Structural Analyses of *Anaphe* Silk Fibroin and Several Model Peptides Using ^{13}C NMR and X-ray Diffraction Methods

Chikako Tanaka,[†,‡] Rui Takahashi,[†] Atsushi Asano,[‡] Takuzo Kurotsu,[‡] Hiromu Akai,[§] Kazuhiko Sato,[||] David P. Knight,[#] and Tetsuo Asakura*[,†]

Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan, Department of Applied Chemistry, National Defense Academy, Yokosuka, Kanagawa 239-8686, Japan, Laboratory of Entomology, Tokyo University of Agriculture, Setagaya, Tokyo 156-8502, Japan, Material Analysis Research Laboratories, Teijin Ltd., Hino, Tokyo 191-8512, Japan, and Oxford Biomaterials Ltd., Units 14-15 Galaxy House, Thatcham RG19 6HR, U.K.

Received October 1, 2007; Revised Manuscript Received November 21, 2007

ABSTRACT: The Thaumetopoeid silkmoth *Anaphe reticulata* is highly abundant in equatorial and southern Africa and a potential commercial source of silk. In this paper, we report detailed structural characteristics of the silk fibroins. Comparison of the ^{13}C solution NMR spectra of *Anaphe* silk fibroin and several model peptides with Ala and Gly residues indicates that $(AAG)n_1$ and $(AG)n_2$ are the main sequences. In addition this analysis also indicates that the sequence contains $(A)_m$ (where $m > 2$) such as $(AAAG)n_3$, $(AAGAG)n_4$, and $(AAAGAG)n_5$. GG sequences were absent at a level that could be detected by our NMR method. The ^{13}C CP/MAS NMR study shows that the fiber structure is heterogeneous, but predominantly β-sheet structure and the length of $(AG)n_2$ is too short to form the Silk I structure detected in *Bombyx mori* silk fibroin. X-ray diffraction analyses gave information on the higher order structure and hydrogen-bonding character of *Anaphe* silk fiber.

1. Introduction

Silks are fibrous proteins with properties that have intrigued scientists ranging from structural engineers to polymer chemists and biomedical researchers.[1] Many kinds of silks with a wide range of structures and properties are produced by different species of silkworms and spiders. Silk spun naturally from the commercial Mulberry Silkworm *Bombyx mori* has good tensile properties including high strength and Young's modulus.[2] In the past decade *B. mori* silk has come to the forefront as a biomaterial with high environmental stability, good biocompatibility and workability.[2,3] It is readily formed into fibers, films, gels and sponges and has considerable potential for biomedical applications. Silk fibroins from Wild Silkworms such as *Samia cynthia ricini*[4−11] and *Antheraea pernyi*[12−15] have also been studied and compared with that of *B. mori*.

Large numbers of the final instar larvae of *Anaphe reticulata* form into a procession and spin cocoons communally first forming a very large common silk shell in which individuals then form their own cocoons (Figure 1). The common silk shell is divided into three parts: membranous outer layer; soft middle layer; and hard inner layer. The individual cocoons formed within the common shell are a little smaller than typical cocoons of *B. mori*. It is unfortunate that there are no reports giving detailed information on the main amino acid sequences in *Anaphe* silk fibroin as this would be very useful for assessing the potential use of this fibroin as a biomaterial.

In this paper, we seek to rectify this problem, reporting a detailed structural analysis of the silk fibroin from *A. reticulata* using the novel approach described below. The amino acid composition reported previously[16−18] indicates that Ala and Gly strongly predominate. We therefore synthesized several polypeptides containing Ala and Gly but with different primary structures [$(AG)_{15}$; $(AAG)_{10}$; $(AAGAG)_6$; $(AAAG)_7$; $(AAAGAG)_5$ and $(AGG)_{10}$[20−22]] for determination of the main sequences in the *Anaphe* fibroin by comparing ^{13}C solution NMR spectra in 60% LiSCN solution.[23,24] In addition, ^{13}C CP/MAS NMR and wide-angle X-ray diffraction (WAXD) were used to study the higher order structure of *Anaphe* silk fibroin in the solid state.[25]

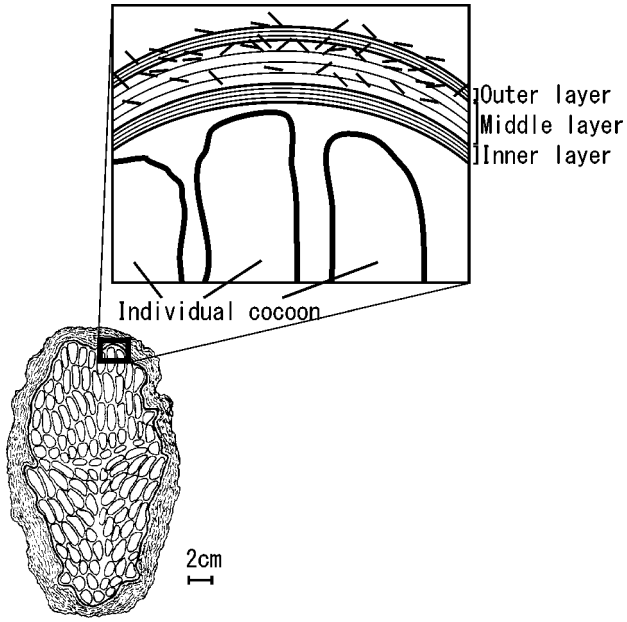

Figure 1. Diagramatic cross section of an *Anaphe* cocoon. The silk fibers were obtained from the outer, middle, and inner layers of the common shell and from an individual cocoon within the common shell.

2. Materials and Methods

2.1. Silk Fibroin from *Anaphe* Cocoons. The outer, middle and inner layers of the common silk shell and individual cocoons of *A.*

* To whom correspondence should be addressed. E-mail: asakura@cc.tuat.ac.jp.
[†] Department of Biotechnology, Tokyo University of Agriculture and Technology.
[‡] Department of Applied Chemistry, National Defense Academy.
[§] Laboratory of Entomology, Tokyo University of Agriculture.
[||] Material Analysis Research Laboratories, Teijin Ltd.
[#] Oxford Biomaterials Ltd.

10.1021/ma7021875 CCC: $40.75
Published on Web 01/01/2008

reticulata were separated by peeling (Figure 1). The raw silks were subjected to two degumming cycles in boiling water containing sodium carbonate (0.3% w/v) for a total of 4 h.[19] The silks were then bleached by immersion in an aqueous solution containing 2% Marseilles soap, 3% hydrogen peroxide, and 2% sodium silicate at 60 °C. The resulting silks were dried after rinsing with boiling water containing sodium carbonate (0.3% w/v) and with boiling deionized water. *S. c. ricini* cocoons were degummed and bleached in the same way as *Anaphe* cocoons. *B. mori* cocoons were degummed in boiling water containing sodium bicarbonate (0.05% w/v). The resulting material was rinsed thoroughly with distilled water and then air-dried at room temperature. The amino acid composition of the outer layer of *Anaphe* cocoon was determined with a Shimadzu Shim-pack Amino-Na column after hydrolysis with 6 N HCl at 110 °C for 22 h in a sealed tube.

2.2. Synthesis of Model Polypeptides with Different Primary Structures of Ala and Gly Residues. Using solid-phase Fmoc-chemistry, we synthesized the following polypeptides: $(AG)_{15}$, $(AAG)_{10}$, $(AAGAG)_6$, $(AAAG)_7$, $(AAAGAG)_5$, and $(AGG)_{10}$. A fully automated Pioneer peptide synthesis system (Applied Biosystem Ltd.) was used throughout. After synthesis, the polypeptides were dissolved in aqueous 60% w/v LiSCN solution and dialyzed against distilled water for 4 days using cellulose tubes (MWCO; 1000). Thereafter, the contents of the dialysis tubes were freeze-dried.

2.3. ^{13}C Solution NMR Measurements. Here, 40 mg of each silk fibroin from *Anaphe*, *B. mori*, and *S. c. ricini*, and 10 mg of each polypeptide, $(AG)_{15}$, $(AAG)_{10}$, $(AAGAG)_6$, $(AAAG)_7$, $(AAAGAG)_5$, and $(AGG)_{10}$, were separately dissolved in 960 μL of 60% w/w aqueous LiSCN solution after which 40 μL of D_2O was added. Similarly, solutions containing a 2:3 molar ratio of the mixture of $(AG)_{15}$ and $(AAG)_{10}$, and a 7:15 molar ratio of the mixture of $(AG)_{15}$ and $(AAAG)_7$ were prepared in the same solvent. All ^{13}C solution NMR spectra were measured on a Bruker DMX 500 MHz spectrometer at 125.78 MHz. A recycle delay of 2 s and 90° pulse width of 5.8 μs was employed. To avoid the large peak derived from the thiocyanate carbon from the LiSCN used to dissolve the silk, the ^{13}C NMR spectra were measured by dividing into two regions: one centered on 22.5 kHz with a spectral width of 8.5 kHz for carbonyl regions, and the other centered on 7.5 kHz with a width of 16 kHz for hydrocarbon regions. ^{13}C chemical shifts were calibrated indirectly through the thiocyanate carbon peak observed at 135.3 ppm and converted to the value relative to 3-trimethylsilylpropionate-d_4 (TSP-d_4) at 0 ppm.

2.4. ^{13}C CP/MAS NMR Measurements. ^{13}C CP/MAS NMR spectra of the three kinds of silk fibroins, and five synthetic polypeptides were recorded on a Chemagnetics Infinity or Varian 400 MHz spectrometer with an operating frequency of 100.0 MHz for ^{13}C at a sample spinning rate of 10 kHz. The number of acquisitions was 8000, and the recycle delay 5 s. For ^{1}H decoupling, 50 kHz of radio frequency field strength was used during the acquisition period of 12.8 ms. A 90° pulse width of 5 μs with 1 ms CP contact time was employed. ^{13}C chemical shifts were calibrated indirectly through the methine peak of adamantane observed at 28.8 ppm relative to TMS at 0 ppm.[20]

2.5. Wide-Angle X-Ray Diffraction Measurements. The X-ray diffraction patterns of silk fibers from the three species were obtained using a Rigaku RINT-TTR3 X-ray diffractometer with Cu Kα radiation (λ = 1.542 nm) from a rotating anode source monochromatized by a multilayer mirror. The voltage and current of the X-ray source were 50 kV and 300 mA, respectively.

3. Results and Discussion

3.1. Amino Acid Composition of *Anaphe* Silk Fibroin. Table 1 compares the amino acid composition of *Anaphe* silk fibroin with published values for *B. mori* and *S. c. ricini* silk fibroins.[26] The silk fibroin from *Anaphe* (outer layer) is exceptional in three respects: The Ala content of 59.1 mol % is very high for a silk; the sum of Ala and Gly content accounts for more than 90 mol %; and the molar ratio of Ala to Gly is about 2:1.

Table 1. Amino Acid Compositions of Silk Fibroins[a] from *Anaphe*, *B. mori*, and *S. c. ricini*[18]

	composition (in mol %)		
amino acid	*Anaphe*	*B. mori*	*S. c. ricini*
Ala	59.1	30.0	48.4
Gly	32.3	42.9	33.2
Ser	2.4	12.2	5.5
Asp	1.7	1.9	2.7
Glu	0.9	1.4	0.7
Tyr	0.8	4.8	4.5
Pro	0.6	0.5	0.4
Leu	0.6	0.6	0.3
His	0.5	0.2	1.0
Val	0.4	2.5	0.4
Thr	0.3	0.9	0.5
Ile	0.2	0.6	0.4
Phe	0.1	0.7	0.2
Lys	0.1	0.4	0.2
Met	0.02	0.1	0.01
Arg	-	0.5	1.7
Cys	-	0.03	0.01
Trp	-	-	0.3

[a] Note that during acid hydrolysis asparagine and glutamine are converted to asparatic acid and glutamic acid respectively while tryptophan is at least partially destroyed.

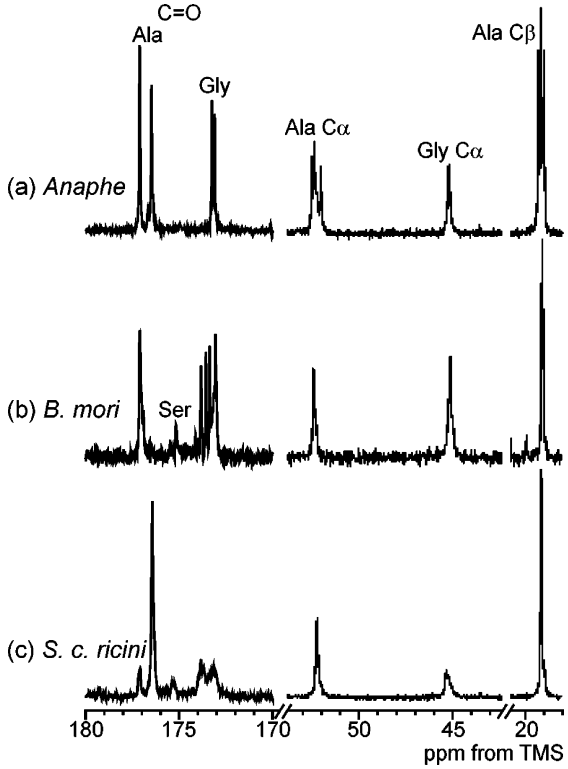

Figure 2. ^{13}C solution NMR spectra of silk fibroins from (a) *Anaphe*, (b) *B. mori*, and (c) *S. c. ricini* in aqueous 60% LiSCN solution. The assignments are shown.

3.2. Primary Structure Determined from ^{13}C Solution NMR. The ^{13}C solution NMR spectra of the silk fibroins prepared from the outer shell of the *Anaphe* cocoon is shown in Figure 2a. The spectra for the inner layer and individual cocoons were closely similar (Data not shown), indicating a similar primary structure in each layer. As is expected from the amino acid composition of *Anaphe* fibroin, only Ala and Gly peaks could be observed in this spectral region. Each carbon in Figure 2a gave rise to roughly two or three split peaks except for Gly Cα peak. The conformation of *Anaphe* silk fibroin is

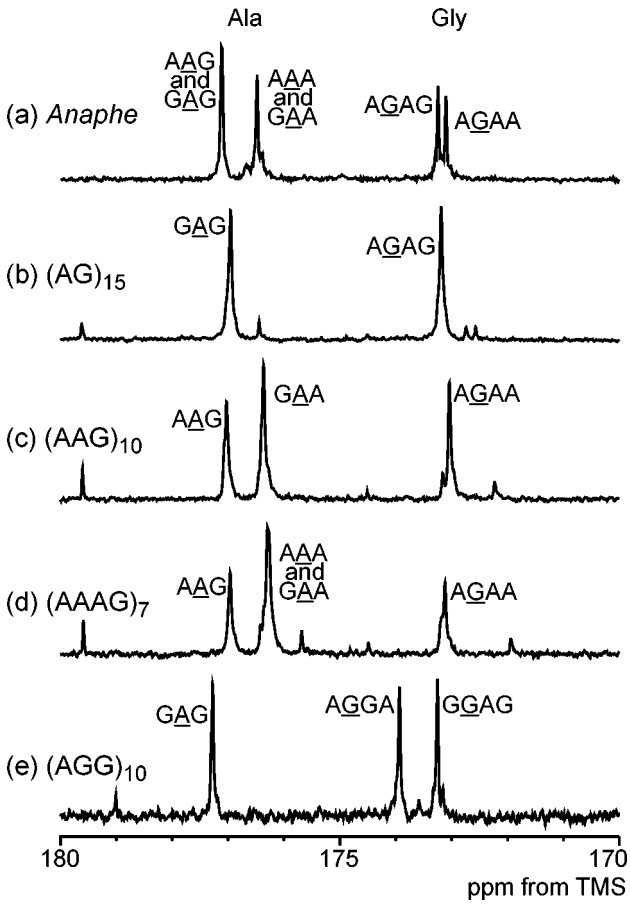

Figure 3. Carbonyl carbon regions in the ^{13}C solution NMR spectra of (a) silk fibroin from *Anaphe*, (b) $(AG)_{15}$, (c) $(AAG)_{10}$, (d) $(AAAG)_7$, and (e) $(AGG)_{10}$ in aqueous 60% LiSCN solution. The details of the assignments are described in the text.

considered to be random coil in 60% LiSCN aqueous solution as in other lepidopteran silks.[23,24]

Dealing first with Gly carbonyl region, (172−174 ppm) of the ^{13}C solution NMR spectrum of *Anaphe* fibroin, it is striking that this region showed only two sharp peaks. We have reported that the Gly carbonyl peaks are very sensitive to the primary structure in silk fibroins.[23,24,28] Thus, the many peak splittings with chemical shift distributions observed for *B. mori* and *S. c. ricini* silk fibroins in random coil state indicate a complex primary structure. The peak splitting in the Gly carbonyl carbon region in *B. mori* silk fibroin clearly originates from the repetitive pentapeptide in the primary structure (Figure 2b).[24] Similarly the Gly carbonyl splitting in *S. c. ricini* silk fibroin also originates from the repetitive tripeptide structure (Figure 2c).[27,28] In contrast, the appearance of only two sharp Gly peaks in *Anaphe* silk fibroin indicates a very simple repetitive sequences containing Gly residues in the silk fibroin. Because of the very small chemical shift difference between the two Gly peaks, we propose that both sequences giving rise to these peaks are G−X, where X is the same residue. The carbonyl carbon regions in the ^{13}C solution NMR spectrum of *Anaphe* silk fibroin are expanded in Figure 3a. We used ^{13}C solution NMR spectra of the four polypeptides, $(AG)_{15}$, $(AAG)_{10}$, $(AAAG)_7$, and $(AGG)_{10}$ we synthesized in order to assign the observed peaks in *Anaphe* silk fibroin in detail. The ^{13}C NMR spectra of the carbonyl carbon region are shown in Figure 3, parts b, c, d, and e, respectively. There are several small peaks which can be mostly assigned to the carbonyl carbons at the terminal residues and the residues adjacent to the terminal residues. For example, the small peaks observed at the lowest field of all

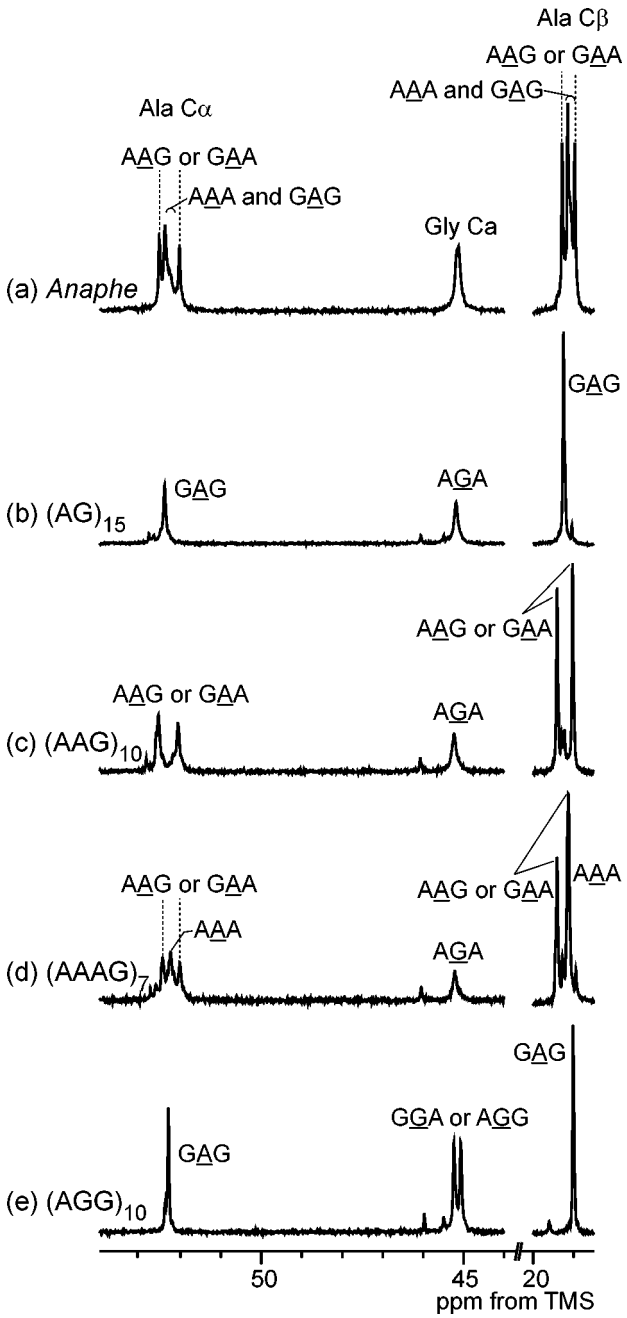

Figure 4. Cα and Cβ carbon regions in the ^{13}C solution NMR spectra of (a) silk fibroins from *Anaphe*, (b) $(AG)_{15}$, (c) $(AAG)_{10}$, (d) $(AAAG)_7$, and (e) $(AGG)_{10}$ in 60% LiSCN aqueous solutions. The details of the assignments are described in the text.

polypeptide spectra are assigned to the carboxylic acid carbons of the C-terminal residues (the last Gly residues) of the peptides. In this paper, we focus on the splitting of main peaks. These peaks are influenced by the adjoining residues on their carbonyl side rather than amino side. Thus, the absence of the peak at 173.9 ppm assigned to the underlined Gly carbonyl carbon in $(AGG)_{10,}$ the one immediately adjacent to the next glycine, strongly indicates that there is no detectable quantity of Gly−Gly in *Anaphe* silk fibroin. The small peak at around 175 ppm in the *Anaphe* spectrum can be assigned to Ser carbonyl peak from the chemical shift.[23,24] This peak is considerably smaller that the equivalent peaks in *B. mori* and *S. c. ricini* and this correlates with the relatively small content of Ser residues in *Anaphe*, only 2.4 mol % compared with 12.2 mol % and 5.5 mol % respectively in the other two silks as determined by amino acid analysis (Table 1).

Table 2. ^{13}C Solution NMR Chemical Shifts of Ala and Gly Residues in Silk Fibroins from *Anaphe, B. mori, and S. c. ricini* Dissolved in Aqueous 60% LiSCN[a]

	Ala						Gly		
sample	C=O		Cα		Cβ		C=O		Cα
Anaphe	176.5	(GAA and AAA)	52.0	(AAG or GAA)	19.0	(AAG or GAA)	173.1	(AGAA)	45.1
	177.3	(GAG and AAG)	52.4	(GAG and AAA)	19.2	(GAG and AAA)	173.2	(AGAG)	
			52.5	(AAG or GAA)	19.3	(AAG or GAA)			
B. mori	177.1		52.5		19.1		173.1		45.1
S. c. ricini	176.4		52.2		19.1				45.3
$(AG)_{15}$	177.0	(GAG)	52.4	(GAG)	19.3	(AGAGA)	173.2	(AGAG)	45.2
$(AAG)_{10}$	176.4	(GAA)	52.1	(AAG or GAA)	19.0	(AAG or GAA)	173.0	(AGAA)	45.2
	177.0	(AAG)	52.5	(AAG or GAA)	19.4	(AAG or GAA)			
$(AAAG)_7$	176.3	(GAA and AAA)	52.0	(AAG or GAA)	19.1	(AAG or GAA, and AAA)	173.1	(AGAA)	45.2
	177.0	(AAG)	52.3	(AAA)	19.4	(AAG or GAA)			
			52.4	(AAG or GAA)					
$(AAGAG)_6$	176.3		52.0	(AAG or GAA)	19.0	(AAG or GAA)	173.0	(AGAA)	45.2
	177.0	(AAG and GAG)	52.4	(GAG)	19.2	(GAG)	173.2	(AGAG)	
			52.5	(AAG or GAA)	19.4	(AAG or GAA)			
$(AAAGAG)_5$	176.3	(AAA and GAA)	52.0	(AAG or GAAA)	19.1	(AAG or GAA)	173.1 (173.12)	(AGAA)	45.2
	177.0	(AAG and GAG)	52.3	(AAAA)	19.2 (19.18)	(AAA)	173.1 (173.15)	(AGAG)	
			52.4	(AAG or GAA and GAG)	19.2 (19.21)	(GAG)			
					19.4	(AAG or GAA)			
$(AGG)_{10}$	177.3	(GAG)	52.3	(GAG)	19.0	(GGAGG)	173.3	(GGAG)	45.1
							173.9	(AGGA)	45.2

[a] The chemical shifts of several polypeptides with different primary structures which consist of Ala and Gly residues, $(AG)_{15}$, $(AAG)_{10}$, $(AAAG)_7$, $(AAGAG)_6$, $(AAAGAG)_5$, and $(AGG)_{10}$ in aqueous 60% LiSCN are also listed. The detailed assignments for the sequences are described in the text.

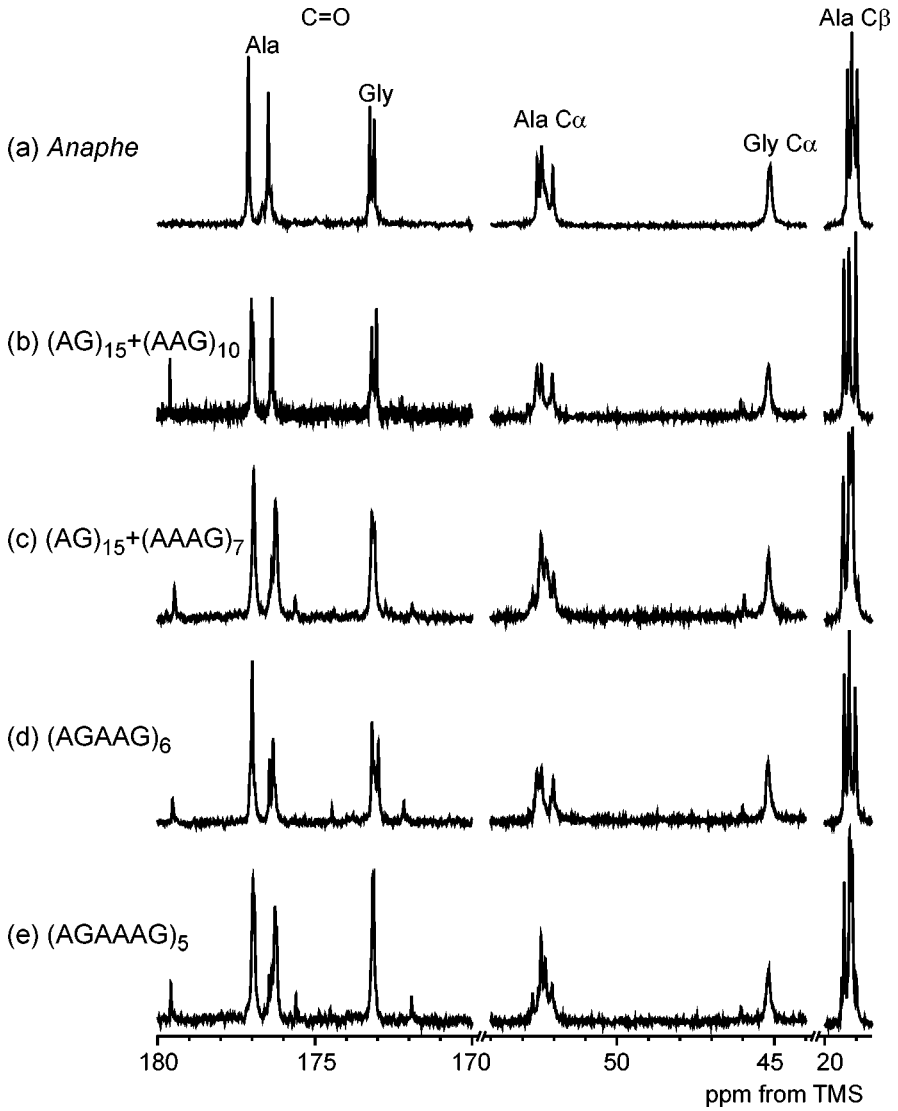

Figure 5. ^{13}C solution NMR spectra of (a) *Anaphe* silk fibroin, (b) 2:3 mixture of $(AG)_{15}$ and $(AAG)_{10}$, and (c) 7:15 mixture of $(AG)_{15}$ and $(AAAG)_7$ together with (d) $(AAGAG)_6$ and (e) $(AAAGAG)_5$ in aqueous 60% LiSCN solution.

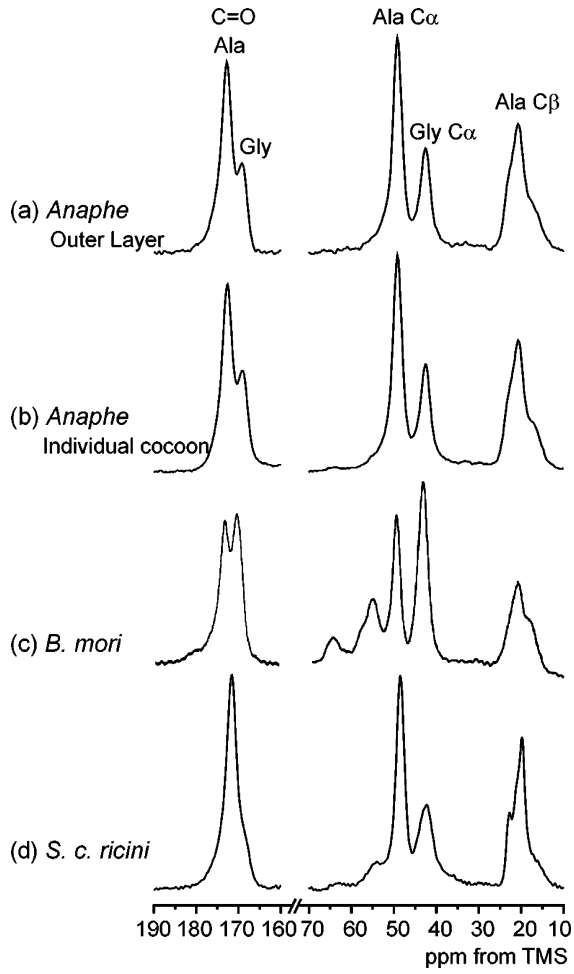

Figure 6. The ^{13}C CP/MAS NMR spectra of silk fibroin fiber from (a) *Anaphe* outer layer, (b) *Anaphe* individual cocoons, (c) *B. mori*, and (d) *S. c. ricini*.

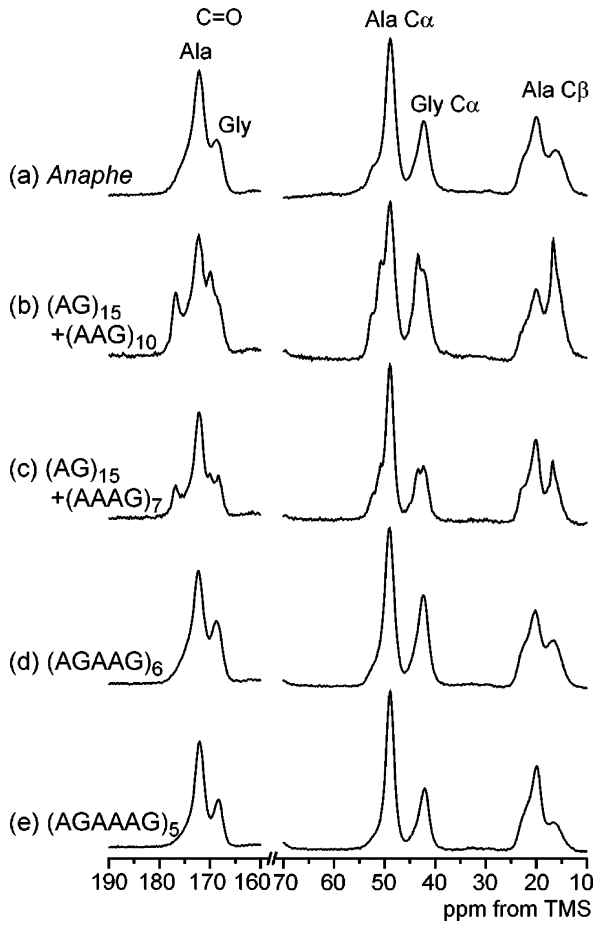

Figure 7. ^{13}C CP/MAS NMR spectra of (a) *Anaphe* silk fibroin, (b) 2:3 mixture of $(AG)_{15}$ and $(AAG)_{10}$, (c) 7:15 mixture of $(AG)_{15}$ and $(AAAG)_7$ together with (d) $(AAGAG)_6$ and (e) $(AAAGAG)_5$.

The separation of the two peaks in the Gly carbonyl region in *Anaphe* was very small, 0.1 ppm. However, it is possible to assign these peaks by reference to the peaks of the spectra of $(AG)_{15}$, $(AAG)_{10}$, $(AAAG)_7$, and $(AGG)_{10}$. The chemical shift of the lower field peak in *Anaphe* silk fibroin was 173.2 ppm which is in agreement with that of $(AG)_{15}$. In addition, the chemical shift of the higher field peak at 173.1 ppm is in agreement with that of $(AAAG)_7$ and close to the chemical shift, 173.0 ppm of $(AAG)_{10}$. Thus, a comparison of Figure 3, parts a−d, shows that the lower field peak can be assigned to the underlined Gly residue of the sequence, A$\underline{G}$AG. The higher field peak is assigned to the underlined Gly residue of the sequence, A$\underline{G}$AA. The chemical shifts and the assignments of the carbonyl carbon peaks are summarized in Table 2.

Turning now to the Ala carbonyl carbon region, and dealing first with the lower field peak of *Anaphe* silk fibroin, this had a chemical shift of 177.3 ppm close to the single Ala carbonyl chemical shift, 177.0 ppm of $(AG)_{15}$, and the lower field peaks of $(AAG)_{10}$ and $(AAAG)_7$; both 177.0 ppm (Figure 3 and Table 2). A comparison of parts b−d of Figure 3 indicates that the lower field peaks in the region of 177.0−177.3 can be assigned to A$\underline{A}$G and G$\underline{A}$G. Thus, we assign the 177.3 ppm peak in *Anaphe* fibroin to the underlined central Ala residue of the sequences, A$\underline{A}$G and G$\underline{A}$G. The higher field peak in the Ala carbonyl carbon region of *Anaphe* silk fibroin had a chemical shift 176.5 ppm, close to the chemical shifts of the higher field peaks of $(AAG)_{10}$ and $(AAAG)_7$, 176.4 and 176.3 ppm, respectively (Figure 3a−d and Table 2). Thus, the higher field peak of *Anaphe* silk fibroin can be assigned to the central Ala residue of the sequences A$\underline{A}$A and G$\underline{A}$A.

Such a sequence analysis was also applied to Ala Cα and Cβ peaks as shown in Figure 4. Both Ala Cα and Cβ peaks of *Anaphe* silk fibroin split into roughly three peaks. The comparison of the spectra in Figure 4b and 4c indicates that the Cβ peak of Ala residue in $(AG)_{15}$ appears between the two peaks of $(AAG)_{10}$. Thus, the center peak of the underlined Ala Cβ carbon is assigned to G$\underline{A}$G. The higher field and lower field peaks are assigned respectively to the underlined Ala residues of either A$\underline{A}$G or G$\underline{A}$A. This assignment is also applied to Ala Cα peak. The chemical shifts and the assignments of Ala Cα and Cβ peaks are summarized in Table 2. By reference of the Ala Cβ peak of $(AAAG)_7$ as shown in Figure 4d the shoulder of the central Ala Cβ peak may indicate the presence of small amounts of the sequence $(A)_m$ (where $m > 2$). The Gly Cα peak of *Anaphe* silk fibroin is a single peak which is similar to the Gly Cα single peak observed for $(AG)_{15}$, $(AAG)_{10}$ and $(AAAG)_7$. The Gly Cα peak of $(AGG)_{10}$ clearly split into two sharp peak, which can be assigned to the sequence either A$\underline{G}$G or G$\underline{G}$A. Thus, it is also concluded that the sequence G−G seen in *B. mori* and *S. c. ricini* silk fibroin is absent in *Anaphe* at concentrations that can be detected by NMR. This is in agreement with the evidence from our analysis of the carbonyl carbon described above. The Gly Cα peak in the *Anaphe* spectrum is slightly broadened, indicating a chemical shift distribution due to the presence of either Gly or Ala residues adjacent to the sequence A$\underline{G}$A.

Parts a−e of Figure 5 show ^{13}C solution NMR spectra of *Anaphe* silk fibroin (a), a 2:3 mixture of $(AG)_{15}$ and $(AAG)_{10}$

Table 3. Distances and Half-Widths for Each Reflection in X-ray Diffraction Patterns of Silk Fibroin Fibers from *Anaphe* Outer Layer, *Anaphe* Individual Cocoons, and *B. mori* and *S. c. ricini* Fibers

sample	2θ (deg)	dist (Å)	half-width (deg)	cryst dimens (Å)	crystallinity (%)
Anaphe	17.7[a]	5.01	1.50	29.8	33.0
outer layer	20.7[b]	4.29	0.96	46.7	
Anaphe	17.7[a]	5.01	1.42	31.4	35.1
individual cocoon	20.7[b]	4.29	0.96	46.7	
B. mori	19.0[a]	4.67	3.22	13.9	37.1
	20.8[b]	4.27	1.53	29.3	
S. c. ricini	16.9[a]	5.25	1.18	37.8	25.2
	20.5[b]	4.32	0.99	45.3	

[a] Hydrobonding distance within β-sheet plane. [b] Inter-sheet distance.

(b), a 7:15 mixture of $(AG)_{15}$ and $(AAAG)_7$ (c), and $(AAGAG)_6$ (d) and $(AAAGAG)_5$ (e). Figure 5b is markedly similar to the spectrum of *Anaphe* silk fibroin (Figure 5a) indicating that these are the main sequences of the silk fibroin. However, there are shoulders at the higher field of both the central Ala Cα and central Ala Cβ peaks in Figure 5a, which could not be observed in the spectrum shown in Figure 5b. This feature in Figure 5a was not found in the spectrum of $(AAGAG)_6$ but is present in the spectra of $(AAAG)_7$ and $(AAAGAG)_5$. Thus, the shoulders at the higher field in both the central Ala Cα and Cβ peaks suggest the presence of small amounts of the $(AAAG)_7$ and $(AAAGAG)_5$ in *Anaphe* fibroin. Although it is difficult to evaluate the amount of these sequences exactly, this evidence does confirm the presence of $(A)_m$ (where $m > 2$) in *Anaphe* fibroin.

3.3. Higher-Order Structure in the Solid-State Determined from ^{13}C CP/MAS NMR. As shown in parts a and b of Figure 6, the ^{13}C CP/MAS NMR spectra of the silk fibroin fibers in the outer layer and individual cocoon of *Anaphe* are quite similar. Thus, both the primary and higher order structures are similar in both parts of this material. The ^{13}C CP/MAS spectra of silk fibroin fibers from *B. mori* and *S. c. ricini* are shown in Figures 6c and 6d respectively. In our previous papers[29,30] we have shown that Ala Cβ region in the ^{13}C CP/MAS NMR spectra of several silk fibroins contains much information on intra- and intermolecular chain arrangement. The spectral pattern of silk fibroin fiber from *Anaphe* is fairly similar to those of silk fibers from *B. mori* and *S. c. ricini* both of which have a β-sheet structure, indicating that all three silks have a similar predominantly β-sheet structure. However, when we compare the spectra in detail, the CP/MAS spectrum of silk fibroin fibers from *Anaphe* is closer to that of *S. c. ricini* than to *B. mori*. This similarity correlates with the Ala and Gly contents in the three silks (Table 1); the Ala content is higher than Gly in both *Anaphe* and *S. c. ricini* but this is not the case in *B. mori*.

Parts a−e of Figures 7a-7e show ^{13}C CP/MAS NMR spectra of *Anaphe* silk fibroin (a), a 2:3 mixture of $(AG)_{15}$ and $(AAG)_{10}$ (b), and a 7:15 mixture of $(AG)_{15}$ and $(AAAG)_7$ (c) together with $(AAGAG)_6$ (d) and $(AAAGAG)_5$ (e). The mixtures containing $(AG)_{15}$ clearly showed sharp peaks at 16.7 ppm in Ala Cβ region and 176.8 ppm in Ala C═O region indicating Silk I structure. Thus, when $(AG)_{15}$ peptides containing a long Ala-Gly repeat was present in the mixture, the higher order structure, Silk I was observed. Our previous paper[31] showed that ^{13}C CP/MAS NMR spectrum of the peptides longer than $(AG)_9$ gave sharp Ala Cβ peak with line width of about 110 Hz which was assigned to Silk I structure. In contrast, the spectrum of the shorter peptide $(AG)_6$ gave a broadened single peak with line width of 192 Hz for the Ala Cβ carbon[31]

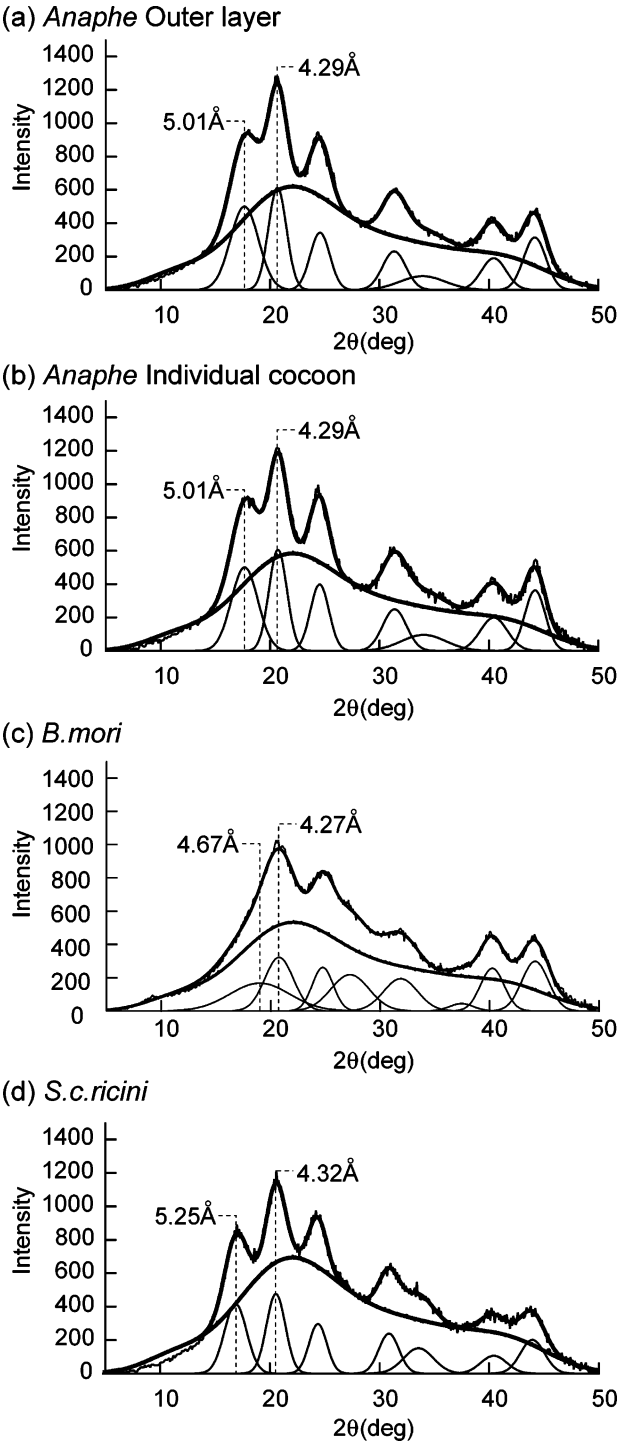

Figure 8. Powder Wide-angle X-ray diffraction spectra of silk fibroin fibers from (a) *Anaphe* outer layer, (b) *Anaphe* individual cocoons, (c) *B. mori*, and (d) *S. c. ricini*. The amorphous patterns were subtracted from the original X-ray diffraction patterns. Upper trace, spectrum; middle trace, amorphous background; lower trace, peaks after Gaussian deconvolution.

indicating the presence of distorted β-turn or random coil rather than Silk I in $(AG)_6$. However, the line width, 192 Hz, was sharper than the broad shoulder peak at 16.7 ppm of the Ala Cβ carbon in the *Anaphe* spectrum in the present study. Thus, the length, n_2 of $(AG)_{n2}$ in *Anaphe* silk fibroin is likely to be less than 6. As is expected, there are no peaks corresponding to Silk I structure in the spectra of $(AAGAG)_6$ and $(AAAGAG)_5$ (see Figure 7, parts d and e), which are similar to the spectrum of *Anaphe* silk fibroin in the solid state.

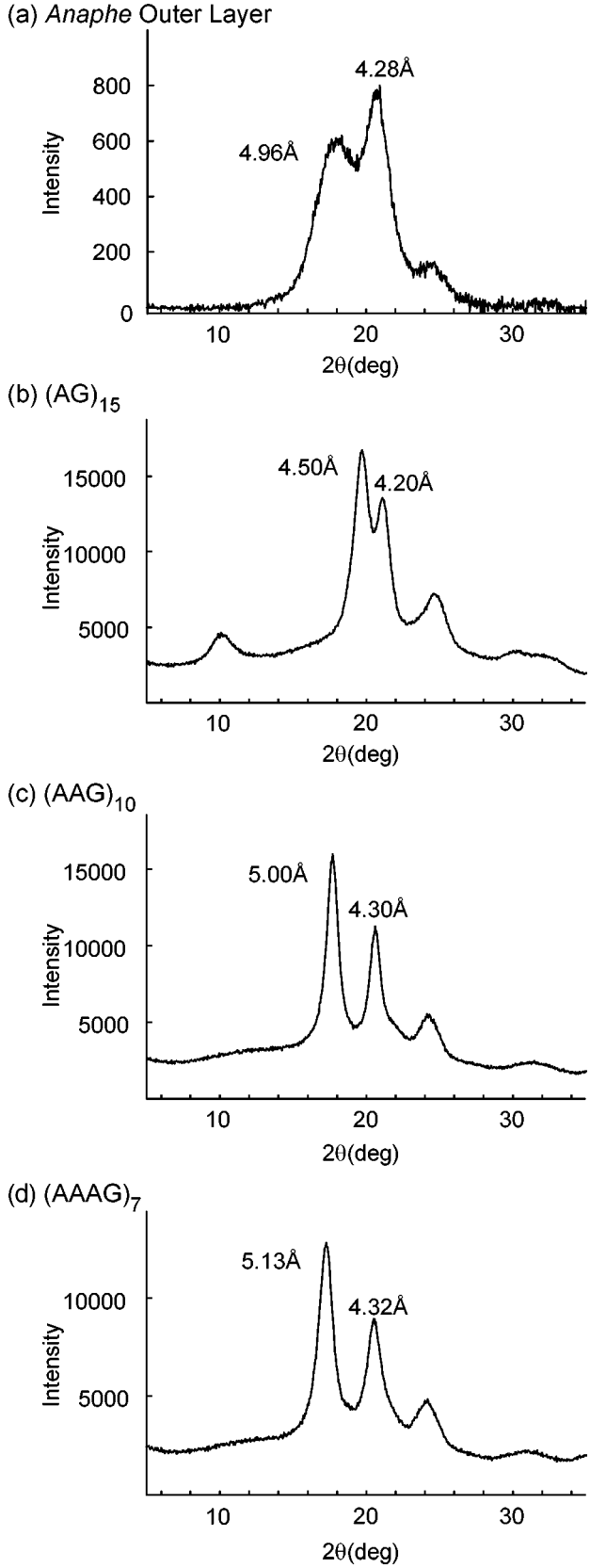

Figure 9. Powder Wide-angle X-ray diffraction spectra of (a) *Anaphe* outer layer after subtraction of the amorphous pattern from Figure 8a, (b) $(AG)_{15}$, (c) $(AAG)_{10}$, and (d) $(AAAG)_7$.

3.4. Differences in Higher-Order Structure in the Solid-State Revealed by WAXD. The results of WAXD of the four different silk fibroin samples are summarized in Table 3 and Figure 8. The patterns are almost the same for the outer layer (Figure 8a) and individual cocoons (Figure 8b) in *Anaphe* as expected from the ^{13}C CP/MAS NMR data (see above). In addition, the patterns are also similar in *Anaphe* and *S. c. ricini* silk fibroin fibers (Figure 8d), but are quite different in *B. mori* fibers (Figure 8c). The hydrogen-bonding distances within β-sheet planes were almost the same for *Anaphe*, *S. c. ricini* and *B. mori* silk fibers; That is approximately 4.3 Å, the 2θs being 20.7, 20.5, and 20.8 deg, respectively. However the inter-sheet distances were considerably different; 4.67 Å ($2\theta = 19.0°$) for *B. mori*, 5.01 Å ($2\theta = 17.7°$) for *Anaphe* and 5.25 Å ($2\theta = 16.9°$) for *S. c. ricini*. The value for *Anaphe* lies between those of the two other silks, but is closer to that of *S. c. ricini.* The half-height-peak width corresponding to the inter β-sheet distances increased in the order *S. c. ricini* < *Anaphe* < *B. mori* indicating an increase in the distribution of inter-sheet distance. The crystal sizes of the silk fibers calculated using the Scherrer formula decreased in the same order. The X-ray diffraction patterns were also observed for $(AG)_{15}$, $(AAG)_{10}$, and $(AAAG)_7$ as shown in Figure 9. The results are in agreement with WAXD inter-sheet distance data for different poly peptides reported by Lotz et al.[32] The WAXD pattern of *Anaphe* silk fibroin fiber can be reproduced qualitatively by the mixtures of the patterns of $(AG)_{15}$ and $(AAG)_{10}$ in agreement with the above analysis of the ^{13}C CP/MAS NMR of *Anaphe* fibroin.

4. Summary

Comparison of the ^{13}C solution NMR spectra of *Anaphe* silk fibroin and several model peptides indicates that the sequence of this silk's heavy chain fibroin predominantly contains a mixture of $(AAG)n_1$ and $(AG)n_2$. In addition the fibroin may also contain $(A)_m$ (where $m > 2$) contained in sequences such as $(AAAG)n_3$, $(AAGAG)n_4$ and $(AAAGAG)n_5$, while GG was not present at concentrations detectable by NMR. Lucas and co-workers[33] proposed the sequence of *Anaphe* silk fibroin as $((G-A)_{p1}-A)_{p2}$, but in contrast to their proposal, our data indicates that $(AAG)n_1$ and $(AG)n_2$ are the main sequences. This is supported by the Ala content of approximately 60% which shows that there must be −AA− sequences in this fibroin. Our results conclude that *Anaphe* silk is intermediate in primary and higher order structure between the Ala-Gly silk of *B. mori* and the polyalanine silks typified by *Antheraea* and *Samia* Wild silks and the MaSp1 silk fibroins of spiders. This difference may be useful in the design of biomedical materials.

Acknowledgment. T.A. acknowledges Dr Hajime Saito for stimulating discussions. T.A. also acknowledges support from a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Culture, and Supports of Japan (18105007). D.P.K acknowledges support from the EU FP6 Silkbone project.

References and Notes

(1) O'Brien, J.; Fahnestock, S.; Termonia, Y.; Gardner, K. *Adv. Mater.* **1998**, *10*, 1185−1195.
(2) Asakura, T.; Kaplan, D. L. *Encyclopedia of Agricultural Science*; Arutzen, C. J., Ed.; Academic Press: New York, 1994; Vol. 4, pp 1−11.
(3) Altman, G. H.; Diaz, F.; Jakuba, C.; Calabro, T.; Horan, R. L.; Chen, J.; Lu, H.; Richmond, J.; Kaplan, D. L. *Biomaterials* **2003**, *24*, 401−416 and references therein.
(4) Asakura, T.; Nakazawa, Y. *Macromol. Biosci.* **2004**, *4*, 175−185.
(5) Asakura, T.; Murakami, T. *Macromolecules* **1985**, *18*, 2614−2619.
(6) Asakura, T.; Kashiba, H.; Yoshimizu, H. *Macromolecules* **1988**, *21*, 644−648.
(7) Kameda, T.; Ohkawa, Y.; Yoshizawa, K.; Nakano, E.; Hiraoki, T.; Ulrich, A. S.; Asakura, T. *Macromolecules* **1999**, *32*, 8491−8495.
(8) van Beek, J. D.; Beaulieu, L.; Schafer, H.; Demura, M.; Asakura, T.; Meier, B. H. *Nature* **2000**, *405*, 1077−1079.
(9) Nakazawa, Y.; Asakura, T. *J. Am. Chem. Soc.* **2003**, *125*, 7230−7237.

(10) Asakura, T.; Okonogi, M.; Nakazawa, Y.; Yamauchi, K. *J. Am. Chem. Soc.* **2006**, *128*, 6231−6238.
(11) Lefevre, T.; Rousseau, M-E.; Pezolet, M. *Biophys. J.* **2007**, *92*, 2885−2895.
(12) Magoshi, J.; Magoshi, Y.; Nakamura, S. *J. Appl. Polym. Sci.* **1977**, *21*, 2405−2407.
(13) Taddei, P.; Arai, T.; Boschi, A.; Monti, P.; Tsukada, M.; Freddi, G. *Biomacromolecules* **2006**, *7*, 259−267.
(14) Nakazawa, Y.; Asakura, T. *Macromolecules* **2002**, *35*, 2393−2400.
(15) Sezutsu, H.; Yukuhiro, K. *J. Mol. Evol.* **2000**, *51*, 329−338.
(16) Lucas, F.; Shaw, T. B.; Smith, S. G. *J. Mol. Biol.* **1960**, *2*, 339−349.
(17) Warwicker, J. O. *J. Mol. Biol.* **1960**, *2*, 350−362.
(18) Komatsu, K.; Yamada, M.; Hashimoto, Y. *J. Sericult. Sci. Jpn.* **1969**, *38*, 219−229.
(19) Akai, H.; Nagashima, T.; Mugenyi, G. *Int. J. Wild Silkmoth Silk* **1999**, *4*, 7−12.
(20) Asakura, T.; Ashida, J.; Yamane, T.; Kameda, T.; Nakazawa, Y.; Ohgo, K.; Komatsu, K. *J. Mol. Biol.* **2001**, *306*, 291−305.
(21) Ashida, J.; Ohgo, K.; Komatsu, K.; Kubota, A.; Asakura, T. *J. Biomol. NMR* **2003**, *25*, 91−103.
(22) Asakura, T.; Ohgo, K.; Komatsu, K.; Kanenari, M.; Okuyama, K. *Macromolecules* **2005**, *38*, 7397−7403.
(23) Asakura, T.; Watanabe, Y.; Uchida, A.; Minagawa, H. *Macromolecules* **1984**, *17*, 1075−1081.
(24) Asakura, T.; Watanabe, Y.; Itoh, T. *Macromolecules* **1984**, *17*, 2421−2426.
(25) Asakura, T.; Yamane, T.; Nakazawa, Y.; Kameda, T.; Ando, K. *Biopolymers* **2001**, *58*, 521−525.
(26) Hojyo, N., Ed. *Zoku Kenshi no Kozo (Structure of Silk Fibers)*; Shinshu University: Ueda, Japan 1980.
(27) Tao, W.; Li, M.; Zhao, C. *Int. J. Biol. Macromol.* **2007**, *40*, 472−478.
(28) Nakazawa, Y.; Asakura, T. *FEBS Lett.* **2002**, *529*, 188−192.
(29) Asakura, T.; Yao, J.; Yamane, T.; Umemura, K.; Ulrich, A. S. *J. Am. Chem. Soc.* **2002**, *124*, 8794−8795.
(30) Yao, J.; Nakazawa, Y.; Asakura, T. *Biomacromolecules* **2004**, *5*, 680−688.
(31) Ohgo, K.; Kurano, T. L.; Kumashiro, K. K.; Asakura, T. *Biomacromolecules* **2004**, *5*, 744−750.
(32) Lotz, B.; Brack, A.; Spach, G. *J. Mol. Biol.* **1974**, *87*, 193−203.
(33) Lucas, F.; Shaw, J. T. B.; Smith, S. G. *J. Mol. Biol.* **1960**, *2*, 339−349.

MA7021875