

Comparative Structure Analysis of Tyrosine and Valine Residues in Unprocessed Silk Fibroin (Silk I) and in the Processed Silk Fiber (Silk II) from *Bombyx mori* Using Solid-State ^{13}C , ^{15}N , and ^2H NMR[†]

Tetsuo Asakura,* Rena Sugino, Juming Yao, Hidehiko Takashima, and Raghuvansh Kishore[‡]

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

Received October 9, 2001

ABSTRACT: The solid-state ^{13}C CP-MAS NMR spectra of biosynthetically labeled [$^{13}\text{C}_\alpha$]Tyr, [$^{13}\text{C}_\beta$]Tyr, and [$^{13}\text{C}_\alpha$]Val silk fibroin samples of *Bombyx mori*, in silk I (the solid-state structure before spinning) and silk II (the solid-state structure after spinning) forms, have been examined to gain insight into the conformational preferences of the semicrystalline regions. To establish the relationship between the primary structure of *B. mori* silk fibroin and the “local” structure, the conformation-dependent ^{13}C chemical shift contour plots for Tyr C_α , Tyr C_β , and Val C_α carbons were generated from the atomic coordinates of high-resolution crystal structures of 40 proteins and their characteristic ^{13}C isotropic NMR chemical shifts. From comparison of the observed Tyr C_α and Tyr C_β chemical shifts with those predicted by the contour plots, there is strong evidence in favor of an antiparallel β -sheet structure of the Tyr residues in the silk fibroin fibers. On the other hand, Tyr residues take a random coil conformation in the fibroin film with a silk I form. The Val residues are likely to assume a structure similar to those of Tyr residues in silk fiber and film. Solid-state ^2H NMR measurements of [3,3- $^2\text{H}_2$]Tyr-labeled *B. mori* silk fibroin indicate that the local mobility of the backbone and the C_α – C_β bond is essentially “static” in both silk I and silk II forms. The orientation-dependent (i.e., parallel and perpendicular to the magnetic field) solid-state ^{15}N NMR spectra of biosynthetically labeled [^{15}N]Tyr and [^{15}N]Val silk fibers reveal the presence of highly oriented semicrystalline regions.

Silk, a fibrous protein produced by the domestic silkworm *Bombyx mori* (*B. mori*),¹ has a number of desirable properties, i.e., luster, comfort, high strength, and elasticity for textiles. In addition, recent studies clearly show that silk fibroin may find potential applications in the development and construction of various biotechnological and biomedical devices (1). In fact, the extensive use of the silk protein essentially originates from its unique amino acid composition translated into an unusual primary structure and hierarchical structural organization. The amino acid composition (in mol %) of *B. mori* fibrous protein showed the predominance of five amino acids: Gly (42.9%), Ala (30.0%), Ser (12.2%), Tyr (4.8%), and Val (2.5%) (2). Using the cDNA sequencing method, Mita et al. (3) and later Zhou et al. (4), employing shotgun sequencing strategy combined with traditional physi-

cal map-directed sequencing of the fibroin gene of the heavy chain, predicted the presence of unusual repeat sequences in the silk fibroin, shown in Figure 1. Their analysis revealed that the primary structure of *B. mori* silk fibroin may be approximately divided into four regions. The repetitive region R (i, ii, and iii) and amorphous region A (iv) appear alternatively along the chain: (i) highly repetitive GAGAGS sequences constituting the crystalline region, (ii) relatively less repetitive GAGAGY and/or GAGAGVGY sequences comprising semicrystalline regions containing mainly hydrophobic moieties, (iii) sequences similar to i plus an extension by AAS, and (iv) amorphous regions containing negatively charged, polar, bulky hydrophobic, and aromatic residues, for example, TGSSGFGPYVANGGYSGYEYAW-SSESDFGT. The superior nature of silk fibroin, making it an excellent natural fiber, basically originates from the combination of these unique amino acid sequences translating them into higher order structures.

The two crystalline forms, silk I and silk II, are reported as dimorphs of silk fibroin on the basis of extensive investigations from X-ray fiber diffraction (5–11), electron diffraction (8, 9, 11), conformational energy calculations (12, 13), infrared (14), and ^{13}C and ^{15}N solid-state NMR spectroscopy (14–25). The silk I form is the silk fibroin structure in the solid state before spinning, where the sample is obtained either as a film from liquid silk in the silk gland of the silkworm or as an aqueous solution of the regenerated silk fibroin. The silk II form is the structure after spinning

[†] This work was financially supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan. The work was carried out partly under a JSPS Invitation Fellowship offered to R.K.

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

[‡] On leave from the Institute of Microbial Technology, Sector 39-A, Chandigarh-160 036, India.

¹ Abbreviations: *B. mori*, *Bombyx mori*; CP-MAS, cross-polarization-magic angle spinning; DNA, deoxyribonucleic acid; FID, free induction decay; Fmoc, 9-fluorenylmethoxycarbonyl; HATU, *N*-(dimethylamino)-1*H*-1,2,3-triazolo[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide; NMR, nuclear magnetic resonance; PEG-PS, poly(ethylene glycol)–polystyrene; ppm, parts per million; τ_c , motional correlation time; TMS, tetramethylsilane; TFA, trifluoroacetic acid.

was collected from *B. mori* cocoon pieces after removing the sericin protein by degumming it twice with 0.5 w/v % Marseilles soap solution at 100 °C for about 30 min and washing it thoroughly with distilled water (14). The fibers thus obtained were used directly for the analysis. To obtain the film, frequently referred as regenerated silk (14), the silk fibroin was treated with 9 M LiBr solution separately, and after extensive dialysis against distilled water the clear solution was dried gently at room temperature for several hours.

Peptide Synthesis. Fmoc amino acids and all reagents used in peptide synthesis were procured from PerSeptive Biosystems, Warrington, U.K. The resin and HATU were from Applied Biosystems and PE Biosystems, respectively. The solvents of high-purity grade and other chemicals were available locally from Wako Pure Chemical Industries Ltd.

The two model peptides, **1** and **2**, were synthesized by solid-phase Fmoc chemistry on a fully automated Pioneer peptide synthesis system (24). The peptide was assembled on Fmoc-Gly-PEG-PS resin (0.19 mmol/g). The coupling of Fmoc amino acids were performed by HATU. After synthesis, the free peptides were released from the resin by treatment with a 40 mL mixture of TFA, phenol, triisopropylsilane, and water (88:5:2:5 vol %) for 2 h at room temperature. The crude peptide was precipitated and washed repeatedly with cold diethyl ether. The precipitate, collected by centrifugation, was dried in a vacuum, dissolved in 9 M LiBr, and dialyzed extensively against distilled water. The desired peptide was recovered by filtration and/or lyophilization. This is a widely accepted method for the preparation of *B. mori* silk fibroin with silk I structure (14, 15). To transform this structural form to silk II, the peptide was dissolved in a minimum amount of formic acid and dried at ambient temperature (34). In addition, the fraction of *B. mori* silk fibroin precipitated after chymotrypsin enzymatic cleavage (Cp fraction) was obtained with a silk I form, according to the method described earlier (35). The silk I structure of the Cp fraction was converted to the silk II form by formic acid treatment.

NMR Spectroscopy. ^{13}C CP-MAS NMR measurements were conducted on a Chemagnetics CMX-400 spectrometer operating at 100 MHz for ^{13}C . CP was employed for sensitivity enhancement with high-power ^1H decoupling during the signal acquisition interval. A ^1H 90° pulse width of 3 μs duration was used with a 1 ms contact time and a 3 s repetition time. Approximately 15K FIDs were added to generate the spectra. The chemical shifts were represented in parts per million with respect to the external reference adamantane. For direct comparison with previous data (14–17, 21, 26), we added 28.8 ppm to the observed chemical shifts to account for TMS as a reference on the silk I and silk II as well as the solution chemical shift data of the silk fibroins. The revised chemical shift values for the (AG)₁₅ sample of our previous work (24) are also listed in Table 1.

^2H NMR measurements were made on a 61.06 MHz Chemagnetics CMX Infinity 400 spectrometer equipped with suitable accessories. The details of the pulse sequence used have been described elsewhere (31, 32). The spectral width was 2000 kHz, 8192 data points were collected, and the spectra represent 8K–10K scans acquired at room temperature. Lorentzian line broadening of 4 kHz was applied. For the line-shape analysis the MXQET program was used to

simulate quadrupole echo ^2H NMR powder spectra as a function of the exchange rate and libration angle according to a suitable dynamic model, as previously described (31, 32).

The solid-state ^{15}N NMR measurements were performed at 25 °C on a JEOL GX400 spectrometer operating at 40.4 MHz equipped with a probe of 7 mm inner diameter. A contact time of 3 ms and a repetition time of 5 s were used to accumulate a total of 5K–20K scans at room temperature. $^{15}\text{NH}_4\text{NO}_3$ was used as an external reference. The ^{15}N solution NMR spectra were observed with a JEOL FX-90Q NMR spectrometer operating at an ambient temperature. Cross-polarization with high-power ^1H decoupling was employed as described (36).

Preparation of ^{13}C Chemical Shift Contour Maps. Contour plots of the conformation-dependent chemical shifts of the Tyr C $_{\alpha}$, Tyr C $_{\beta}$, and Val C $_{\alpha}$ carbons were generated from the known chemical shift data of 40 proteins whose X-ray coordinates were available from the Brookhaven Protein Data Bank. The relevant information about the proteins with their PDB codes has been provided in our previous publications (27). The conformation-dependent secondary shift ($\Delta\delta = \delta_{\text{observed}} - \delta_{\text{random}}$) is defined as the difference between the observed ^{13}C NMR chemical shift and the shift of the corresponding residues in the random/unordered structures. Such random coil chemical shift values were taken from the chemical shifts of the corresponding amino acid residues in *B. mori* silk fibroin in aqueous solution (37): 55.4 ppm for Tyr C $_{\alpha}$, 36.1 ppm for Tyr C $_{\beta}$, and 59.8 ppm for Val C $_{\alpha}$. The steps involved in the construction of the contour maps have been described in another paper (27).

RESULTS

Solid-State ^{13}C CP-MAS NMR Spectra of [$^{13}\text{C}_{\alpha}$]Tyr-, [$^{13}\text{C}_{\beta}$]Tyr-, and [$^{13}\text{C}_{\alpha}$]Val-Labeled Silk Fibroins. The ^{13}C CP-MAS NMR spectra of native *B. mori* silk fibroin (i.e., nonlabeled) and of silk fibroin enriched with ^{13}C at the Tyr C $_{\alpha}$ (or [$2\text{-}^{13}\text{C}$]Tyr), Tyr C $_{\beta}$ (or [$3\text{-}^{13}\text{C}$]Tyr), and Val C $_{\alpha}$ (or [$2\text{-}^{13}\text{C}$]Val) sites are shown in Figures 2 and 3 for both the silk I (film) and silk II (fiber) forms, respectively. Only the expanded spectra between 10 and 70 ppm are shown, and the observed chemical shifts are summarized in Table 1. For the sake of comparison, the solution-state ^{13}C NMR chemical shifts of the silk fibroin have also been tabulated along with the ^{13}C isotropic chemical shifts from the CP-MAS experiments. By comparison of spectra **b**, **c**, and **d** with that of the nonlabeled material, **a**, the relatively increased peaks are assigned to the labeled carbons. To confirm these resonances, the spectrum of the native *B. mori* silk fibroin recorded under similar conditions was subtracted. As expected, the resulting difference spectra in Figure 4 contain the resonances arising predominantly from the labeled sites, while only small peaks in spectrum **a** are due to differences in the line shapes of the labeled and nonlabeled materials. Therefore, it is evident that the biosynthetic labeling technique described here successfully targets the C $_{\alpha}$ - and C $_{\beta}$ -labeled sites of the Tyr residues in the silk fibroin and may be extended to other related biological model systems. Characteristic solid-state ^{13}C NMR chemical shifts of the Tyr C $_{\alpha}$ carbon are ~56 and 54.2 ppm for film and fiber, respectively, and for Tyr C $_{\beta}$ ~36.0 and 40 ppm.

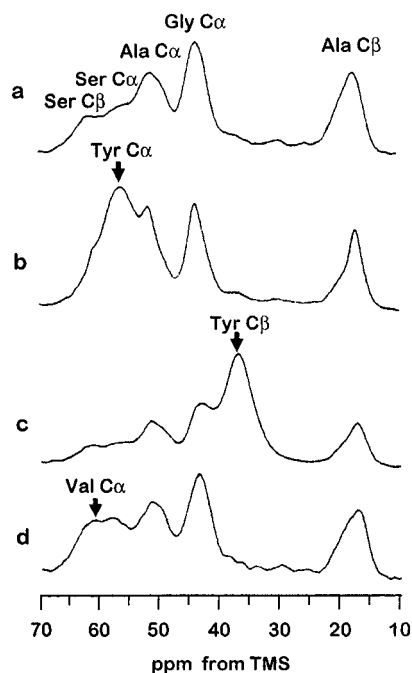


FIGURE 2: Solid-state ^{13}C CP-MAS NMR spectra (10–70 ppm) of *B. mori* silk fibroin with the silk I form (a) in its native state, i.e., unlabeled sample, (b) enriched with ^{13}C at the C_α site of the tyrosine residues, (c) enriched with ^{13}C at the C_β site of the tyrosine residues, and (d) enriched with ^{13}C at the C_α site of the valine residues.

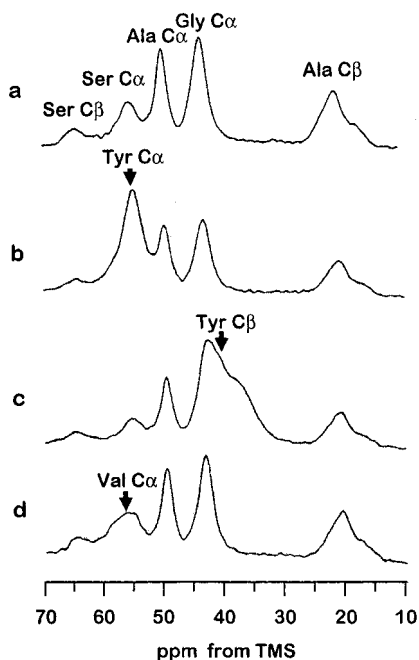


FIGURE 3: Solid-state ^{13}C CP-MAS NMR spectra (30–70 ppm) of *B. mori* silk fibroin with the silk II form (a) in its native state, i.e., unlabeled sample, (b) enriched with ^{13}C at the C_α site of the tyrosine residues, (c) enriched with ^{13}C at the C_β site of the tyrosine residues, and (d) enriched with ^{13}C at the C_α site of the valine residues.

Figures 2d and 3d also show the ^{13}C CP-MAS spectra of the *B. mori* silk fibroin that has been enriched at the C_α site of Val. To ascertain that the assigned resonances (~59 and 56.9 ppm for film and fiber, respectively) arise from the isotopically enriched Val C_α sites, again the spectrum of the nonlabeled native *B. mori* silk fibroin sample was

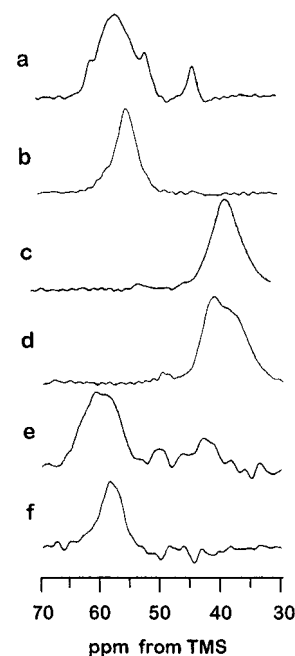


FIGURE 4: Solid-state ^{13}C CP-MAS NMR difference spectra (10–70 ppm) of labeled *B. mori* silk fibroins with the silk I (a, c, e) and silk II (b, d, f) structural forms: spectra a and b are the C_α sites of the tyrosine residues; spectra c and d are the C_β sites of the tyrosine residues, and spectra e and f represent the C_α sites of the valine residues.

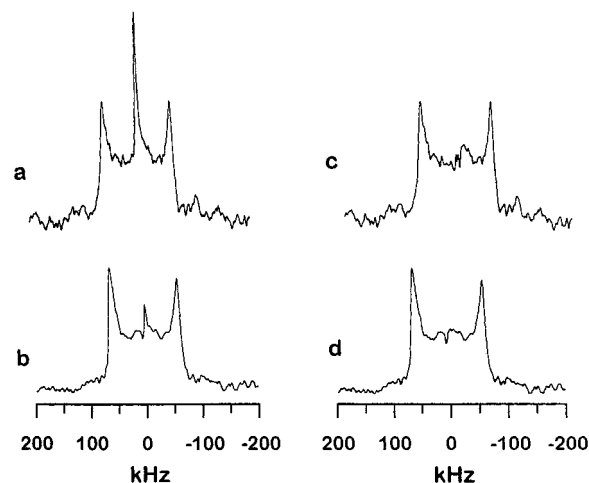


FIGURE 5: Solid-state ^2H NMR spectra (–200–200 ppm) of $[3,3\text{-}^2\text{H}_2]\text{Tyr}$ -labeled *B. mori* silk fibroin with the silk I (a and c) and silk II (b and d) structural forms. Spectra c and d represent results after the water peak at the center is subtracted.

subtracted from the spectrum of the corresponding labeled sample. The difference NMR spectra of the silk I and silk II structural forms (Figure 4) clearly indicate that the labeling technique has successfully targeted the Val residues, too. From these results it appears that in solid state the Tyr C_α , Tyr C_β , and Val C_α resonances are conformationally sensitive.

Solid-State ^2H NMR Spectra of $[3,3\text{-}^2\text{H}_2]\text{Tyr}$ -Labeled Silk Fibroins. We examined the solid-state ^2H NMR spectra of $[3,3\text{-}^2\text{H}_2]\text{Tyr}$ -labeled silk fibroin in Figure 5 in the silk I (film) and silk II (fiber) structural forms. The observed line shapes could be satisfactorily simulated with an asymmetry parameter of $\eta = 0.00$ and a quadrupole coupling constant

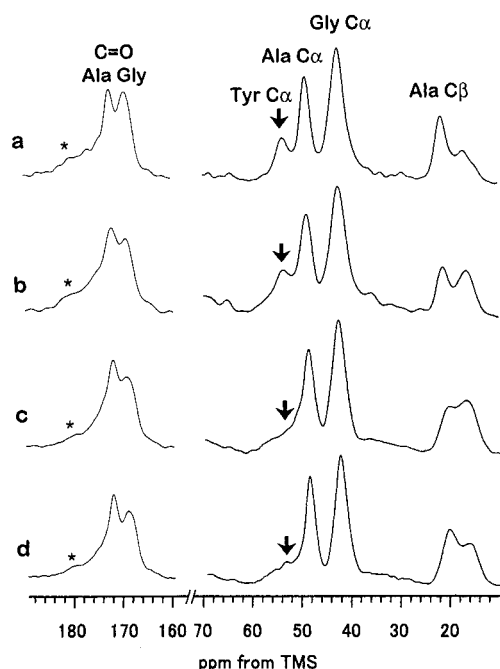


FIGURE 6: Solid-state ^{13}C CP-MAS NMR spectra (10–70 ppm and 160–190 ppm) of the two model peptides, **1** (a, b) and **2** (c, d). Spectra **a** and **c** were obtained after LiBr treatment and dialysis against distilled water for **1** and **2**, respectively. Spectra **b** and **d** were obtained after formic acid treatment for **1** and **2**, respectively. For both peptides the characteristic doublets of the Ala C_β resonances are clearly seen. The broad appearance of the Tyr C_α resonance in **2** is indicated by an arrow.

of $Q_{\text{cc}} = 164$ kHz (31). On the basis of the model of a three-site jump around the $\text{C}_\alpha\text{--C}_\beta$ bond, a very slow rate constant of 10^3 Hz was obtained as reported previously for $[3,3\text{-}^2\text{H}_2]$ -Tyr-labeled silk fibroin in the silk II form (31). Thus, the rotation about the $\text{C}_\alpha\text{--C}_\beta$ bond of the Tyr residues can be considered as essentially “static” for both of the samples. Therefore, we suggest that in *B. mori* silk fibroin the origin of the peak broadening may be due to the distribution of the “local” structural environment of the ^{13}C -labeled Tyr nuclei, presumably associated with multiple secondary structures of Tyr. We attribute a similar explanation also to the observed broad peaks of the Val C_α resonance.

Solid-State ^{13}C CP-MAS NMR Spectra of Model Peptides **1 and **2**.** Two model peptides, **1** and **2**, were synthesized to examine the influence of the presence of Tyr and Val residues on the structure of the (Ala-Gly) $_{15}$ sequence. Such information is usually difficult to obtain directly from silk fibroin samples. The solid-state ^{13}C CP-MAS NMR spectra of these peptides, after dissolution in 9 M LiBr solution and extensive dialysis against water, are shown in Figure 6, and the characteristic chemical shifts are summarized in Table 1, along with formic acid-treated samples. The resonance assignments of the constituent amino acids could be readily made from their relative peak area and by comparison with the spectra reported previously for several related silk fibroin model peptides (25). Both peptides gave remarkably similar ^{13}C chemical shifts; however, a marked difference in the relative intensities of the conformationally sensitive Ala C_β resonance (~ 16.5 and ~ 21.0 ppm) was apparent. It should be noted that the 9 M LiBr-treated (AG) $_{15}$ sequence yields a single sharp peak at ~ 16.5 ppm for the Ala C_β resonance which could be assigned *exclusively* to the silk I structural

form (14–16). Thus, the observed chemical shifts and the relative peak intensities of the Ala, Gly, and Tyr residues in **1** and the Ala, Gly, Tyr, and Val residues in **2** clearly indicate that both peptides assume a fiberlike silk II conformation rather than silk I, even after LiBr treatment, according to the summarized values in Table 1. The Tyr residues exhibit a dramatic destabilizing effect on the silk I structure, as judged from the characteristic doubling of the Ala C_β resonance at ~ 16.5 and ~ 21.0 ppm representing silk I and silk II, respectively. Of particular interest is the finding that the Tyr C_α resonance at 53.8 ppm (marked by arrow in peptide **1**) is closer to the fibrous form, which may be characteristic of predominantly β -sheet structures. The Tyr C_β peaks at ~ 38 ppm remain significantly broadened. A careful interpretation of the results, in conjunction with the appearance of broad peaks for the Ala C_β and Ala carbonyl carbons at ~ 16.5 and ~ 177 ppm (marked by an asterisk), respectively, may suggest the existence of diverse structural forms of a “distorted β -turn” type (unpublished results). This interpretation is fully consistent with our recent proposal, based on experimental NMR as well as theoretical molecular simulation structural characterizations, of a “heterogeneous” structure of the silk II form. Further, the analysis of NMR spectra of **2** suggests that the presence of a single Val residue, as such, may not have any dramatic influence, particularly when it coexists with almost uniformly distributed Tyr residues. The simultaneous appearance of the characteristic silk II resonances in both **1** and **2**, as also analyzed from the split patterns of the Ala C_β resonance, provides evidence that these residues stabilize the silk II structure, i.e., an antiparallel β -sheet. The overall results on the model peptides clearly show that the effect of the Val residue is significant and nonnegligible. The present data, however, do not allow a quantitative evaluation of the destabilizing and stabilizing effects on the silk I and silk II structures, respectively, exerted by these two residues.

^{13}C NMR Chemical Shift Contour Maps of Tyr and Val Residues. Figures 7 and 8 show the contour plots of the conformation-dependent ^{13}C chemical shifts of the Tyr C_α and C_β carbons, respectively. The $\Delta(\phi, \psi)$ values in the left half of the Ramachandran map (i.e., $-180^\circ < \phi < 0^\circ$ and $-180^\circ < \psi < 180^\circ$) are reported. The analysis shows that indeed there exists a conformation dependence of the chemical shifts. For example, the chemical shifts falling in the helical region ($\phi \sim -57 \pm 20^\circ$; $\psi \sim -47 \pm 20^\circ$) of the ϕ, ψ map appear at much lower field (~ 58.0 ppm) for the Tyr C_α resonance than those values (~ 53.5 ppm) appearing in the β -sheet region ($\phi \sim -139 \pm 20^\circ$; $\psi \sim 135 \pm 20^\circ$). Interestingly, a reverse trend was observed for the Tyr C_β resonance, i.e., upfield chemical shifts for the helical region (~ 35.5 ppm) and downfield shifts (~ 38.5 ppm) for the extended β -sheet region. The Tyr carbon chemical shifts of silk fiber in the silk II form are examined in the contour plots in Figures 7 and 8. The observed chemical shifts, 54.2 ppm (Tyr C_α) and ~ 40 ppm (Tyr C_β) are marked by assuming an error in the chemical shifts of ± 0.5 ppm. The overlapped ϕ and ψ region, which satisfies both the Tyr C_α and Tyr C_β chemical shifts of silk fibers simultaneously, is located close to an extended β -sheet region in the Ramachandran map.

There are several areas which satisfy both the Tyr C_α and C_β chemical shifts of the silk film with silk I form. The

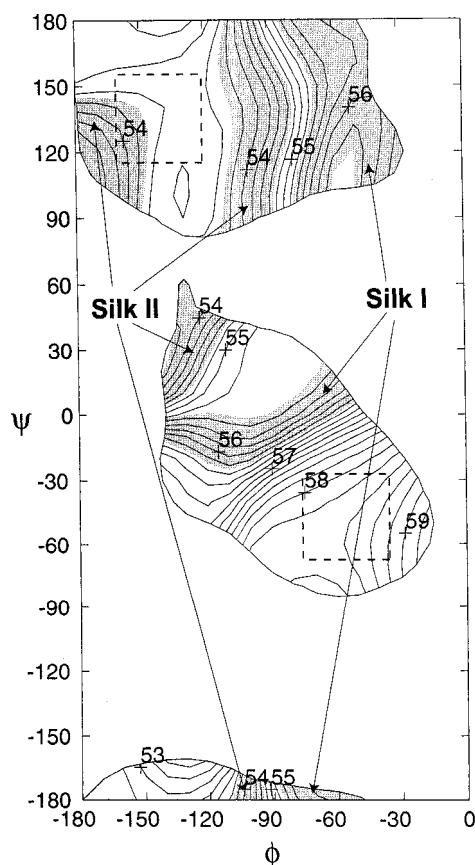


FIGURE 7: Contour plot of the conformation-dependent ^{13}C chemical shifts (in ppm) for the C_α carbon of Tyr residues in 40 proteins, shown as a function of torsion angles ϕ and ψ . Chemical shift values in the region $-180^\circ < \phi < 0^\circ$ and $-180^\circ < \psi < 180^\circ$ are shown where the density function is >1 . The random coil chemical shift for the Tyr C_α carbon considered is 55.4 ppm. The areas of silk I chemical shift (56.0 ± 0.5 ppm) and of silk II chemical shift (54.2 ± 0.5 ppm) are shown in gray. The boxes with broken lines indicate α -helical (lower) and β -sheet (upper) regions.

chemical shifts are the same as those in liquid silk, but in the latter case there is a rapid exchange among various conformations on the NMR time scale. However, in the solid state the ^2H NMR analysis of $[3,3\text{-}^2\text{H}_2]\text{Tyr}$ -labeled silk fibroin in the silk I form has shown that there is no interconversion among different conformations. Thus, it is likely that the Tyr residues assume several torsion angles in the silk film with the silk I form.

A similar conformation-dependent contour plot was constructed from the ^{13}C chemical shift values of the Val C_α carbon of known protein structures, as shown in Figure 9. The results indicate that the torsion angles, ϕ and ψ , exert a significant effect on the chemical shift values of the Val C_α carbons. The typical chemical shift values for α -helices are ~ 63.0 ppm (shift to lower field) and ~ 58.0 ppm (shift to higher field) for antiparallel β -sheets. A comparison of the observed ^{13}C chemical shifts of the Val C_α carbons with the contour map suggests that the most likely conformation for the Val residues in *B. mori* silk fiber (silk II form) is the extended β -sheet. On the other hand, we speculate that the Val residues in the silk film (silk I form) take several conformations that satisfy the chemical shift constraints, just as in the case of the Tyr residues in silk film.

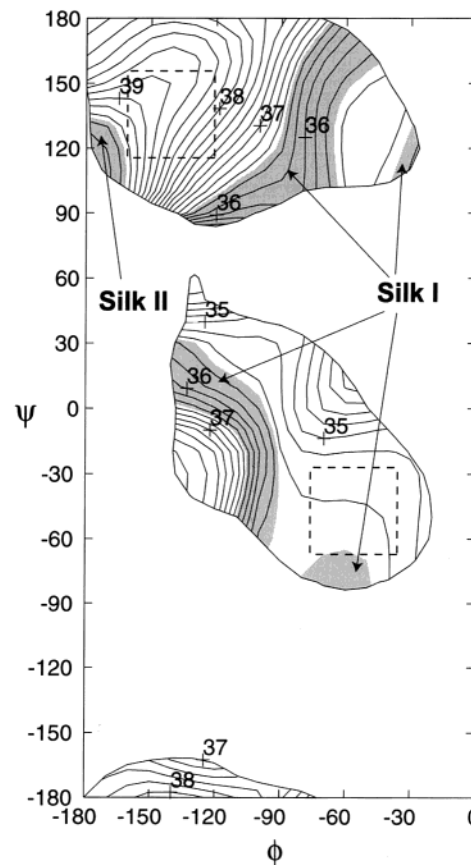


FIGURE 8: Contour plots of the conformation-dependent ^{13}C chemical shifts (in ppm) for the C_β carbon of Tyr residues in 40 proteins, shown as a function of torsion angles ϕ and ψ . Chemical shift values in the region $-180^\circ < \phi < 0^\circ$ and $-180^\circ < \psi < 180^\circ$ are shown where the density function is >1 . The random coil chemical shift for the Tyr C_β carbon considered is 36.1 ppm. The areas of silk I chemical shift (36.0 ± 0.5 ppm) and of silk II chemical shift (40.0 ± 0.5 ppm) are shown in gray. The boxes with broken lines indicate α -helical (lower) and β -sheet (upper) regions.

Solid-State ^{15}N NMR Spectra of Oriented Silk Fibroin Fibers. Figure 10 shows the ^{15}N NMR spectra of ^{15}N Tyr- and ^{15}N Val-labeled *B. mori* silk fibroins in aqueous solution, along with the spectrum of ^{15}N Ser-labeled fibroin (33). For comparison, the natural abundance ^{15}N NMR spectrum (lower most) of the native *B. mori* silk fibroin, obtained under similar conditions, is also shown. In the spectrum of each silk fibroin-labeled sample, only one major ^{15}N resonance is observed. The small peaks in the NMR spectra are attributed to the natural abundance in the samples. It should be mentioned that the nitrogen nuclei of the amide bonds in proteins, as donors, frequently participate in diverse hydrogen-bonding interactions in different secondary structures, along with other long-range interactions. A wide distribution of the shifts in the native silk fibroin may indicate that these moieties are likely to be distributed in different conformational states. We observed relatively low enrichment of ^{15}N Val silk fibroin samples which could be due to low contents of the Val residues in the silk fibroin molecules; however, the intensity of the peak is still significant to the extent that the sample can be exploited for the solid-state NMR analysis. Thus, the results demonstrate that isotopic labeling of selected residues during the cultivation of the posterior silk gland of *B. mori* silk fibroin can yield

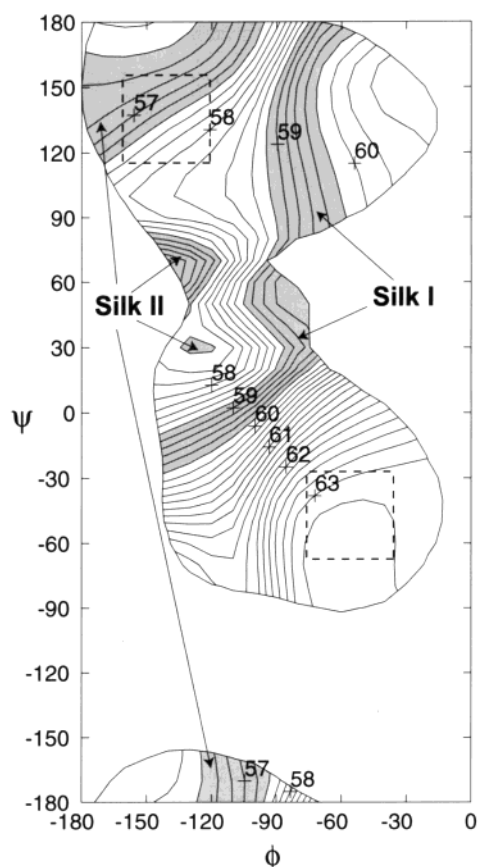


FIGURE 9: Contour plots of the conformation-dependent ^{13}C chemical shifts (in ppm) for the C_α carbon of Val residues in 40 proteins, shown as a function of torsion angles ϕ and ψ . Chemical shift values in the region $-180^\circ < \phi < 0^\circ$ and $-180^\circ < \psi < 180^\circ$ are shown where the density function is >1 . The random coil chemical shift for the Val C_α carbon is considered 59.8 ppm. The areas of the silk I chemical shift (59.0 ± 0.5 ppm) and the silk II chemical shift (56.9 ± 0.5 ppm) are shown in gray. The boxes with broken lines indicate α -helical (lower) and β -sheet (upper) regions.

high ^{15}N incorporation for these three residues. This approach may be useful for the preparation of other related ^{15}N -labeled biomaterials in their native forms for structural investigations.

To gain insight into the structural organization of the fibroin molecules, we investigated the ^{15}N spectra of ^{15}N -labeled Tyr, Val, and Ser silk fibers in the silk II forms in the solid state. Oriented samples were obtained as a block, which was placed with the fiber axis either parallel or perpendicular to the applied magnetic field (Figure 11). All of the experimental spectra were found to contain a contribution of approximately 30% of a nonoriented powder pattern which was subtracted from the observed spectra (18, 23). The resulting spectral patterns in Figure 11 show significant differences between the parallel and perpendicular alignments, indicating that the Tyr and Val sites are well oriented within the macroscopic fiber. Since the spectral patterns are very similar to those of ^{15}N -labeled Gly- and Ala-labeled silk fiber samples reported previously, we conclude that the long silk fibroin chain represents an antiparallel β -sheet structure. Thus, most of the Tyr and Val sites, incorporated into the semicrystalline domains of repeated -GAGAGS- sequences, are suggested to assume the silk II structure.

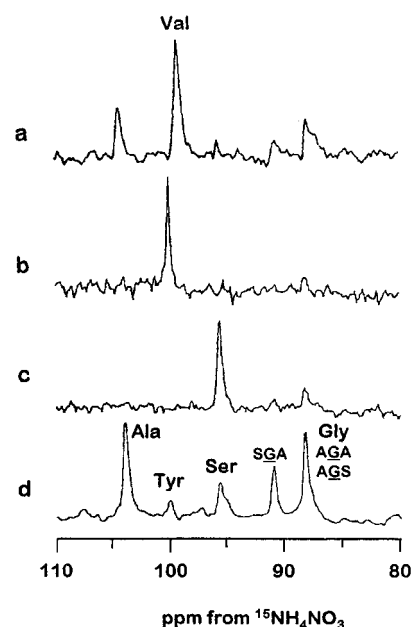


FIGURE 10: ^{15}N NMR solution spectra (80–110 ppm) of selectively labeled *B. mori* silk fibroin samples: (a) ^{15}N -labeled Val silk fibroin, (b) ^{15}N -labeled Tyr silk fibroin, and (c) ^{15}N -labeled Ser silk fibroin. For comparison, the ^{15}N NMR spectrum (d) of an unlabeled (i.e., natural abundance) silk fibroin sample is also shown.

DISCUSSION

The conformation of the silk II form, reported by Marsh et al. using X-ray diffraction analysis, has been established as an array of antiparallel β -sheets oriented such that all of the Ala methyl groups and Gly residues face opposite sides, thus giving rise to a “polar-antiparallel” arrangement (5). A monoclinic unit cell was proposed where the chains in consecutive sheets are staggered when viewed along the intersheet distance. These general features have largely been supported by Fraser and co-workers (6) by an independent analysis of the X-ray diffraction patterns of an $(\text{AG})_n$ model system, and the intersheet distances were further clarified by Lotz and Cesari (11, 12). Therefore, to a first approximation, the amino acid sequence of the crystalline region comprising the β -sheet structure is an alternating sequence of Ala and Gly residues. Interestingly, a reexamination of the X-ray diffraction pattern of *B. mori* silk fibroin fiber by Takahashi et al. (7) revealed an “antipolar antiparallel” model, composed of antiparallel chains where the methyl groups of the Ala residues alternately lie on both sides of the β -sheet, since this refined model provided a “better fit” to the observed diffraction pattern. Fraser et al. (6) reported that the diffraction pattern of the crystalline Cp fraction (the precipitate obtained from an aqueous solution of fibroin treated with pancreatic chymotrypsin) was similar to that of powdered silk fibroin and to that of the model polypeptide $(\text{GAGAGS})_n$ containing Ser residues. This finding is substantiated by our own studies (32) where we showed that the Ser residues act as a stabilizing entity of the basic $(\text{AG})_n$ β -sheet structure through hydrogen bonding of the side-chain hydroxyl groups with carbonyl groups on adjacent chains.

The effect of individual amino acid residues on the conformational preferences within globular proteins has been extensively studied by a number of researchers (38). According to these studies, the distribution of Tyr residues is significantly below average in helical regions, and Tyr does

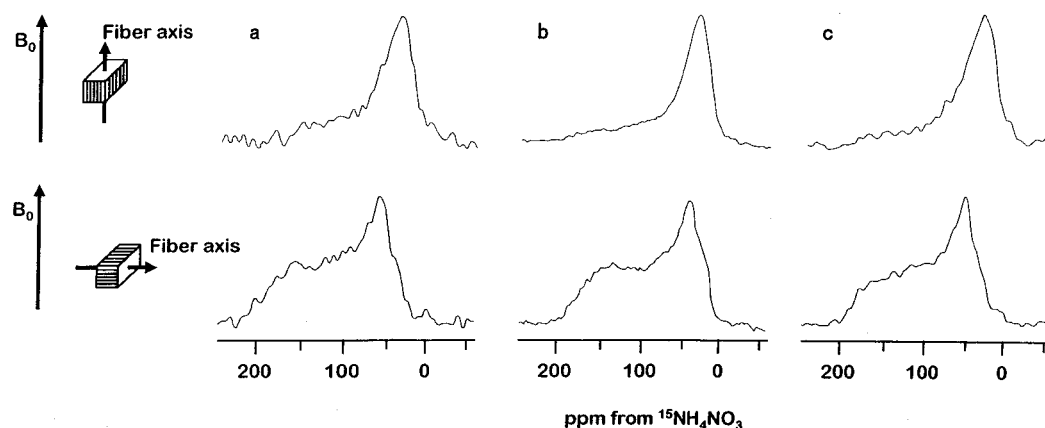


FIGURE 11: Solid-state ^{15}N NMR spectra (0–200 ppm) of the labeled, uniaxially aligned *B. mori* silk fibroin fiber: (a) ^{15}N [Tyr] fiber, (b) ^{15}N [Val] fiber, and (c) ^{15}N [Ser] fiber. The fiber axis was placed parallel (upper spectra) and perpendicular (lower spectra) to the static magnetic field direction B_0 . Approximately 30% of the intensity arising from nonoriented powder patterns was subtracted from the experimental spectra.

not exhibit a propensity for forming helical structures. Similarly, the distribution of Val residues was shown to be significantly above average in β -sheet regions. Hence, our findings substantiate the propensity of Tyr and Val residues in globular proteins, which is reported to be considerably higher for β -sheet structures (39, 40).

Relatively few studies have reported on the structural behavior of the Tyr residue within *B. mori* silk fibroin. Using ^{13}C spin–lattice relaxation, the dynamic feature of the Tyr side chains in *B. mori* silk fibroin was investigated by Saito et al. (41). The dynamic behavior of the Tyr side chains in *B. mori* silk, reported by Kameda et al. (31) using solid-state ^2H NMR spectroscopy, revealed that $[3,3\text{-}^2\text{H}_2]$ Tyr-labeled silk fibroins showed a typical rigid powder pattern, indicating that there is essentially no motion about the $\text{C}_\alpha\text{—C}_\beta$ bond axis, thereby revealing the existence of distinct rotameric populations within the oriented domains. On the contrary, the ^2H NMR spectra of silk fibers with ^2H -labeled aromatic rings, that is, $[3',5'\text{-}^2\text{H}_2]$ Tyr-labeled silk fibroins, exhibited two dynamic components. Here, the presence of a rigid powder contribution and a motionally averaged contribution indicated that some of the Tyr side chains are in fact mobile at the phenolic ring. This motion was characterized to be a π -flip, typical of aromatic rings. Moreover, the ^2H NMR line shape could be simulated by attributing 20% of the signal to a motionally averaged component with a fast rate of $\sim 10^6$ Hz, while the remaining 80% reflected a much slower component ($< 10^3$ Hz). To our knowledge, there has been no report on the dynamics of Val residue(s) in *B. mori* silk fibroin, yet.

An apparent $\text{p}K_a$ value of ~ 10.1 for the phenolic hydroxyl group of Tyr in silk fibroin indicated a favorable deprotonation (37). This largely precludes the formation of hydrogen-bonding interactions involving the hydroxyl groups. The absence of hydrogen bonding is likely to promote the π -flip motions of the phenolic ring as revealed by our ^2H NMR study of $[3',5'\text{-}^2\text{H}_2]$ Tyr-labeled silk fibroin (31). Indeed, by enabling the π -flip motions to occur, the model proposed by Takahashi et al. (7), where the methyl groups of the Ala residues alternately point to either side of the sheet, is more feasible. The proposed antipolar-antiparallel arrangements of the fibroin chains are expected to bring relatively more local disorder in the stacking of sheets and, consequently, the bulky side-chain dispositions.

The driving force for self-aggregation of the fibroin molecules into a β -sheet structure is most likely a hydrophobic exclusion of the solvent by the $(\text{AG})_n$ sequence, and the role of the Ser residue is presumably to stabilize the β -sheet structure through hydrogen bonding. The occurrence of bulky aromatic side chains of Tyr within the silk I structure possibly exerts a disruptive effect on the hydrophobic interactions between the basic $(\text{AG})_n$ chains. Interestingly, our results show that the Tyr residues in the “oriented domains” of *B. mori* silk fibroin (silk II) are highly oriented in such a way that the ϕ , ψ torsional angles of this residue can easily accommodate a β -sheet conformation.

By comparing the ^{13}C NMR chemical shifts and the X-ray structures of proteins, Spera and Bax (42) and we (27) suggested that the $\Delta\delta$ values for the C_α and C_β resonances correlate quantitatively well with the backbone torsional angle obtained from the crystal structures. The ϕ and ψ values determined from the ^{15}N NMR experiments of silk fibroin are nicely consistent with the range predicted by the contour plots constructed from the observed ^{13}C chemical shifts of the C_α and C_β carbons of the Tyr residues in other proteins. However, the packing of the polypeptide chains into β -sheets in the vicinity of the Tyr residue is suggested to be less tight than around the Ser residues, probably due to the disordered side-chain orientations, described above (7). It is also worth mentioning that the diffuse streak scattering observed between the layer lines of the X-ray diffraction patterns by Takahashi et al. has been interpreted as arising from the periodic occurrence of the Ser residue within the β -sheet structure. Nevertheless, the presence of Tyr residues within the β -sheet structure may equally explain this effect.

CONCLUSION

The results of the present investigation have established that *B. mori* native silk fibroin can be successfully labeled biosynthetically at specific main-chain positions with ^{13}C [Tyr], ^{13}C [Val], ^{15}N [Tyr], ^{15}N [Val], and ^{15}N [Ser] amino acids. The simple method can be extended to other related biological systems for site-specific labeling of particular residues in fibroin-like molecules for structural characterizations. Given that the repeated $(\text{AG})_n$ sequence in the silk II form confers the basic β -sheet structure to *B. mori* silk fibroin, the role

of the short polar side chain of Ser is suggested to stabilize the β -sheet structure via side-chain to main-chain hydrogen bonding (32). Employing ^{13}C CP-MAS NMR, we have established that the bulky hydrophobic side chains of Tyr and Val may readily be included in the β -sheet structure of the *B. mori* silk fiber. Here, it appears that in the silk I form (and/or LiBr-treated samples) particularly the Tyr residues exhibit a larger distribution of their torsion angles (i.e., unordered structures), whereas in the silk II form the distribution is significantly restricted to an extended β -sheet conformation. The ^{13}C CP-MAS NMR analysis of the model peptides **1** and **2** also indicates that both the position and the relative content of Tyr in the $(\text{AG})_n$ model system can be critical in causing the structural change from silk I to silk II, in a quantitative manner (unpublished data). A more systematic analysis of the simultaneous destabilizing effects, exerted by these residues on the silk I structure, is currently being analyzed in our laboratory with the help of designed model peptides. The current analysis also suggests that the ^{13}C NMR chemical shift values of the Tyr C_α resonance may be more sensitive to the backbone conformation accommodating the β -sheet structures, whereas the dependence of the Tyr C_β and Val C_α chemical shifts on the backbone conformation met with limited success. Finally, our experiments clearly indicate that a number of different solid-state NMR techniques are required to obtain detailed knowledge about the intra- and intermolecular conformational geometry, particularly in the case of fibrous proteins such as silk.

REFERENCES

- Asakura, T., and Kaplan, D. L. (1994) in *Encyclopedia of Agricultural Science* (Arutzen, C. J., Ed.) Vol. 4, pp 1–11, Academic Press, New York.
- Shimura, K. (1980) *Zoku Kenshi no Kozo* (Hojoyo, N., Ed.) Shinshu University, Ueda.
- Mita, K., Ichimura, S., and James, C. T. (1994) *J. Mol. Evol.* 38, 583–592.
- Zhou, C., Confalonieri, F., Medina, N., Zivanovic, Y., Esnault, C., Yang, T., Jacquet, M., Janin, J., Duguet, M., Perasso, R., and Li, Z. (2000) *Nucleic Acids Res.* 28, 2413–2419.
- Marsh, R. E., Corey, R. B., and Pauling, L. (1955) *Biochim. Biophys. Acta* 16, 1–34.
- Fraser, R. D. B., and MacRae, T. P. (1973) in *Conformations of Fibrous Proteins and Related Synthetic Polypeptides*, Academic Press, New York.
- Takahashi, Y., Gehoh, M., and Yuzuriha, K. (1999) *Int. J. Biol. Macromol.* 24, 127–138.
- Konishi, T., and Kurokawa, M. (1968) *Sen'i Gakkaishi* 24, 550–554.
- Okuyama, K., Takanashi, K., Nakajima, Y., Hasegawa, Y., Hirabayashi, K., and Nishi, N. (1988) *J. Seric. Sci. Jpn.* 57, 23–30.
- Anderson, J. P. (1998) *Biopolymers* 45, 307–321.
- Lotz, B., and Keith, H. D. (1971) *J. Mol. Biol.* 61, 201–215.
- Lotz, B., and Cesari, F. C. (1979) *Biochimie* 61, 205–214.
- Fossey, S. A., Nemethy, G., Gibson, K. D., and Scheraga, H. A. (1991) *Biopolymers* 31, 1529–1541.
- Asakura, T., Kuzuhara, A., Tabeta, R., and Saito, H. (1985) *Macromolecules* 18, 1841–1845.
- Saito, H., Tabeta, R., Asakura, T., Iwanaga, Y., Shoji, A., Ozaki, T., and Ando, I. (1984) *Macromolecules* 17, 1405–1412.
- Asakura, T., and Yamaguchi, T. (1987) *J. Seric. Sci. Jpn.* 56, 300–304.
- Ishida, M., Asakura, T., Yokoi, M., and Saito, H. (1990) *Macromolecules* 23, 88–94.
- Nicholson, L. K., Asakura, T., Demura, M., and Cross, T. A. (1993) *Biopolymers* 33, 847–861.
- Asakura, T., Aoki, A., Demura, M., Joers, J. M., Rosanske, R. C., and Gullion, T. (1994) *Polym. J.* 26, 1405–1408.
- Asakura, T., Demura, M., Hiraishi, Y., Ogawa, K., and Uyama, A. (1994) *Chem. Lett.*, 2249–2252.
- Asakura, T., Demura, M., Date, T., Miyashita, N., Ogawa, K., and Williamson, M. P. (1997) *Biopolymers* 41, 193–203.
- Asakura, T., Minami, M., Shimada, R., Demura, M., Osanai, M., Fujito, T., Imanari, M., and Ulrich, A. S. (1997) *Macromolecules* 30, 2429–2435.
- Demura, M., Minami, M., Asakura, T., and Cross, T. A. (1998) *J. Am. Chem. Soc.* 120, 1300–1308.
- Asakura, T., Ashida, J., Yamane, T., Kameda, T., Nakazawa, Y., Ohgo, K., and Komatsu, K. (2001) *J. Mol. Biol.* 306, 291–305.
- Asakura, T., Ashida, J., and Yamane, T. (2001) in *High-Resolution NMR Spectroscopy of Polymers*, American Chemical Society, Washington, DC (in press).
- Asakura, T., Iwade, M., Demura, M., and Williamson, M. P. (1999) *Int. J. Biol. Macromol.* 24, 167–171.
- Iwade, M., Asakura, T., and Williamson, M. P. (1999) *J. Biomol. NMR* 13, 199–211.
- Ulrich, A. S., and Grage, S. L. (1998) in *Solid State NMR of Polymers* (Ando, I., and Asakura, T., Eds.) pp 190–211.
- Kameda, T., and Asakura, T. (2001) in *Annual Report on NMR Spectroscopy* (Webb, G. A., Ed.) Academic Press, London (in press).
- Saito, H., Tabeta, R., Kuzuhara, A., and Asakura, T. (1986) *Bull. Chem. Soc. Jpn.* 59, 3383–3387.
- Kameda, T., Ohkawa, Y., Yoshizawa, K., Nakano, E., Hiraoki, T., Ulrich, A. S., and Asakura, T. (1999) *Macromolecules* 32, 8491–8495.
- Kameda, T., Ohkawa, Y., Yoshizawa, K., Naito, J., Ulrich, A. S., and Asakura, T. (1999) *Macromolecules* 32, 7166–7171.
- Asakura, T., Sakaguchi, R., Demura, M., Manabe, T., Uyama, A., Ogawa, K., and Osanai, M. (1993) *Biotechnol. Bioeng.* 41, 245–252.
- Panitch, A., Matsuki, K., Cantor, E. J., Cooper, S. J., Atkins, E. D. T., Fournier, M. J., Mason, T. L., and Tirrell, D. A. (1997) *Macromolecules* 30, 42–49.
- Asakura, T., Watanabe, Y., and Itoh, T. (1984) *Macromolecules* 17, 2421–2426.
- Asakura, T., Yoshimizu, H., and Yoshizawa, F. (1988) *Macromolecules* 21, 2038–2041.
- Asakura, T., Watanabe, Y., Uchida, A., and Minagawa, H. (1984) *Macromolecules* 17, 1075–1081.
- Richardson, J. S., and Richardson, D. V. (1990) in *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G. D., Ed.) pp 1–98, Plenum Press, New York.
- Minor, D. L., Jr., and Kim, P. S. (1994) *Nature* 367, 660–663.
- Kim, C. A., and Berg, J. M. (1993) *Nature* 362, 267–270.
- Saito, H., Ishida, M., Yokoi, M., and Asakura, T. (1990) *Macromolecules* 23, 83–88.
- Spera, S., and Bax, A. (1991) *J. Am. Chem. Soc.* 113, 5490–5492.

BI0119013