to 8. Molecular dynamics and molecular mechanics calculations were also performed to examine the stability of the aggregated domains of the peptides in the solid state. The conformational change of $(E)_4(AGSGAG)_4$ was monitored in the solid state by decreasing the pH of the aqueous solution during the sample preparation.

Introduction

The fibrous silk fibroin produced by the domestic silkworm *Bombyx mori* has a number of outstanding properties not only for textile applications but also for design and development of various biotechnological and biomedical devices.[1] Zhou et al.[2] reported the primary structure of the heavy chain which is the main part of the *B. mori* silk fibroin chain and reported the presence of unusual repeat sequences in the silk fibroin. The primary structure can be divided into a repetitive region (R) and an amorphous region (A) together with N-terminal and C-terminal regions. The R region consists of highly repetitive AGSGAG sequences constituting the hydrophobic crystalline region, while the A region consists of a variety of amino acids such as TGSSGFGPYVANGGYSGYEYAWSSESDFGT. The latter region adds hydrophilic character to the silk fibroin. The origin of the desirable qualities of the silk fibroin basically comes from the combination of these unique amino acid sequences which interact synergistically to make it an excellent natural fiber.

Two crystalline forms, Silk I and Silk II, have been reported as the dimorphs of the silk fibroin, based on extensive investigations from X-ray fiber diffraction, electron diffraction, conformational energy calculations, infrared, and ^{13}C and ^{15}N solid-state NMR spectroscopy.[3,4] Silk I is the silk fibroin structure in the solid state before spinning and can be obtained as a silk film from liquid silk stored in the silk gland of the silkworm, whereas the Silk II form is the structure after spinning and is obtained as mature silk fibers from the cocoon. We have characterized the molecular structure of Silk I by using a synthetic peptide $(AG)_{15}$ as a simple model for the highly repetitive crystalline domain employing several solid-state NMR techniques and X-ray diffraction analysis.[5−7] A repeat β-turn type II structure, stabilized by a $4 \rightarrow 1$ intramolecular hydrogen bond, was proposed for the Silk I form. This unique structure has been suggested to be important to produce a silk fibroin fiber with high strength and high toughness from the aqueous solution of the silk fibroin. By using similar solid-state NMR techniques, the Silk II structure was also clarified.[8,9] The structure was heterogeneous in contrast to the Silk I structure and consists of roughly 70% antiparallel β-sheet and 30% distorted β-turn and/or random coil. In addition, the antiparallel structure consists of two different intermolecular arrangements although the torsion angles of the backbone chains are essentially the same in both structures. A change in the fraction of Silk I and Silk II including random coil structures produces a variety of silk fibroin samples with different physical properties which have been used for biomaterials in several forms, that is, long fiber, nonwoven fiber, film, sponge, powder, gel, and solution.[10]

Recent developments in biotechnology make it possible to produce new silk-like materials which are designed for biomaterials.[11−14] Since the *B. mori* silk fibroin chain consists of approximately alternating sequences of the hydrophobic crystalline regions $(AGSGAG)_m$ and hydrophilic amorphous region, the alternative introduction of poly(L-glutamic acid) $(E)_n$ as the hydrophilic region into the hydrophobic crystalline regions $(AGSGAG)_m$ seems interesting as a mimic of the silk fibroin.[15]

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

 Published on Web 10/30/2009

Moreover, many block copolymers with alternative hydrophilic and hydrophobic sequences have been developed as new biomaterials.[16−18] Since poly(L-glutamic acid) is involved in the hydroxyapatite-nucleating domains of bone sialoprotein, the resultant silk-like material may have potential use as bone graft substitute materials.[16,17] By changing the relative length n of $(E)_n$, it is possible to control the structures and physical properties of such silk-like materials.

In the present paper, we designed water-soluble silk fibroin model peptides, $(Glu)_n(Ala\text{-}Gly\text{-}Ser\text{-}Gly\text{-}Ala\text{-}Gly)_4$; $(E)_n(AGSGAG)_4$ where $n = 4-8$. The balance of hydrophilic and hydrophobic characters of the samples was controlled by changing the relative length, n, of $(E)_n$ from 4 to 8 with a fixed length of $(AGSGAG)_4$ in the peptides. The structure and structural change were studied by solid-state NMR. Especially, ^{13}C selective isotope labeling of several Ala and Gly carbons in the model peptides coupled with ^{13}C conformation-dependent chemical shifts and ^{13}C solid-state spin-diffusion NMR observations gives detailed information on the local structures.[4,5,7,18] This was applied to the detailed structural analysis of $(E)_n(AGSGAG)_4$ where $n = 4$ or 8. The alternation of hydrophilic and hydrophobic sequences is expected to produce phase-separated domains in the aggregated state, in which the Silk I or β-sheet domains will be stabilized by hydrogen-bonding formation and van der Waals interactions. In order to examine the stability of these aggregated domains, molecular dynamics (MD) simulation and molecular mechanics (MM) calculation[19,20] were performed based on the conformational information on these peptide molecules obtained from solid-state NMR. The conformational change from Silk I to Silk II for the peptide $(E)_4(AGSGAG)_4$ was monitored using solid-state NMR by decreasing the pH in the aqueous solution in the sample preparation process.

Experimental Part

Materials. The natural abundance model peptides, peptides **1**−**5** with the sequences $(E)_n(AGSGAG)_4$ ($n = 4-8$), respectively, were synthesized. The following ^{13}C selectively labeled peptides, peptides **6**−**11**, were also synthesized for the detailed structural analysis. **6**: $(E)_8AGSGAGAGS[2\text{-}^{13}C]GA\text{-}[1\text{-}^{13}C]G[1\text{-}^{13}C]AGSGAG[3\text{-}^{13}C]AGSGAG$; **7**: $(E)_8AGS\text{-}[1\text{-}^{13}C]G[1\text{-}^{13}C]AGAGSGAG[3\text{-}^{13}C]AGSGAGA[2\text{-}^{13}C]GSGAG$; **8**: $(E)_8AGS[2\text{-}^{13}C]G[3\text{-}^{13}C]AGAGSGAGAGSGA[1\text{-}^{13}C]G[1\text{-}^{13}C]\text{-}AGSGAG$; **9**: $(E)_4AGSGAGAGSGA[1\text{-}^{13}C]G[1\text{-}^{13}C]AGSGAG\text{-}[3\text{-}^{13}C]AGSGAG$; **10**: $(E)_4AGS[1\text{-}^{13}C]G[1\text{-}^{13}C]AGAGSGAG\text{-}[3\text{-}^{13}C]AGSGAGAGSGAG$; **11**: $(E)_4AGSG[3\text{-}^{13}C]AGAGSGA\text{-}GAGSGA[1\text{-}^{13}C]G[1\text{-}^{13}C]AGSGAG$.

The synthesis was performed using the Fmoc-solid phase method, and a fully automated Pioneer Peptide Synthesis System (Applied Biosystems Ltd.) was used throughout. The crude peptides were purified by RP-HPLC using acetonitrile/0.1%TFA as an eluant. After purification, acetonitrile was removed by evaporation, and then aqueous solutions of the peptides were lyophilized for 24 h and used for solid-state NMR observation. The peptides synthesized here are summarized in Table 1.

In the experiment on conformational change of the peptide **11**, the peptide was first dissolved in water at pH 8 and then the pH was changed to 6, 4, or 2. Within short time, the sample was adjusted to the desired pH, and then the aqueous solutions were lyophilized. Thus, three forms of the peptide with different conformations in the solid state were obtained.

^{13}C CP/MAS NMR Measurements. ^{13}C CP/MAS NMR experiments were performed on a Chemagnetics Infinity 400 spectrometer and a Bruker DSX-400 AVANCE spectrometer with an operating frequency of 100.0 MHz for ^{13}C at a sample spinning rate of 8 kHz in a 4 mm diameter ZrO_2 rotor. A total of 20K−30 K scans for the nonlabeled sample and 2K−3K scans for ^{13}C-labeled samples were collected over a spectral width of 35 kHz with a recycle delay of 3 or 4 s. All spectra were obtained using a cross-polarization time of 2 ms and broadband proton decoupling. Phase cycling was used to minimize artifacts. ^{13}C chemical shifts were represented relative to TMS (tetramethylsilane).[21]

Table 1. Natural Abundance Model Peptides, Peptides 1−5, and ^{13}C Selectively Labeled Peptides, Peptides 6−11, Synthesized in This Work

1: $(E)_4(AGSGAG)_4$
2: $(E)_5(AGSGAG)_4$
3: $(E)_6(AGSGAG)_4$
4: $(E)_7(AGSGAG)_4$
5: $(E)_8(AGSGAG)_4$
6: $(E)_8AGSGAGAGS[2\text{-}^{13}C]GA\ [1\text{-}^{13}C]G[1\text{-}^{13}C]\ AGSG\ AG[3\text{-}^{13}C]\text{-}AGSGAG$
7: $(E)_8AGS\ [1\text{-}^{13}C]G\ [1\text{-}^{13}C]\ AGAGSGAG[3\text{-}^{13}C]AGSGAGA\text{-}[2\text{-}^{13}C]GSGAG$
8: $(E)_8AGS[2\text{-}^{13}C]G\ [3\text{-}^{13}C]AGAGSGAGAGSGA[1\text{-}^{13}C]\ G[1\text{-}^{13}C]\ AGSGAG$
9: $(E)_4AGSGAGAGSGA\ [1\text{-}^{13}C]G\ [1\text{-}^{13}C]AGSGAG\ [3\text{-}^{13}C]AGSGAG$
10: $(E)_4AGS\ [1\text{-}^{13}C]G\ [1\text{-}^{13}C]AGAGSGAG\ [3\text{-}^{13}C]AGSGAGA\text{-}GSGAG$
11: $(E)_4AGSG\ [3\text{-}^{13}C]AGAGSGAGAGSGA\ [1\text{-}^{13}C]G\ [1\text{-}^{13}C]\text{-}AGSGAG$

2D Spin Diffusion ^{13}C Solid-State NMR Measurements and Simulation. The 2D spin-diffusion NMR spectrum was obtained using a Varian Unity INOVA 400 NMR spectrometer with a 7 mm Jakobsen-type double-tuned MAS probe at off magic angle condition ($\theta_m - 6.2°$) at room temperature.[5] The sample spinning rate was 5 kHz. The scaling factor of the 2D spin-diffusion spectra is $^1/_2(3\cos^2(\theta_m - 6.2°) - 1) = 0.158$. The mixing times were set to 2 s. The contact time was set to 2 ms using the variable-amplitude CP technique. The principal values of the chemical shift tensors for the carbonyl carbon nuclei of the ^{13}C-labeled Ala and Gly residues were determined by analysis of the spinning sidebands under slow MAS conditions using a Varian Unity INOVA 400 NMR spectrometer.[22] For simulation of 2D spin-diffusion NMR spectra, the calculations were performed using a grid of 15° for ϕ and φ values. A calculation program developed in our laboratory was used to simulate the 2D spin-diffusion NMR spectra.[5] The rmsd value was used to quantify the difference between the experimental and calculated spectra.[23]

MD Simulation. MD and MM calculations were performed for aggregated $(E)_4(AGSGAG)_4$ or $(E)_8(AGSGAG)_4$ molecules by using the "Discover" module in Materials Studio 4.1 (Accelrys Inc.).[19,20] All of the simulations were performed by using a pcff force field in vacuo. MD calculations were performed under a periodic boundary condition including 16 molecules with NPT ensemble, that is, number of the peptide molecules is constant, and pressure is controlled at 0.2 GPa along the chain and temperature is controlled at 398 K. The calculations used 50 000 steps over 50 ps. The initial structures of these peptides for MD and MM calculations are shown in Figure 1. If the sequence in the peptide takes a specified conformation, Silk I or β-sheet, based on solid-state NMR results as described later, this conformation was adopted as the initial conformation. The coordinates reported previously for Silk I[7] or β-sheet[3] were used. If the conformation of the sequence is in random coil, an extended structure (ϕ and $\varphi = -180°$, 180°) was used as the initial one. Thus, the initial conformations of $(E)_8(AGSGAG)_4$ and $(E)_4(AGSGAG)_4$ are shown in parts a and d of Figure 1, respectively. Then, the aggregated initial state models were built up. For $(E)_8(AGSGAG)_4$, both parallel and antiparallel β-sheet arrangements of the sequence $(E)_8$ both within sheet and in the stacked state were set as shown in parts b and c of Figure 1, respectively. On the other hand, for $(E)_4(AGSGAG)_4$, the coordinates of Silk I reported previously[16] were used for the structure of the sequence $(AGSGAG)_4$, as shown in Figure 1e. Then the MD simulation was performed. The change of these

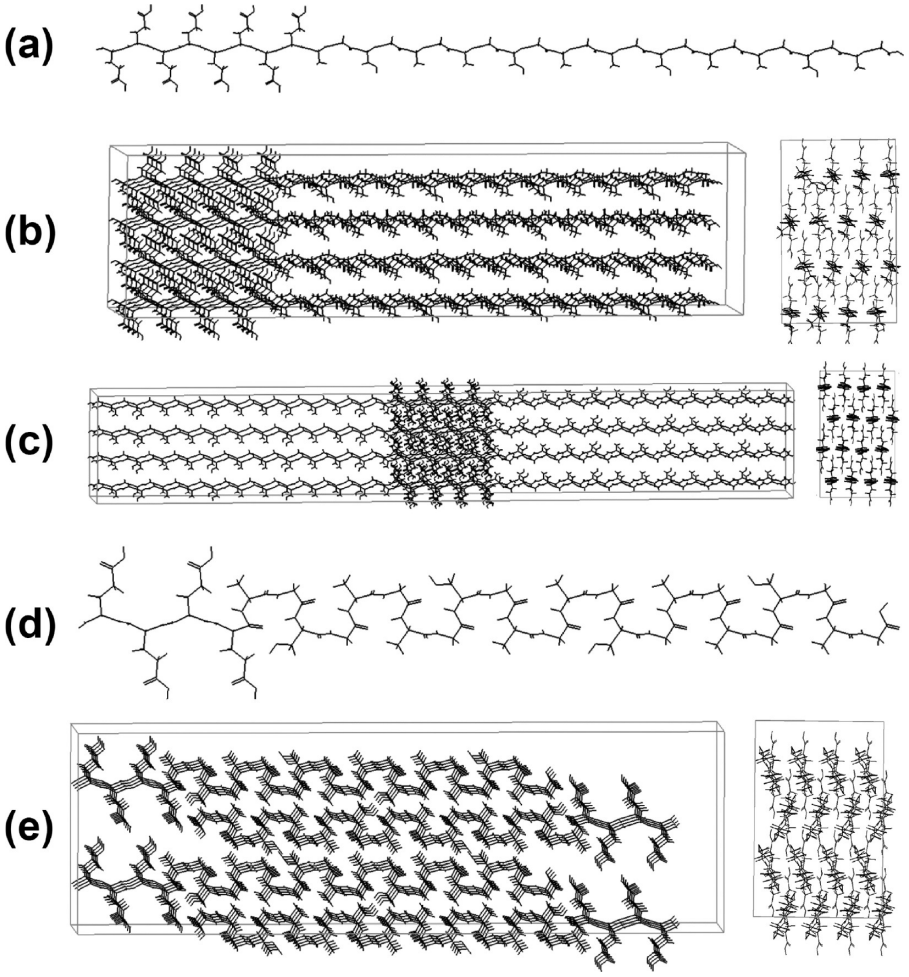

Figure 1. Initial models of $(E)_8(AGSGAG)_4$ and $(E)_4(AGSGAG)_4$ for MD and MM calculations. (a) Initial conformation of an $(E)_8(AGSGAG)_4$ molecule with extended structure (ϕ and ψ = −180°, 180°). The side chains are also extended structure. (b, c) Aggregated initial state models of $(E)_8(AGSGAG)_4$ with parallel and antiparallel chain arrangements within one sheet, respectively, and the sheet was stacked four times. (d) Initial conformation of an $(E)_4(AGSGAG)_4$ molecule with extended structure for the sequence $(E)_4$ and Silk I structure for the sequence, $(AGSGAG)_4$. (e) Aggregated initial state model of $(E)_4(AGSGAG)_4$ where the aggregated Silk I structure for $(AGSGAG)_4$ was assumed and the sequence $(E)_4$ was attached at the N-terminal with extended structure. In (b), (c), and (e), the left structures are side view and the right ones are end view.

structures in the aggregated state including changes in the conformations was monitored . Then MM calculations were performed to obtain the energy-minimized structure in the aggregated state.

Results and Discussion

^{13}C CP/MAS NMR Spectra of Nonlabeled Peptides, $(E)_n(AGSGAG)_4$ (n = 4−8). Figure 2 shows a comparison of the ^{13}C CP/MAS NMR spectra of $(E)_n(AGSGAG)_4$ (n = 4−8), in order to examine the effect of the length of polyglutamic acid $(E)_n$ at the N-terminal site on the structure of the peptides in the solid state. The chemical shift data of the peptides **1** and **5** together with the assignments are summarized in Table 2. The sharp peaks observed at 16.7, 51.1, and 176.8 ppm assigned to Ala Cβ, Cα, and CO carbons, respectively, indicate that the sequence $(AGSGAG)_4$ takes exclusively Silk I form,[16] which is the case of both peptides **1** and **2** (Figure 1, a and b, respectively). The chemical shifts of Gly Cα (43.5 ppm), C=O (170.4 ppm), and Ser Cα (55−60 ppm), Cβ (60.3 ppm), and C=O (173.7 ppm) indicate that these residues are also in Silk I form.[16] The peptide $(AGSGAG)_4$ without $(E)_n$ was insoluble in water and took Silk II form even if the solvents and/or methods for precipitation were changed.[8,9,24] Thus, the presence of $(E)_n$ (n = 4 or 5) at the N-terminus makes the sequence $(AGSGAG)_4$ soluble and induces a Silk I form of the sequence.

The Ala Cβ peak changes dramatically between n = 5 and 6 as shown in Figure 2b,c. The spectral characterization of such an Ala Cβ peak of $(AGSGAG)_m$ when n = 6 has been reported by us previously.[21,26−28] Namely, the broad component at 16.4 ppm was assigned to a distorted β-turn structure, which is characterized by a large distribution in the torsion angles around an average conformation of a type II β-turn. The chemical shift of this peak is also in agreement with that of the random coil peak of Ala in aqueous solution as discussed previously.[21,26,27] Thus, this component is assigned to random coil/distorted β-turn. The other component with a chemical shift around 19.2 ppm is assigned to β-sheet structures. The details of the assignments have also been reported previously.[21,26−28] Thus, by changing the length of poly(L-glutamic acid) from n = 5 to 6, the Silk I structure of the sequence $(AGSGAG)_4$ changes to a mixture of random coil/distorted β-turn and β-sheet structure. The fraction of β-sheet structure increases slightly by increasing the number n from 6 (Figure 2c) to 8 (Figure 2e). The peak of the Gly Cα carbon shifts to higher field slightly with peak broadening. The Ser Cβ peak also shifts to lower field with remarkable peak broadening. These results indicate that the conformations of these residues in the sequence $(AGSGAG)_4$ also change to a mixture of random coil/distorted β-turn and β-sheet structure. The carbonyl carbons also show the conformational change.

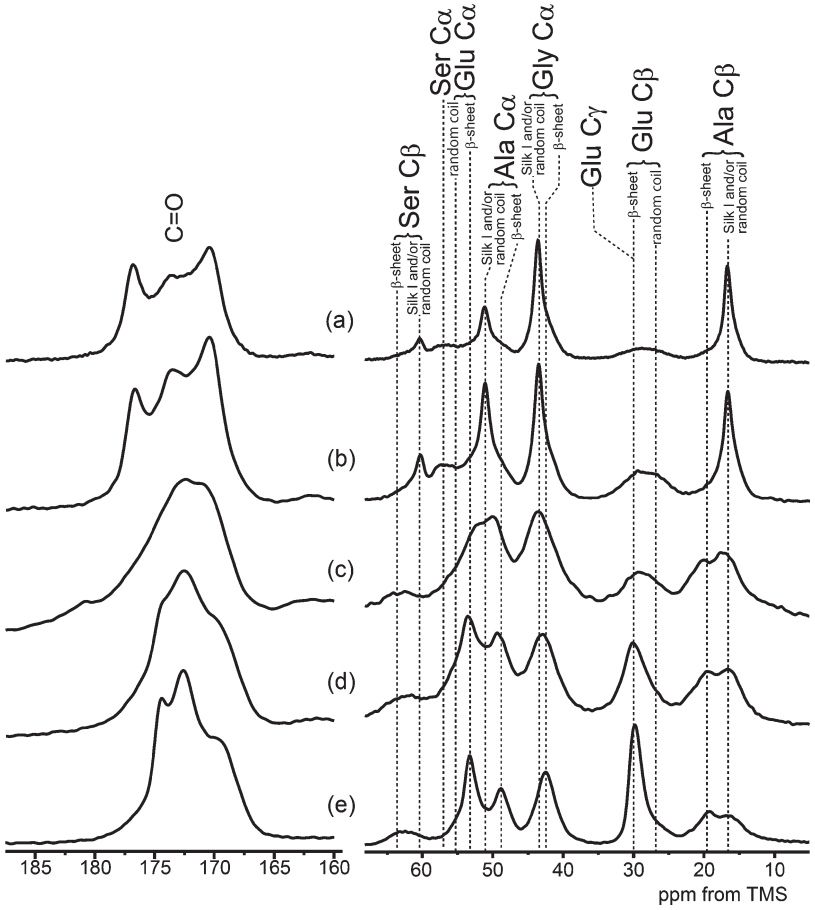

Figure 2. ^{13}C CP/MAS NMR spectra of natural abundance $(E)_n(AGSGAG)_4$ ($n = 4-8$). The spectra from (a) to (e) correspond to the samples from $n = 4$ to $n = 8$, respectively.

Table 2. Summary of Chemical Shifts (in ppm) of Nonlabeled Peptides $E_n(AGSGAG)_4$ Where $n = 4, 8$

	peptide **1**	peptide **5**		Silk I	Silk II
residue	$E_4(AGSGAG)_4$	$E_8(AGSGAG)_4$	random coil	film	fiber
GlyCα	43.5	42.6	42.7	43.8	42.5
Gly C=O	170.4	169.5	171.3	170.7	169.3
Ala Cα	51.1	48.9	50.1	51.4	48.9
Ala Cβ	16.7	16.4, 19.2	16.6	16.5	16.6, 19.6, 21.9
Ala C=O	176.8	174.4	175.2	177.0	172.2
Ser Cα	55−60 (br)[a]	—[b]	55.9	58.0	54.6
Ser Cβ	60.3	60−65 (br)[a]	61.3	60.7	63.9
Ser C=O	173.7	172.6	172.4	173.7	172.3
GluCα	—[c]	53.3	55.2		53.2
GluCβ	27 (br)[a]	29.9	26.9		30.0
GluCγ	30 (br)[a]	30	30.5		

[a] br, broad peak. [b] Overlapped with Glu Ca peak. [c] Overlapped with Ser Ca peak.

In order to monitor the conformational transition from random coil to β-sheet of the sequence $(E)_n$, the Glu Cβ and Glu Cγ peaks can be used because there are no other peaks in this area. Both peaks of Glu Cβ (27 ppm) and Cγ (30 ppm) were broad in $(E)_4(AGSGAG)_4$, and therefore we observed one broad overlapped peak at the center of these two peaks. In this case, the conformation of $(E)_4$ is random coil. With increasing the number n of $(E)_n$, the Glu Cγ peak (30 ppm) does not shift because of the conformation-independent character, but the Glu Cβ peak shifts from 27 ppm (random coil in $((E)_{4-5}(AGSGAG)_4)$) to 30 ppm (β-sheet in $((E)_{7-8}(AGSGAG)_4)$). This is in agreement with the reported chemical shift data of random coil and β-sheet forms of poly(L-glutamic acid) (Table 2). Therefore, the Glu Cα peak should also shift by this conformational change. The Glu Cα random coil peak of $(E)_4(AGSGAG)_4$ should appear at 55 ppm, but it is difficult to assign clearly because of the presence of Ser Cα peak and also broadness of the Glu Cα peak. The latter broadness corresponds to the broad peaks of Glu Cβ and Glu Cγ carbons of $(E)_4(AGSGAG)_4$. The sharp Glu Cα peak at 53.3 ppm in $(E)_8(AGSGAG)_4$ can be assigned to β-sheet peak. The sharpness of the peak also corresponds to the sharp peaks of Glu Cβ and Glu Cγ carbons. The Glu Cα peak of $(E)_6(AGSGAG)_4$ is also broad, which corresponds to the broad peaks of Glu Cβ and Glu Cγ, and therefore it is difficult to point out the peak position clearly. Thus, the chemical shift behavior including shape of the peak can monitor the conformational change of the region $(E)_n$ with increasing the number of n.

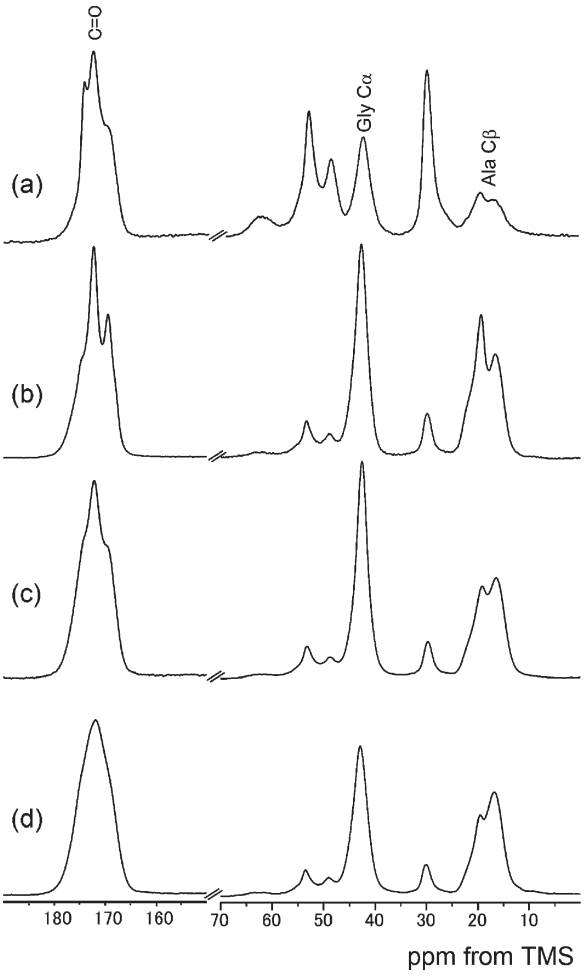

Figure 3. ^{13}C CP/MAS NMR spectra of the peptide **5** (a) and a series of ^{13}C-labeled peptides, $E_8(AGSGAG)_4$ with different ^{13}C labeling positions, namely the peptides, **6** (b), **7** (c), and **8** (d).

^{13}C CP/MAS NMR Spectra of ^{13}C Selectively Labeled Peptides $E_8(AGSGAG)_4$. The ^{13}C CP/MAS NMR spectra of the peptide **5** and a series of ^{13}C-labeled peptides, $E_8(AGSGAG)_4$ with different ^{13}C labeling position, the peptides **6**−**8**, are shown in Figure 3. The local conformations of Ala^{27}, Ala^{21}, and Ala^{13} residues in $(AGSGAG)_4$ are discussed from the ^{13}C-labeled Ala Cβ peaks (Figure 3b−d), respectively. Although the conformation of these residues is a mixture of β-sheet and random coil/distorted β-turn, the fraction of β-sheet is the largest for Ala^{27} and decreases slightly toward the N-terminal sequence $(E)_n$. Figure 4 shows the deconvolution of these Ala Cβ peaks to determine the fraction of random coil/distorted β-turn and β-sheet structure. Here, the former fraction was determined from the relative peak intensity at 16.4 ppm (shaded in gray). As discussed in previous papers,[8,9] the lower field peak at 19.2 ppm can be deconvoluted into two peaks, and the deconvolution was performed by assuming two peaks as shown in Figure 4. However, the sum of fractions of these two peaks is considered to be the fraction of β-sheet. This fraction decreases 55% (a), 43% (b), and 25% (c) in Ala^{27}, Ala^{21}, and Ala^{13}, respectively. The ^{13}C labeling was also performed for the Cα carbons of Gly^{18}, Gly^{28}, and Gly^{12} residues in peptides **6**−**8**. However, there are no significant changes in both the chemical shift and peak broadening. This is due to the small conformation-dependent chemical shift of the Gly Cα carbon.[29]

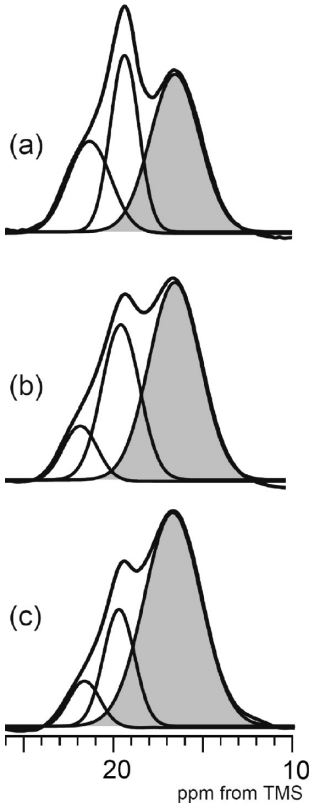

Figure 4. Deconvolution of Ala Cβ peaks of peptides **6** (a), **7** (b), and **8** (c) to determine the fraction of random coil/distorted β-turn and β-sheet structure.

2D Spin-Diffusion NMR Analyses of ^{13}C Selectively Double-Labeled Peptides $E_8(AGSGAG)_4$. 2D spin diffusion ^{13}C solid-state NMR was used to investigate the structures in more detail through the determination of the torsion angles of the labeled residues.[4,5,7,18] Figure 5 summarizes the theoretical spin-diffusion patterns of Ala residues as a function of the torsion angles, ϕ and ψ. The pattern changes significantly depending on the torsion angles, and therefore it is possible to obtain information on the torsion angle from the observed spin-diffusion NMR spectrum. For this purpose, we synthesized the peptides **6**−**8**, where the carbonyl carbons of the Gly^{20}Ala^{21}, Gly^{12}Ala^{13}, and Gly^{26}Ala^{27} residues were ^{13}C double-labeled. The observed 2D spin-diffusion NMR spectra are shown in Figure 6a−c for peptides **6**−**8**, respectively. However, because of a mixture of random coil/distorted β-turn and β-sheet structure, it is difficult to determine the torsion angles of these Ala residues exactly by matching the spin-diffusion spectra. Therefore, we tried to reproduce the spectral patterns by using the fractions of random coil/distorted β-turn and β-sheet structure that were determined from Figure 4 discussed above. For example, for the Ala^{27} residue using the spin-diffusion spectrum of Gly^{26}Ala^{27} in peptide **8**, the fractions determined from the Ala^{27} Cβ peak in peptide **6** were used. Similarly, for the Ala^{21} residue using Gly^{20}Ala^{21} in peptide **6**, and Ala^{13} residue using Gly^{12}Ala^{13} in peptide **7**, the fractions determined from the Ala^{21}Cβ peak in peptide **7** and Ala^{13} Cβ peak in peptide **8**, respectively, were used. In the selection of the spectral patterns from Figure 5, β-sheet structure is assumed to take the torsion angle $(\phi, \psi) = (-150°, 150°)$, and the random coil/distorted β-turn is the torsion angle $(\phi, \psi) = (60°, 125°)$.[16] The agreement between the observed and simulated spectra is good as shown in Figure 6d−f, confirming that the fractions of random

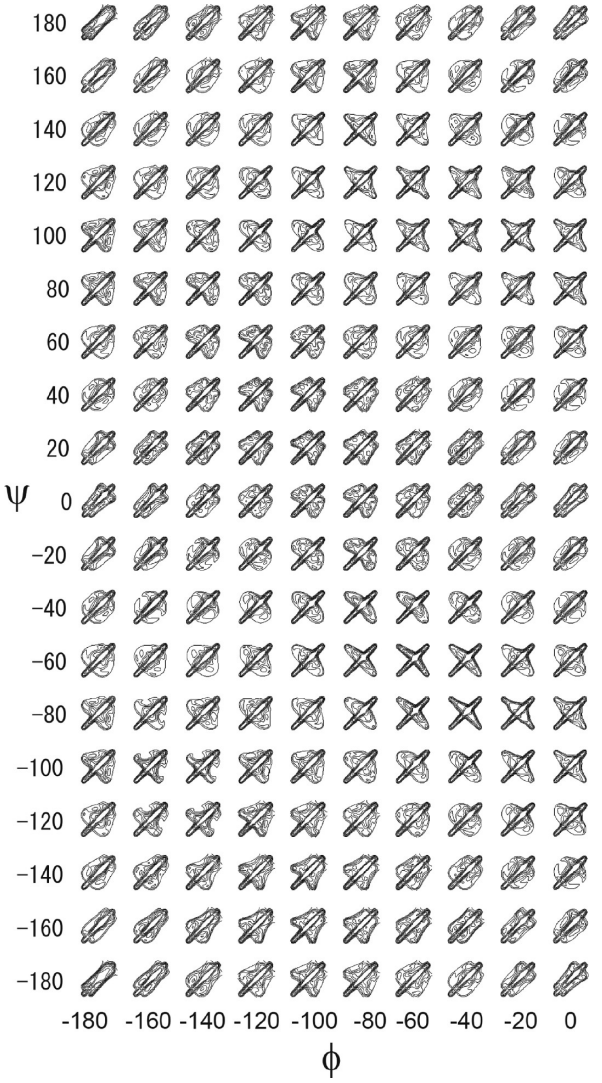

Figure 5. Theoretical ^{13}C spin-diffusion NMR spectral patterns of Ala residues as a function of the dihedral angles, ϕ and ψ.

coil/distorted β-turn and β-sheet structure determined are approximately correct.

^{13}C CP/MAS NMR Spectra of ^{13}C Selectively Labeled Peptides $E_4(AGSGAG)_4$. The ^{13}C CP/MAS NMR spectra of peptides with $n = 4$, namely the peptide **1** and a series of ^{13}C-labeled peptides, $E_4(AGSGAG)_4$ with different ^{13}C labeling position, peptides **9**−**11**, are shown in Figure 7. The local conformations of Ala^{23}, Ala^{17}, and Ala^{9} residues in $(AGSGAG)_4$ can be discussed from the ^{13}C-labeled Ala Cβ peaks in parts b, c, and d, respectively. Contrary to the spectra of peptides **6**−**8**, the spectra do not change among peptides **9**−**11**. This means that the Silk I structure of $(AGSGAG)_4$ was basically maintained independently of the position. However, the peak is slightly broader for Ala^{9} Cβ carbon as shown in Figure 7d, which suggests that this has the highest amount of random coil form among the three labeled Ala residues.

2D Spin-Diffusion NMR Analyses of ^{13}C Selectively Double-Labeled Peptides $E_4(AGSGAG)_4$. We synthesized peptides **9**−**11**, where the carbonyl carbons of the $Gly^{16}Ala^{17}$, $Gly^{8}Ala^{9}$, and $Gly^{22}Ala^{23}$ residues were ^{13}C double-labeled. The observed spin-diffusion NMR spectra are shown in parts a-1, a-2, and a-3 of Figure 8 for peptides **9**−**11**, respectively, which are significantly different from the spectra in Figure 6. The conformation is basically Silk I judging

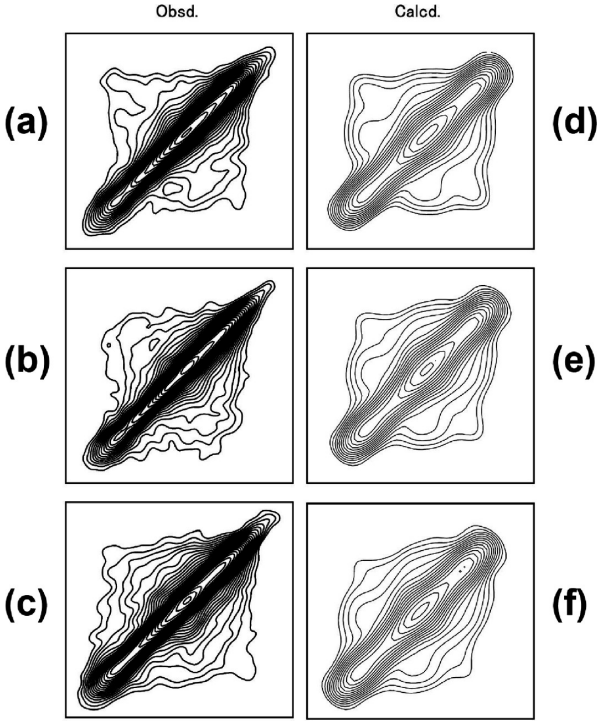

Figure 6. 2D spin-diffusion ^{13}C solid-state NMR spectra of the expanded carbonyl region of (a) [1-^{13}C]G^{20}[1-^{13}C]A^{21} in peptide **6**, (b) [1-^{13}C]G^{12}[1-^{13}C]A^{13} in peptide **7**, and (c) [1-^{13}C]G^{26}[1-^{13}C]A^{27} in peptide **8** observed under off-MAS conditions. The corresponding simulated spectra are shown as (d), (e), and (f), respectively. In the simulation, the fraction of random coil/distorted β-turn and β-sheet structure determined from Figure 4 was used.

from Figure 7 as mentioned above. Therefore, in this case it is possible to estimate the exact torsion angles of the Silk I form of the three Ala residues in the sequence $(AGSGAG)_4$ in peptides **9**−**11**. For this purpose, an rmsd map was prepared using the difference between the observed and calculated 2D spin diffusion NMR spectra in the range of $\phi = -180°-0°$ and $\psi = -180°-180°$. The simulated spectra b-1, b-2, and b-3 are in excellent agreement with the observed ones. All the rmsd maps c-1, c-2, and c-3 indicate that the torsion angles of Silk I are $(\phi, \psi) = (-60°, 135°)$ for all three Ala residues (white circle). These torsion angles obtained here are in agreement with the angles $(\phi, \psi) = (-62°, 125°)$ reported for Silk I form previously within experimental error.[7] The relatively broad error distribution in the rmsd map c-2 was obtained for the 2D spin-diffusion NMR spectrum of the carbonyl carbons of $Gly^{8}Ala^{9}$ residues, which corresponds to the slightly broader peak of Ala^{9} Cβ carbon in Figure 7d. This seems reasonable that Ala^{9} residue is the closest to the N-terminal $(E)_4$ and has the largest proportion of random coil form among the three Ala residues.

MD Simulation. Using solid-state NMR, the conformations of the peptides $(E)_4(AGSGAG)_4$ and $(E)_8(AGSGAG)_4$ were clarified in detail. However, the information on the aggregated states of these peptides is limited. Therefore, MD simulation and then MM calculation were performed complementarily to examine the stability of these peptide molecules in the aggregated state based on the conformations obtained from the solid-state NMR. For $(E)_8(AGSGAG)_4$, the conformation of the sequence $(E)_8$ is concluded to take β-sheet structure clearly, but it is difficult to judge whether the sheet is parallel or antiparallel. Thus, the initial structures before MD calculation were placed in parallel or antiparallel β-sheet forms for the aggregated state of the sequence $(E)_8$, as

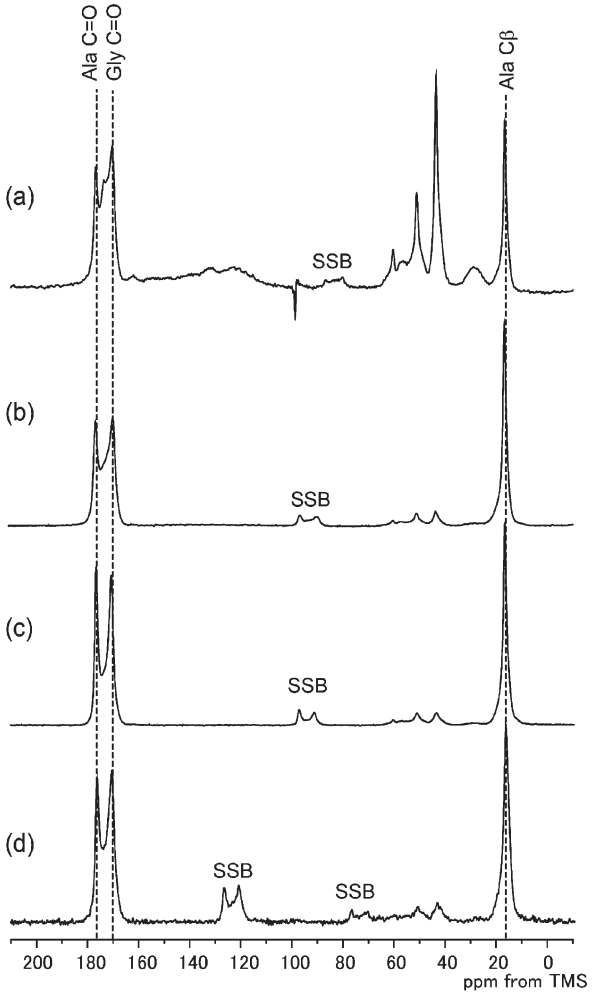

Figure 7. ^{13}C CP/MAS NMR spectra of the peptide **1** (a) and a series of ^{13}C labeled peptides, $E_4(AGSGAG)_4$ with different ^{13}C labeling positions, peptides, **9** (b), **10** (c), and **11** (d).

shown in parts b and c of Figure 1, respectively. On the other hand, a mixture of β-sheet and random coil/distorted β-turn was the local conformation of the sequence $(AGSGAG)_4$, and therefore an extended conformation for the sequence was assumed as the initial structure. If an initial model of the aggregated state changes to unstable state during the MD simulation, the final structure will not be able to be obtained. However, we could obtain the aggregated structures after the MD simulation as shown in Figure 9a,b. These are the final models of the aggregated $(E)_8(AGSGAG)_4$ molecules with parallel or antiparallel arrangements after MD and then MM calculations. Regular arrangements of the sequences of poly(L-glutamic acid) with intermolecular hydrogen bonding are formed and stabilized in both arrangements. On the other hand, although an extended structure was used for the sequence $(AGSGAG)_4$ in the initial state, the structure was destroyed and adopted random coil form after the calculation.

A similar calculation was performed for $(E)_4(AGSGAG)_4$ molecules. Figure 9c shows the final results of the aggregated $(E)_4(AGSGAG)_4$ molecules. The initial structure was Silk I for the sequence $(AGSGAG)_4$, and $(E)_4$ takes an extended form. After the MD calculation, the regular arrangement of the sequence $(AGSGAG)_4$ in Silk I form remains, with intra- and intermolecular hydrogen-bonding formation. On the other hand, the E_4 sequence had an extended form in the initial state, but the structure was destroyed and adopted a random coil form after the calculation. This is also the same as observed by solid-state NMR.

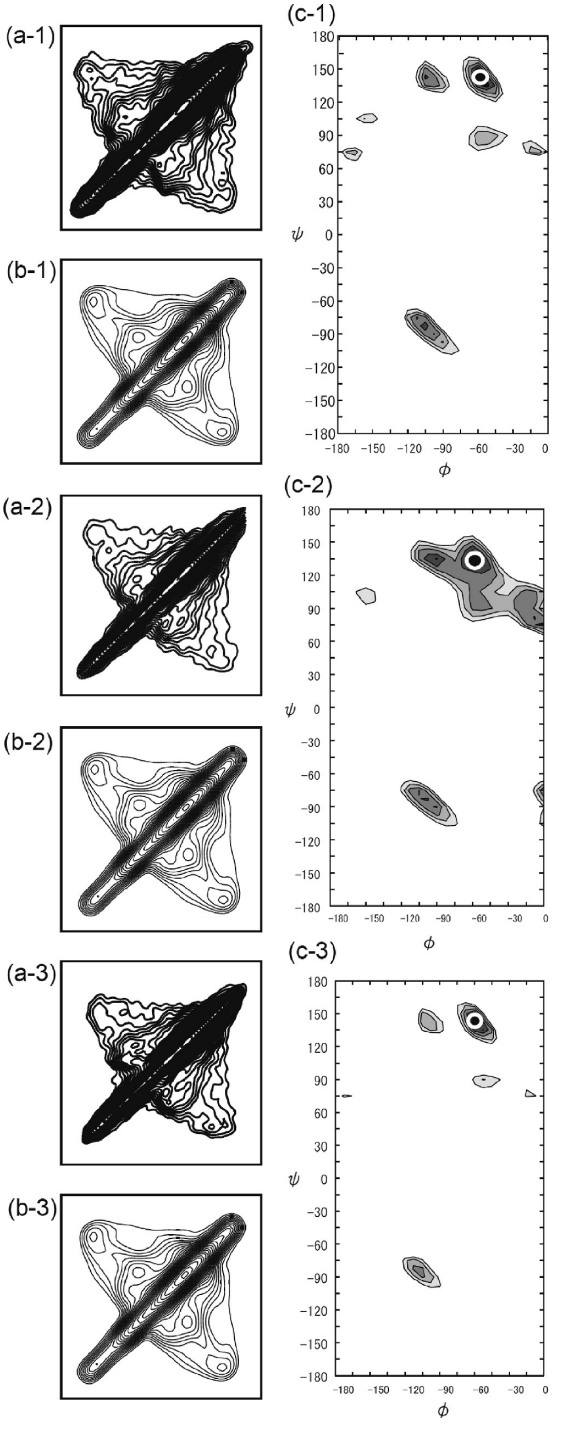

Figure 8. Three sets of spin-diffusion NMR analyses. 2D spin-diffusion ^{13}C solid-state NMR spectra of the expanded carbonyl region of (a-1) [1-^{13}C]G^{16}[1-^{13}C]A^{17} in peptide **9**, (a-2) [1-^{13}C]G^{8}[1-^{13}C]A^{9} in peptide **10**, and (a-3) [1-^{13}C]G^{22}[1-^{13}C]A^{23} in peptide **11** observed under off-MAS conditions. The corresponding simulated spectra are shown as (b-1), (b-2), and (b-3), respectively. The rmsd maps for the determination of the torsion angles, (ϕ, ψ) of Ala residues from the 2D spin-diffusion NMR spectra are also shown as (c-1) for Ala^{17}, (c-2) for Ala^{9}, and (c-3) for Ala^{23}, respectively. The smallest rmsd regions, and therefore the most likely (ϕ, ψ) combination, are indicated as the area enclosed by the white circle.

Structural Change of the Peptide $(E)_4(AGSGAG)_4$ by Changing the pH Value in the Aqueous Solution. The conformational transition from Silk I to Silk II of *B. mori* silk

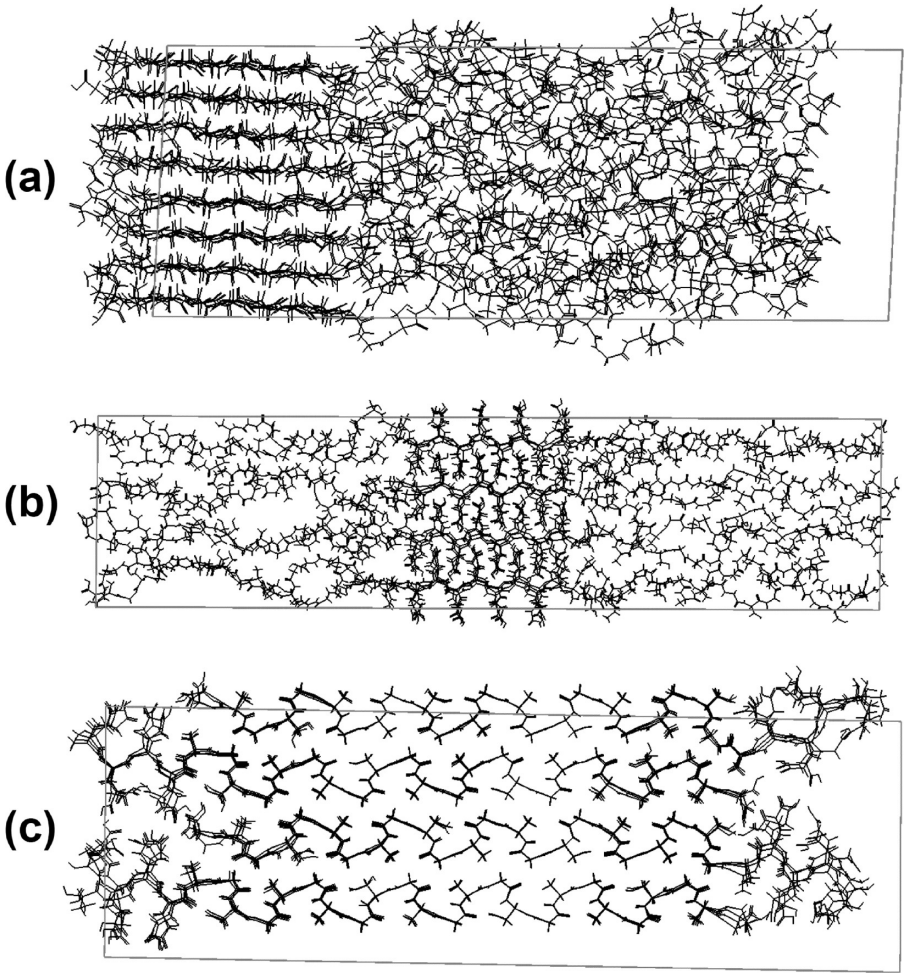

Figure 9. Final results of (a) $(E)_8(AGSGAG)_4$ parallel, (b) $(E)_8(AGSGAG)_4$ antiparallel, and (c) $E_4(AGSGAG)_4$ molecules in the aggregated states after MD and then MM calculations. The initial structures are shown in (b), (c), and (e) of Figure 1, respectively, in detail.

fibroin is important in relation to the fiber formation mechanism to prepare silk fibroin fiber with high strength and toughness.[19,20] We have reported MD and MM calculations about the conformational change of the model peptides for the crystalline domain as well as many experimental works with model peptides.[5,6,9,16,29] However, peptides like $(AGSGAG)_m$ are generally insoluble in water because of their hydrophobic character. Therefore, it seems interesting to monitor the structural change from Silk I to Silk II for the sample $(E)_4(AGSGAG)_4$, which is soluble in water and takes Silk I structure in the solid state. For this purpose, we used peptide **11**: $(E)_4$AGSG[3-^{13}C]AGAGSGAGGSGA[1-^{13}C]-G-[1-^{13}C]AGSGAG. The $^{13}C\beta$ peak from Ala^9 gives information on the local conformation close to the N-terminal $(E)_4$ sequence, while the carbonyl peaks from both Gly^{22} and Ala^{23} give information on the local conformation away from the N-terminal sequence. The peptide was dissolved in water at pH = 8, and then the pH was changed to 6, 4, or 2 in order to monitor the structural change to Silk II structure. The ^{13}C CP/MAS NMR spectra are summarized in Figure 10. At pH = 6, the structure seems to be mainly random coil/distorted β-turn judging from the chemical shift and broadening of the Ala^9 Cβ carbon (Figure 10b) although the spectrum (Figure 10a) shows Silk I structure when the peptide was lyophilized after acetonitrile was removed by evaporation. Namely, a structural transition from Silk I to random coil/distorted β-turn occurs with decreasing pH slightly. With further decreasing of pH value, a conformational change from random coil/distorted β-turn to β-sheet occurs. From the peak simulations of the spectra b−d, the fraction of β-sheet was determined as 0% (pH = 6), 27% (pH = 4), and 44% (pH = 2). Similar tendencies were observed for the carbonyl carbons of both Gly^{22} and Ala^{23} residues although the fraction of β-sheet seems slightly increased because of an increase in the intensities of the sharp peaks assigned to β-sheet structure.

Thus, we could monitor the structural change from Silk I to Silk II in the solid state for the model peptide of the crystalline domain of *B. mori* silk fibroin, $(AGSGAG)_n$, dissolved in water by the presence of a hydrophilic sequence such as poly(L-glutamic acid) in the chain. A change in the pH in the aqueous solution is one of the factors to prepare the silk fibroin with β-sheet structure when poly(L-glutamic acid) is present in the chain.

Conclusion

For design of various biotechnological and biomedical devices based on the excellent properties of *B. mori* silk fibroin, the introduction of hydrophilic sequences to the hydrophobic sequences of the fibroin is important to develop such a material by mimicking the primary structure of the silk fibroin. We synthesized water-soluble silk fibroin model peptides, $(E)_n(AGSGAG)_4$ where n = 4−8. The balance of hydrophilic and hydrophobic characters of the samples was controlled by changing the relative length, n, of $(E)_n$ from 4 to 8 with a fixed length of $(AGSGAG)_4$ in the peptides. The alternation of hydrophilic and hydrophobic

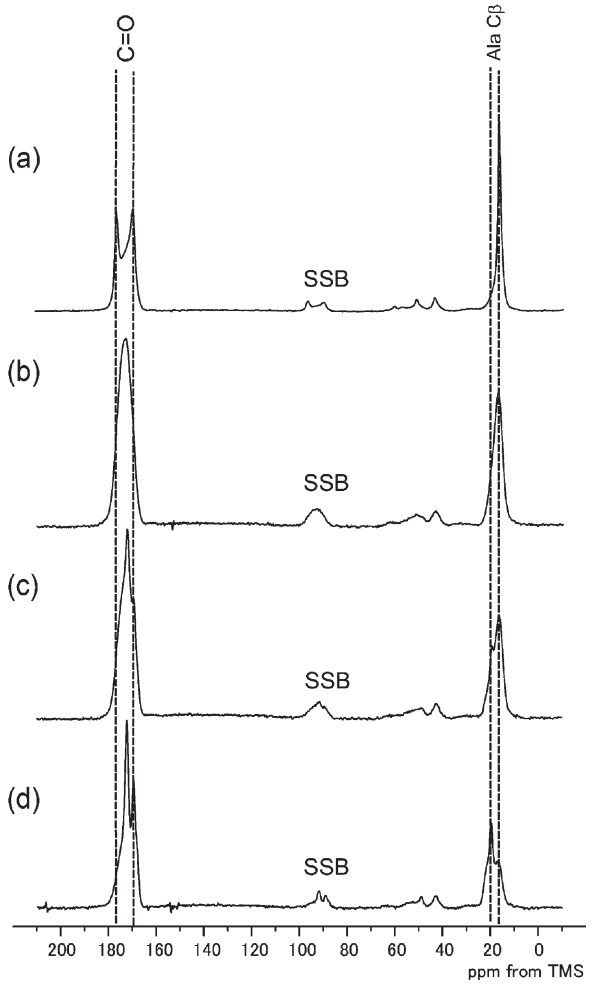

Figure 10. ^{13}C CP/MAS NMR spectra of the peptide **11**, $(E)_4$-AGSG[3-^{13}C]AGAGSGAGAGSGA-[1-^{13}C]G[1-^{13}C]AGSGAG: (a) lyophilized after acetonitrile was removed by evaporation; (b) lyophilized after the aqueous solution was adjusted to pH 6; (c) lyophilized after the aqueous solution was adjusted to pH 4; and (d) lyophilized after the aqueous solution was adjusted to pH 2.

sequences produced phase-separated domains in the aggregated state, in which the Silk I or β-sheet domains was stabilized by hydrogen-bonding formation and van der Waals interactions.

Acknowledgment. T.A. acknowledges support from the Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Culture and Sports of Japan (18105007). This study was supported in part by grants from Grant-in-Aid for Scientific Research (20750174) from JSPS. We also acknowledge Prof. Mike Williamson (Univ. Sheffield, UK) for many useful discussion and careful reading of our draft.

References and Notes

(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) Zhou, C.; Confalonieri, F.; Medina, N.; Zivanovic, Y.; Esnault, C.; Yang, T.; Jacquet, M.; Janin, J.; Duguet, M.; Perasso, R.; Li, Z. *Nucleic Acids Res.* **2000**, *28*, 2413–2419.
(3) Marsh, R. E.; Corey, R. B.; Pauling, L. *Biochem. Biophys. Acta* **1955**, *16*, 1–34.
(4) Zhao, C.; Asakura, T. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 301–352.
(5) Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komatsu, K. *J. Mol. Biol.* **2001**, *306*, 291–305.
(6) Asakura, T.; Yamane, T.; Nakazawa, Y.; Kameda, T.; Ando, K. *Biopolymers* **2001**, *58*, 521–525.
(7) Ashida, J.; Ohgo, K.; Komatsu, K.; Kubota, A.; Asakura, T. *J. Biomol. NMR* **2003**, *25*, 91–103.
(8) Asakura, T.; Yao, J.; Yamane, T.; Umemura, K.; Ulrich, A. S. *J. Am. Chem. Soc.* **2002**, *124*, 8794–8795.
(9) Asakura, T.; Yao, J. *Protein Sci.* **2002**, *11*, 2706–2713.
(10) Altman, G. H.; Diaz, F.; Jakuba, C.; Calabro, T.; Horan, R. L.; Chen, J.; Lu, H.; Richmond, J.; Kaplan, D. L. *Biomaterials* **2003**, *24*, 401–16.
(11) Ferrari, F. A.; Cappello, J. *Protein-Based Materials*; Birkhauser: Boston, 1997.
(12) Parkhe, A. D.; Cooper, S. J.; Atkins, E. D. T.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *Int. J. Biol. Macromol.* **1998**, *23*, 251–258.
(13) O'Brien, J.; Fahnestock, S.; Termonia, Y.; Gardner, K. *Adv. Mater.* **1998**, *10*, 1185–1195.
(14) Valluzzi, R.; Szela, S.; Avtges, P.; Kirschner, D.; Kaplan, D. *J. Phys. Chem. B* **1999**, *103*, 11382–11392.
(15) Wilson, D.; Valluzzi, R.; Kaplan, D. *Biophys. J.* **2000**, *78*, 2690–701.
(16) Domurado, D.; Vert, M. *J. Biomater. Sci., Polym. Ed.* **2007**, *18*, 287–301.
(17) Rijcken, C. J. F.; Soga, O.; Hennink, W. E.; van Nostrum, C. F. *J. Controlled Release* **2007**, *120*, 131–148.
(18) Khalatur, P. G.; Khokhlov, A. R. In *Advances in Polymer Science - Conformation-Dependent Design of Sequences in Copolymers I*; Springer: Berlin 2006; Vol. *195*, pp 1−100.
(19) Tye, C. E.; Rattray, K. R.; Warner, K. J.; Gordon, J. A. R.; Sodek, J.; Hunter, G. K.; Goldberg, H. A. *J. Biol. Chem.* **2003**, *278*, 7949–7955.
(20) Fisher, L. W.; Torchia, D. A.; Fohr, B.; Young, M. F.; Fedarko, N. S. *Biochem. Biophys. Res. Commun.* **2001**, *280*, 460–5.
(21) Kummerlen, J.; vanBeek, J. D.; Vollrath, F.; Meier, B. H. *Macromolecules* **1996**, *29*, 2920–2928.
(22) Yamane, T.; Umemura, K.; Nakazawa, Y.; Asakura, T. *Macromolecules* **2003**, *36*, 6766–6772.
(23) Yamane, T.; Umemura, K.; Asakura, T. *Macromolecules* **2002**, *35*, 8831–8838.
(24) Asakura, T.; Sato, H.; Moro, F.; Yang, M.; Nakazawa, Y.; Collins, A. M.; Knight, D. *Macromolecules* **2007**, *40*, 8983–8990.
(25) Herzfeld, J.; Berger, A. E. *J. Chem. Phys.* **1980**, *73*, 6021–30.
(26) Ohgo, K.; Niemczura, W. P.; Ashida, J.; Okonogi, M.; Asakura, T.; Kumashiro, K. K. *Biomacromolecules* **2006**, *7*, 3306–10.
(27) Ishida, M.; Asakura, T.; Yokoi, M.; Saito, H. *Macromolecules* **1990**, *23*, 88–94.
(28) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67–81.
(29) Saito, H.; Tabeta, R.; Asakura, T.; Iwanaga, Y.; Shoji, A.; Ozaki, T.; Ando, I. *Macromolecules* **1984**, *17*, 1405–1412.
(30) Asakura, T.; Kuzuhara, A.; Tabeta, R.; Saito, H. *Macromolecules* **1985**, *18*, 1841–1845.
(31) Yao, J.; Ohgo, K.; Sugino, R.; Kishore, R.; Asakura, T. *Biomacromolecules* **2004**, *5*, 1763–9.
(32) Iwadate, M.; Asakura, T.; Williamson, M. P. *J. Biomol. NMR* **1999**, *13*, 199–211.