J. Biochem. **133**, 147–154 (2003)
DOI: 10.1093/jb/mvg014

Synthesis and Structural Characterization of Silk-Like Materials Incorporated with an Elastic Motif

Juming Yao and Tetsuo Asakura*

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

Received October 19, 2002; accepted November 13, 2002

Genetic engineering strategies were applied to synthesize silk-like materials, $[(GVPGV)_2GG(GAGAGS)_3AS]_n$. The primary structure of these materials represents the repetitive crystalline region of *Bombyx mori* silk fibroins incorporated with an elastic motif selected from animal elastin. The oligonucleotides were designed to encode the desired recombinant proteins and then expressed in the *Escherichi coli* system. The expression and purification conditions for the production of the recombinant proteins were optimized. ^{13}C CP/MAS NMR was used for structural characterization in the solid state, where the isotope labeling was performed using a modified M9 medium. The secondary structures of these materials are primarily governed by the designated amino acid sequence, where the *B. mori* silk fibroin block, $(GAGAGS)_3$, tends to form the crystalline region, which is interrupted by the flexible $(GVPGV)_2$ block. The CD data suggested that the structure of these materials was length-dependent in the solution state, *i.e.*, a higher molecule weight leads to a higher ordered structure.

Key words: *Bombyx mori* silk fibroin, circular dichroism, genetic engineering synthesis, nuclear magnetic resonance, silk-like materials.

For centuries, nature has been testing biological polymers as structural and functional materials. Proteins, which are polymers with 20 different possible amino acid monomers, have proved to be extremely adaptable as biomaterials. Silk proteins are usually produced within specialized glands after biosynthesis in epithelial cells, followed by secretion into the lumen of these glands, where the proteins are stored prior to spinning into fibers (*1*). As fibrous proteins, silks differ widely in composition, structure and properties depending on the specific source. The most extensively characterized silk is from the domesticated silkworm, *Bombyx mori*. Silk fibers from *B. mori* fibroin have been used as biomedical suture material for a long time, because of their outstanding mechanical properties, in contrast to the catalytic and molecular recognition functions of globular proteins (*2*, *3*). The relative environmental stability of this protein, in combination with its biocompatibility, unique mechanical properties, and options for genetic control to tailor sequences, provides an important basis for exploiting this protein for biomedical applications (*4*).

The unique properties of *B. mori* silk fibroin fibers should be due to the distribution of crystalline and amorphous domains formed in the process of spinning through protein–protein interactions. The complete primary structure of *B. mori* fibroin has been determined by Mita *et al.* (*5*), and more recently by Zhou *et al*. (*6*). Approximately 70% of the fibroin comprises the amino acid sequence, GAGAGS. A more extended sequence has been reported for the crystalline fraction precipitated after chymotrypsin hydrolysis, *i.e.* $GAGAGSGAAG[SG(AG)_n]_8Y$, where n is usually 2 (*7*, *8*). We recently clarified the heterogeneous structure of this crystalline fraction $(GAGAGS)_n$ after fiber spinning (silk II structure), as well as the heterogeneous structure of the natural fibers by means of ^{13}C solid state NMR (*8*, *9*). Although there is some distorted structure, nearly 70% of the repeated sequences in the crystalline fraction are β-sheet regular structures, which could be responsible for the higher strength of *B. mori* silk fibers.

Considering the high degree of utility of silk, man's desire to capitalize on the potential of silk products is an obvious one. However, the lack of abundant sources has been one of the barriers for further developments. Fortunately, genetic engineering has opened a window for overcoming this limitation, which is being actively explored to construct, clone and express native and synthetic genes encoding recombinant proteins (*10–13*). Moreover, genetic engineering synthesis is a powerful method for varying properties, through appropriate choice of the different units, the number of units in each multimer, the spacing between them, and the number of repeats of the multimer combination assembly. Thus, by varying the number and arrangement of primary monomers, a variety of different physical and chemical properties can be obtained.

In this work, we tried to introduce an elastic motif, GVPGV, into this *B. mori* crystalline fraction to form a new silk-like protein by means of a genetic engineering synthesis method. This elastic motif, GVPGV, found in animal elastin has been produced in a polymeric material that exhibits extraordinary elasticity (*14*). A polypentapeptide elastic material is capable of being stretched to more than 300% of its testing length, with no deforma-

*To whom correspondence should be addressed. Tel/Fax: +81-42-383-7733, E-mail: asakura@cc.tuat.ac.jp

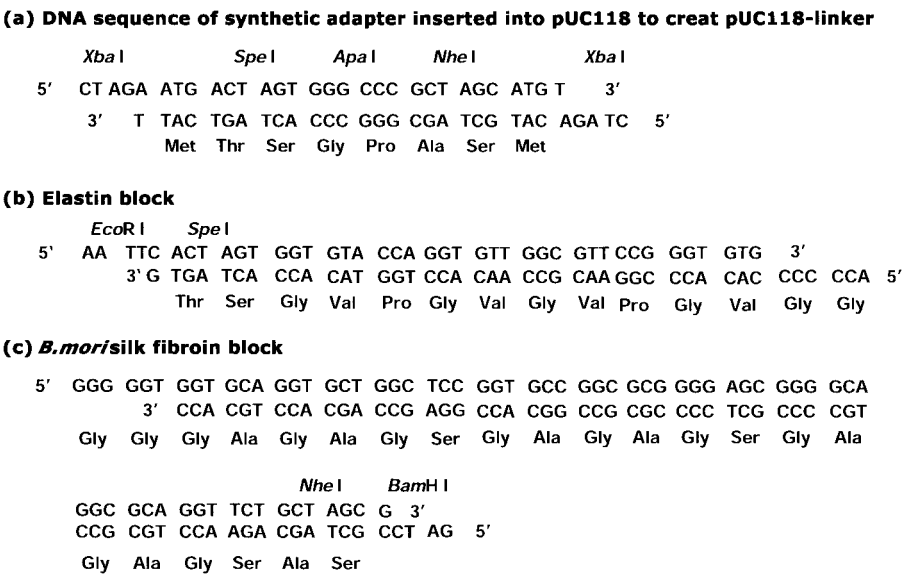

Fig. 1. **The designed oligonucleotide sequences for silk-like proteins (*BcEn*).** (a) DNA sequence of the synthetic adapter inserted into pUC118 to create the pUC118-linker, (b) oligonucleotide sequence of the elastin block, and (c) oligonucleotide sequence of the *B. mori* silk fibroin block used for cloning of the hybrid proteins.

tion. Its capacity to be modified by design has yielded a set of new materials that are able to change their mechanical properties either as the external conditions change or as chemical modifications occur (*14*, *15*). Thus, the amino acid sequence of the new silk-like protein is designated as $[(GVPGV)_2GG(GAGAGS)_3AS]_n$ (defined as *BcEn*), where the elastic block selected from elastin (*En*) interrupts and reduces the crystallinity of a synthetic protein formed by the *B. mori* crystalline fraction (*Bc*). As a result, the insertion of this elastic motif that is unable to form β-sheet could impart flexibility, and change the mechanical properties of the protein polymers. Here, we described a synthetic method for these recombinant silk-like proteins. After optimization of the production conditions, structural characterization of these silk-like materials was carried out by means of ^{13}C CP/MAS NMR in the solid state and CD measurement in the solution state.

EXPERIMENTAL PROCEDURES

Materials—*E. coli* strain DH5α was used for the propagation and construction of plasmids, and *E. coli* strain BL21(DE3)pLysS for the production of proteins. Synthesized DNA fragments with phosphorylation were purchased from Sigma Genosys, Japan. Restriction enzymes and ligase were purchased from Takara Shuzo. Plasmid pUC118 was obtained from Takara Shuzo, and pET30a from Novagen. All other chemicals were of analytical grade.

General Methods—Bacterial growth in rich medium, DNA manipulations, and transformation were performed as described in Sambrook, Frisch, and Maniatis (*16*). DNA sequencing was performed with an ABI PRISM™ 377 Auto Sequencer according to the user's manual. Cell density was measured with a Hitachi U-3200 spectrophotometer in quartz cuvettes with a path length of 1 cm at λ = 600 nm. Batch cultures were performed in an MDL-6C Fermentor (B.E. Marubishi), with a 1.2 liter working volume.

Construction of the pUC118-Linker—The sequence shown in Fig. 1a was obtained as two fragments, which when annealed so as to have *Xba*I compatible ends. Complementary oligonucleotides were mixed in equimolar ratios in TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and heated to 95°C, and then slowly cooled for annealing over 3 h and the formation of duplexes. The duplex DNA was ligated into pUC118, which had been digested with *Xba*I and dephosphorylated with Calf Intestine Alkali Phosphatase (CIAP), with a DNA Ligation Kit Ver. 2 by incubation at 16°C for 3 h. Ten microliters of the ligation mixture was transformed to *E. coli* strain DH5α cells (100 μl), which were grown on LB medium containing ampicillin, β-isopropyl thiogalactoside (IPTG), and chromogenic substrate 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal), respectively, for blue-white screening. The pUC118-linker plasmid was isolated from 1.5 ml 2× YT medium.

Cloning and Amplification of Oligonucleotides—Two oligonucleotide fragments for both *B. mori* crystalline (*Bc*) and elastic motif (*En*) blocks were designed, as shown in Fig. 1, (b) and (c), respectively, and annealed by the same method as that described for construction of the pUC118-linker. The duplex DNAs were ligated into *Bam*HI–*Eco*RI-digested pUC118 with a DNA Ligation Kit Ver. 2, and then used to transform *E. coli* strain DH5α. Cells were grown at 37°C overnight on LB medium containing ampicillin, with blue-white screening. After isolation of the recombinant plasmid from 1.5 ml 2× YT medium, the DNA was digested with *Nhe*I–*Spe*I and the fragments were separated by nondenaturing agarose gel electrophoresis. The DNA monomer was extracted from an agarose slice by low-speed centrifugation.

Polymerization of the DNA Monomer—Purified monomeric DNA was ligated into the *Nhe*I–*Spe*I-digested pUC118-linker with a DNA Ligation Kit Ver. 2. The amplification of pUC118-linker-*BcEn*1 was referred to the amplification of oligonucleotides. The obtained pUC118-linker-*BcEn*1 was digested with both *Nhe*I and *Nhe*I–*Spe*I to obtain linear pUC118-linker-*BcEn*1 and monomer DNA, respectively. The above process was

repeated to obtain the dimer, tetramer, octamer, hexadecamer, *etc*.

Construction of the Bacterial Expression Vector—Recombinant plasmids, pUC118-linker-*BcEn*-multimers, were digested with *Bam*HI and *Hin*dIII to release the multimers. The recovered DNA multimers were ligated into *Bam*HI and *Hin*dIII-digested, dephosphorylated pET30a, and then used to transform *E. coli* strain DH5α. Recombinant plasmids, pET30a-*BcEn*-multimers, were used to transform the expression host, *E. coli* strain BL21(DE3)pLysS.

Protein Expression by Batch Culture and Protein Purification—A culture was grown at 37°C in TB medium containing chloramphenicol (25 μl/ml) and kanamycin (25 μl/ml) in a 2 liter fermentor. Protein expression was induced when the OD_{600} reached 0.6–0.8 by the addition of 0.2 mM IPTG at 30°C. Then, the culture was grown for another few hours prior to harvesting by centrifugation (8,500 rpm, 40 min, 4°C). The cells was collected and stored at –70°C until the next processing.

Frozen cells were thawed, resuspended in 3 ml of lysis buffer [50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10 mM imidazole] per gram of cells (wet weight), and then disrupted by supersonication. After centrifugation (14,500 rpm, 30 min, 4°C), the supernatant was applied to a Ni-NTA agarose column, which was charged with Ni^{2+} to purify the fusion protein by using His-tags. The column with the protein was lysed with the lysis buffer, washed with a wash buffer [50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 20 mM imidazole], and then eluted with an elution buffer [50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 250 mM imidazole]. The solution obtained on elution was dialyzed against distilled water for 3 days and then lyophilized.

Cleavage of the Fusion Protein—Cyanogen bromide cleavage of the fusion proteins was accomplished by the method of Smith (*17*). Approximately 62 mg of fusion protein was dissolved in 10 ml of 99% formic acid and then diluted to 70% formic acid with distilled water. The mixture with 100 mg of CNBr crystals added was stirred at room temperature for 24 h and then transferred to a dialysis membrane with a molecular weight cutoff of 3,500. The solution was dialyzed against distilled water for 3 days, and the cleaved protein (38 mg) was recovered by lyophilization.

Isotopic Labeling of Proteins—To obtain structural information on the $(GAGAGS)_3$ region in the recombinant proteins, ^{13}C isotope labeling was carried out by supplementing a modified M9 medium with [3-^{13}C]Ala (50 mg/liter). The composition of the modified M9 medium is shown in Table 1. The expression conditions were as for the cultivation of cells in TB medium. Protein purification was carried out by nickel-chelate chromatography, as described above.

Peptide Synthesis—To obtain structural information on the mono-blocks, peptide $(VPGVG)_6$ was synthesized by means of solid phase F-moc chemistry with a fully automated Pioneer Peptide Synthesis System (Applied Biosystems) (*9*, *18*). The peptide was assembled on F-moc-Gly-PEG-PS resin and then the coupling of F-moc amino-acids was performed. 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% by volume) for 2 h at room temperature. The crude peptide was precipitated, washed with cold diethyl ether, and then dried under vacuum. The peptide was purified by HPLC and collected by lyophilization. Because of its solubility in water, this sample is considered to take on silk I form without LiBr treatment (*18*). Silk II form treatment was carried out by dissolving the powder in formic acid, followed by drying under vacuum (*9*).

Table 1. **Recipe for 1 liter of modified M9 medium.**

NH_4Cl	1 g
KH_2PO_4	3 g
$Na_2HPO_4{\cdot}7H_2O$	6 g
Glucose	4 g
$MgSO_4$	1 ml 1 M (120 mg)
NaCl	0.5 g
$CaCl_2$	10 mg
VB_1	10 mg

Cp Fraction of B. mori Silk Fibroin—To obtain structural information on the mono-block, $(GAGAGS)_3$, the Cp fraction of *B. mori* silk fibroin [predominantly $(GAGAGS)_n$ (*1*, *6*)] was prepared from a regenerated silk fibroin solution as described elsewhere (*8*, *9*). Chymotrypsin (40 mg), dissolved in a few milliliters of water, was added to an aqueous solution of about 4 g of fibroin buffered with $Na_2HPO_4{\cdot}12H_2O$ and $NaH_2PO_4{\cdot}2H_2O$ at pH 7.8. The solution (200 ml) was incubated at 40°C for 24 h, and the precipitate that formed (Cp fraction) was separated by centrifugation at 10,000 rpm, followed by washing with 0.03 N HCl to stop the enzyme reaction. Then the precipitate was washed several times with distilled water, ethyl alcohol, and ethyl ether. Finally the precipitate was dried under vacuum, yielding 55% of original fibroin. The structure of this Cp fraction is silk II (*8*, *9*). The Cp fraction in the silk I form was prepared by dialyzing the Cp fraction in 60% aqueous LiBr against similar solutions of LiBr diluted progressively with water (*8*, *9*).

NMR Observation—Solid state ^{13}C CP/MAS NMR spectra were acquired with a Chemagnetics CMX-400 spectrometer operating at 100 MHz, with a CP contact time of 1 ms, TPPM decoupling, and magic angle spinning at 5 kHz. A total of 10,000–25,000 scans was collected over a spectral width of 60 kHz, with a recycle delay of 3 s. Chemical shifts are reported relative to TMS as a reference.

Circular Dichroism (CD) Measurements—The solution state CD spectra (185–260 nm) of the proteins were recorded in $HFA{\cdot}xH_2O$ with a JASCO J-805 spectropolarimeter. A cylindrical quartz cell of 0.1 cm in path length was used for spectral measurements. The spectra were recorded at ambient temperature and each spectrum is presented as the average of eight consecutive scans measured with a 1 nm resolution.

RESULTS AND DISCUSSION

Gene Construction—In order to produce the tandem repetitive polypeptides, the DNA sequences shown in Fig. 1, (b) and (c), were designed with the following considerations. The 5′-*Eco*RI site at the elastin-like block and the

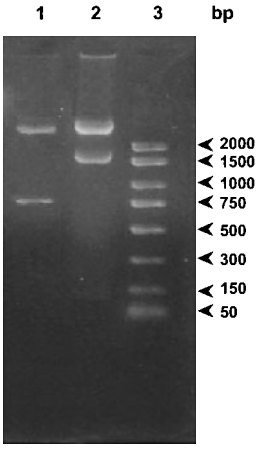

Fig. 2. ***Spe*I and *Nhe*I digestion analysis of multimerized cloning vectors.** Lane 1, pUC118-linker-*BcEn*8; lane 2, pUC118-linker-*BcEn*16; and lane 3, PCR markers.

3′-*Bam*HI site at the *B. mori* silk-like block are positioned at each end for insertion of the DNA into cloning plasmid pUC118 between the *Eco*RI and *Bam*HI sites. The *Spe*I and *Nhe*I restriction sites encoding Thr-Ser and Ala-Ser were inserted after the *Eco*RI and before the *Bam*HI site, respectively, in order to isolate monomer DNA from the recombinant plasmid. Optimal codons for each amino acid are possibly selected using codon usage in *E. coli* as previously reported (*19*). Since the base composition and sequence of the DNA can influence the polypeptide expression level and the stability of the DNA itself, the codons with as high A-T contents as possible were selected. The monomer DNA was inserted into the *Nhe*I and *Spe*I-digested pUC118-linker, flanked by an ATG codon encoding methionine at each site in order to obtain the periodic portion of *BcEn* by cleavage of the expressed recombinant proteins with cyanogen bromide (CNBr). Multimers were formed by insertion of *Nhe*I–*Spe*I fragments into either *Nhe*I- or *Spe*I-digested vectors, since these enzymes generate identical cohesive ends, to build, in a tightly controlled fashion, larger genes of any desired degree of multimerization.

Figure 2 shows the 1.5% agarose electrophoresis results for the *Nhe*I and *Spe*I-digested recombinant cloning vectors, pUC118-linker-*BcEn*8 (lane 1) and pUC118-linker-*BcEn*16 (lane 2), corresponding to the construction of 8 and 16 repeats of the monomer DNA, which comprise 768 and 1,536 bp, respectively, and were verified by the PCR markers (lane 3).

Optimization of the Expression and Purification of BcEn Recombinant Proteins—In this study, purified polymerized DNA fragments were inserted into commercially available expression plasmid pET30a between the *Bam*HI and *Hin*dIII restriction sites. Therefore, the resulting recombinant expression plasmid encodes the *BcEn* DNA sequence of interest, flanked by N- and C-terminal extensions of 53 and 19 amino acids, respectively, as shown in Fig. 3. These terminal regions are derived from the transfer and expression vectors, and can be removed by CNBr cleavage at flanking methionine residues.

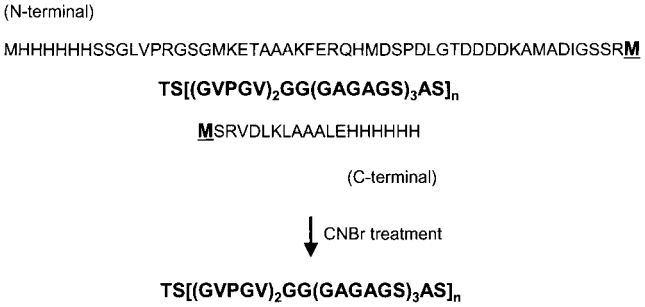

Fig. 3. **Amino acid sequences of recombinant proteins and their cleavage polypeptides.**

The host used for protein expression was *E. coli* strain BL21(DE3)pLysS. In this strain, a gene encoding T7 RNA polymerase is incorporated into the bacterial chromosome under *lacUV5* control, and protein production is induced by the addition of IPTG (*20*). The pLysS plasmid provides low levels of T7 lysozyme, which inhibits T7 RNA polymerase and suppresses the basal level of protein expression.

The protein expression level of *BcEn*8 at 30°C in rich medium (TB) was checked at different induction times. Aliquots of cultures representing approximately equal numbers of cells (as determined from the optical density at 600 nm) were removed periodically after induction of protein synthesis (t = 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, and 4.5 h), and loaded onto a 15% polyacrylamide gel for SDS-PAGE by the method of Laemmli (*21*). Figure 4 shows the Western blotting results after transferring the proteins from the above SDS–polyacrylamide gel, which suggest the expression level increases with the induction time till 4 h, however, further induction is not likely to provide any additional benefit as to the level of expression. Moreover, a single band was observed in each lane, suggesting that there are no other His-tagged proteins due to degradation of the target protein under these expression conditions.

Protein expression of *BcEn*8 by batch culture was performed as described under "EXPERIMENTAL PROCEDURES." Optimization of fermentation can be performed by changing one variable, such as medium, temperature, pH, air-supply (oxygen), concentration of IPTG, *etc.*, while fixing all the others at certain levels. But actually,

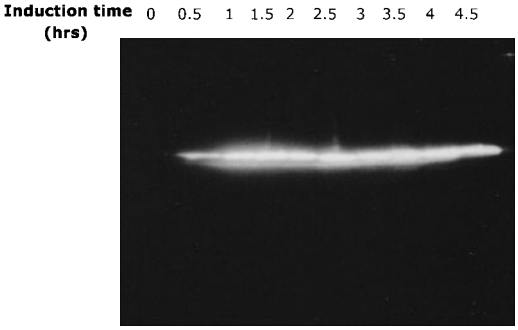

Fig. 4. **Western-blotting analysis of the expression level of *BcEn*8 after different induction times.**

Table 2. **Fermentation conditions for the production of *BcEn*8.**

Temperature before induction	37°C
Temperature after induction	30°C
Induction point	OD_{600} = 0.6–0.8
Concentration of IPTG	0.2 mM
Induction time	4 h
Agitation speed	500 rpm
Air-flow	1 liter/min

it is extremely time consuming and expensive for a large number of variables. In this study, terrific broth (TB) medium was used, which contains K phosphate as a buffer and also as a nutrient source. Lower level expression usually occurs because the protein is toxic or unstable, or because the expression construct is not maintained in the cells during growth. To reduce the effects of protein toxicity on cell growth prior to induction, the level of basal transcription that occurs in the absence of induction should be repressed as much as possible, and the number of generations before induction should be kept to a minimum. Problems related to the loss of plasmids can sometimes be overcome by growing the cells in the presence of high levels of kanamycin (50–100 μg/ml). For instable expression constructs, overnight pre-cultures should be avoided. Colonies from a fresh plate should be inoculated into a small pre-culture and grown for 2–4 h, until the mid-log phase. The pre-culture should then be diluted with prewarmed medium and grown to an OD_{600} of approximately 0.5 before induction. Moreover, the proteins designed in this study contain hydrophobic regions, $(GAGAGS)_3$, which have a toxic effect on host cells due to the association of the proteins with or incorporation into vital membrane systems. This intermolecular association of hydrophobic regions is also believed to play a role in the formation of inclusion bodies (*i.e.,* insoluble proteins), which can be overcome by reducing the growth temperature. In this study, lowering of the growth temperature to 30°C before induction was carried out in order to reduce the toxic effect and to produce soluble proteins. Alternatively, if the production of soluble proteins only is desired, the culture can also be grown to a higher cell density before induction and the expression period can be kept to a minimum. Also the reduction of the IPTG concentration from 1 to 0.005 mM can be selected, which would reduce the expression level by 90–95%.

In this study, the conditions shown in Table 2 were selected. The yield in wet weight of cells was 22 ± 5 g per liter of TB medium under these conditions. The His_6-*BcEn*8 protein was purified by immobilized nickel chelate affinity chromatography through the $[His]_6$ sequences positioned at the N- and C-terminals, being synthesized as a soluble protein in the host (data not shown); none appeared to accumulate in inclusion bodies on reduction for the growth temperature following induction. Consequently, the protein was purified under native conditions given under "EXPERIMENTAL PROCEDURES." The yield of the purified protein was 38 ± 3 mg/liter.

Production of Recombinant Protein BcEn16—Based on the confirmed expression and purification of recombinant protein His_6-*BcEn*8, another expression vector containing a large gene with 16 repeats of the monomer DNA

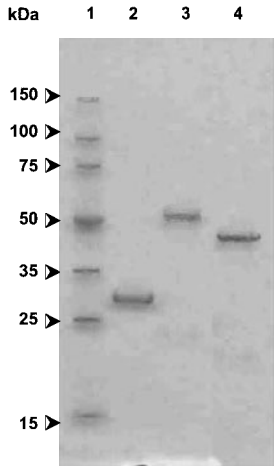

Fig. 5. **SDS-PAGE analysis of purified proteins stained with Coomassie Blue R-250.** Lane 1, Perfect Protein™ Markers; lane 2, recombinant protein His_6-*BcEn*8; lane 3, recombinant protein His_6-*BcEn*16; and lane 4, CNBr-cleaved protein, *BcEn*16.

was constructed for protein production. The yield of purified recombinant protein His_6-*BcEn*16 was 36 ± 3 mg/liter.

Figure 5 shows the results of SDS-PAGE analyses of synthesized proteins stained with Coomassie Blue R-250. Each purified protein was obtained as a single band, the molecular weights being 29 kDa for His_6-*BcEn*8 (Fig. 5, lane 2), and 52 kDa for His_6-*BcEn*16 (Fig. 5, lane 3), respectively. The periodic portion of His_6-*BcEn*16 was liberated from the flanking sequences by cleavage with CNBr. The SDS-PAGE gel stained with Coomassie Blue R-250 for the cleaved protein is shown in Fig. 5 (lane 4). The cleaved product was observed as a single predominant band and the uncleaved protein did not remain under the cleavage conditions. The molecular weight of the cleaved protein, *BcEn*16, is around 48 kDa. The molecular weights of all the proteins, as determined by electrophoresis, were a little higher than the expected values, which could be due to the anomalously low mobility of silk-like periodic proteins, as previously reported (*22*).

Structural Characterization of Recombinant Proteins in the Solid State—Figure 6 shows the ^{13}C CP/MAS NMR spectra of recombinant protein His_6-*BcEn*16 with LiBr treatment (Fig. 6a) and methanol treatment (Fig. 6b). These treatments usually induce the silk fibroin proteins and model peptides to take on the silk I and silk II structures, respectively (*8*, *9*, *18*, *23*, *24*). Each treatment for the recombinant protein His_6-*BcEn*8 and *BcEn*16 led to similar spectra to those for His_6-*BcEn*16 (data not shown). The target proteins produced in this study are flanked by N- and C-terminal extensions of 53 and 19 amino acids, respectively, as shown in Fig. 3. Both of these sequences are relatively hydrophilic and so in theory should not interfere with the protein tertiary structure. In general it appears that target proteins retain full activity even if His·Tag peptides are still attached. For comparison, the ^{13}C CP/MAS NMR spectra of $(VPGVG)_6$ and the Cp fraction [predominantly $(GAGAGS)_n$ (*1*, *6*)], which correspond to the two mono-blocks of this recombi-

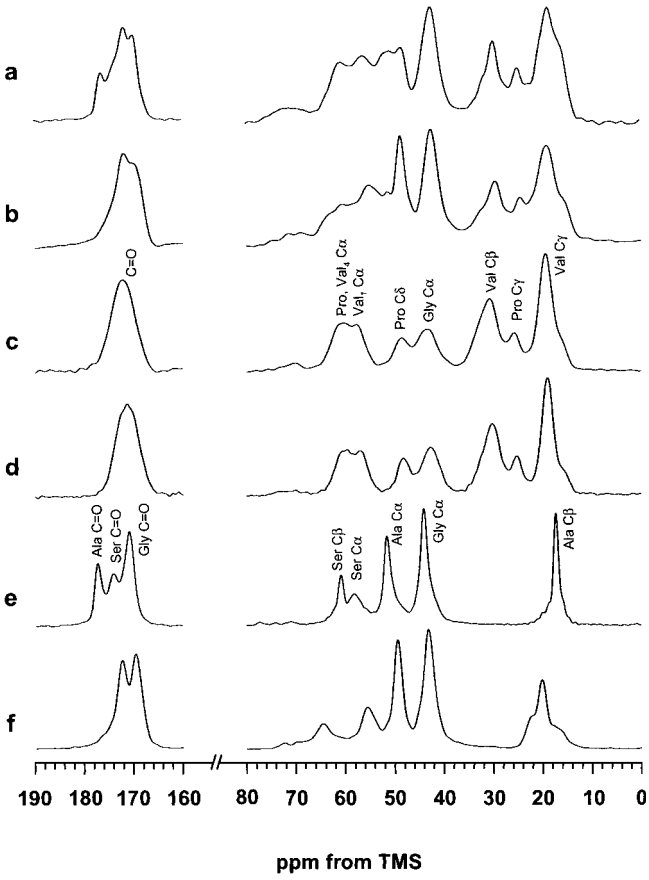

Fig. 6. **^{13}C CP/MAS NMR spectra.** a, His$_6$-*BcEn*16 after 9M LiBr treatment; b, His$_6$-*BcEn*16 after methanol treatment; c, lyophilized (VPGVG)$_6$; d, (VPGVG)$_6$ after formic acid treatment; e, Cp fraction in silk I form; and f, Cp fraction in silk II form.

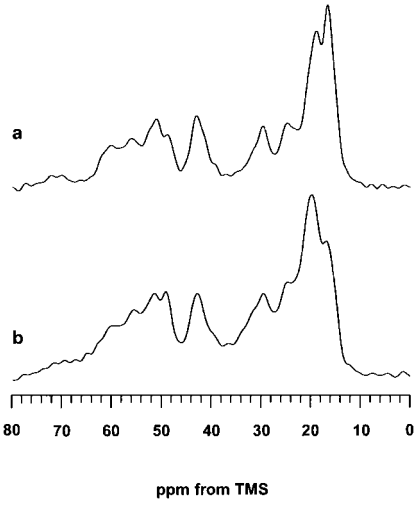

Fig. 7. **^{13}C CP/MAS NMR spectra.** a, [3-^{13}C]Ala-His$_6$-*BcEn*16 after 9 M LiBr treatment; and b, [3-^{13}C]Ala-His$_6$-*BcEn*16 after methanol treatment.

nant protein, are also shown in Fig. 6 (c–f). The ^{13}C NMR peak assignments have been reported previously (*8*, *25*, *26*). The assignment of two mono-blocks was performed based on their chemical shift values and peak intensities.

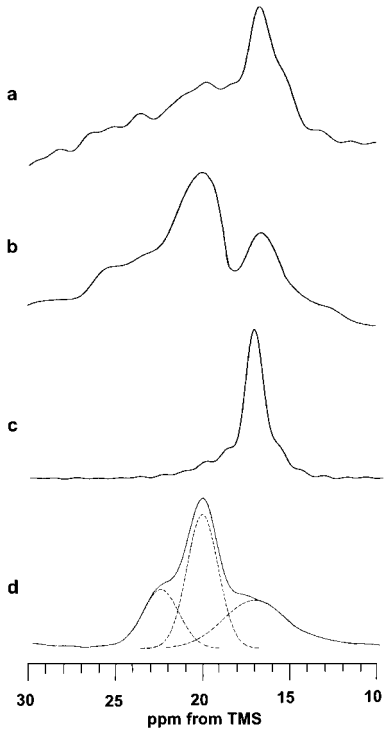

Fig. 8. **Comparison of Ala Cβ peaks.** a, difference spectrum obtained by subtracting Fig. 6a from Fig. 7a; b, difference spectrum obtained by subtracting Fig. 6b from Fig. 7b; c, expanded Ala Cβ peak of the Cp fraction in the silk I form; and d**,** expanded Ala Cβ peak of the Cp fraction in the silk II form. The broken lines indicate the peak deconvolution performed by assuming Gaussian, suggesting there are three kinds of structural component in the silk II form of the Cp fraction (*8*, *9*).

The structural transition of the Cp fraction from silk I to silk II is obvious, while no transition occurs for (VPGVG)$_6$ with the treatments. The broad peaks for (VPGVG)$_6$ suggest that it may take on several conformations, which is supported by our 2D spin-diffusion NMR data (Asakura, T. *et al.*, Polymer J., submitted). The structural transition from silk I to silk II for His$_6$-*BcEn*16 can also be observed in the change in the line-shapes, for example, the methyl group region (around 15–20 ppm) and the carbonyl region (around 169–177 ppm). However, these regions are not well resolved because of peak overlapping.

To obtain more structural information, ^{13}C isotope labeling was applied to the recombinant protein production. Figure 7 shows the ^{13}C CP/MAS NMR spectra of the recombinant protein [3-^{13}C]Ala-His$_6$-*BcEn*16 with LiBr treatment (Fig. 7a) and methanol treatment (Fig. 7b), respectively, where only Ala Cβ peaks were ^{13}C-enriched. The difference spectra obtained on subtracting the natural abundance spectra from the corresponding enriched spectra are shown in Fig. 8 (a and b). On comparison of these two difference spectra with the Cp fraction (Fig. 8, c and d), it can be concluded that His$_6$-*BcEn*16 after LiBr treatment predominantly takes on the silk I structure (type II β-turn), and after methanol treatment, the material takes on heterogeneous structures although the silk II structure is predominant (*8*, *9*, *18*). Some signals were observed on the low field side (24–27 ppm) of Ala Cβ in

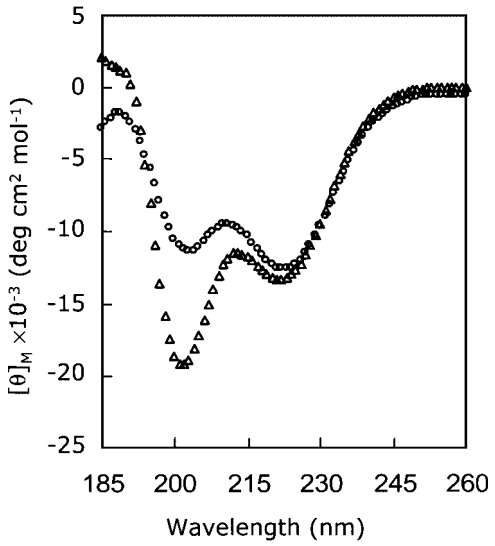

Fig. 9. **Comparison of the CD spectra of His$_6$-*BcEn*8 (open triangle) and His$_6$-*BcEn*16 (open circle) in HFA·hydrate.** Spectra were recorded at protein concentrations of 15 μM for His$_6$-*BcEn*8 and 10 μM for His$_6$-*BcEn*16, respectively.

both Fig. 8 (a and b) due to the amino acid conversion from Ala during the cell metabolism process, which could be assigned to Asp Cβ, for example. Thus, the *B. mori* silk fibroin block $(GAGAGS)_3$ tends to form the crystalline region of a *BcEn* silk-like material, which is interrupted by the flexible $(GVPGV)_2$ block.

Structural Characterization of Recombinant Proteins in the Solution State—The use of hydrates of fluoroketones, as 'special' solvents for synthetic polypeptides and some proteins, has been reported in other literature (*27–30*), because they induce and stabilize the helical structure of the peptides. Recently, HFA-hydrate was proposed to be a suitable solvent for the preparation of regenerated silk fibers (*31*). Because of the hydrophobicity of the trifluoromethyl groups, HFA molecules effectively seclude the peptide in a non-interacting environment, in which intramolecular hydrogen bond formation is energetically favored (*30*). Consequently, the existence of predominant molecular structures of the recombinant proteins is investigated by means of the CD method in a HFA-hydrate solution. Such analysis may provide an opportunity to directly correlate the secondary structural features of silk-derived peptides in the specific hydrophobic environment.

Figure 9 compares the CD spectra of the recombinant proteins, His$_6$-*BcEn*8 and His$_6$-*BcEn*16. The observed CD spectrum of His$_6$-*BcEn*8 (open triangle) is characterized by the presence of two negative bands: a shoulder at 222 nm ($[\theta]_M$ = –13000 deg cm^2 dmol^{-1}) and a relatively sharp band at 202 nm ($[\theta]_M$ = –19000 deg cm^2 dmol^{-1}), as has been found for unordered peptides with a few percentage of helical structure (*32*). However, the spectral features in the case of His$_6$-*BcEn*16 (open circle) are dramatically pronounced, with a change in the short wavelength negative band position, *i.e.,* a negative band at 202 nm shifted to 204 nm, and a positive peak at 190 nm, which are the characteristics of helical structures (*32*). While the CD pattern of His$_6$-*BcEn*8 exhibits a clear inclination for the helical structure, the observed spectrum of His$_6$-*BcEn*16 clearly indicates the predominance of helical conformations stabilized by intramolecular hydrogen bonds. On the contrary, based on the results of CD studies on elastin-like peptides $(VPGVG)_n$ (where *n* varies between 1 to 5), a lack of co-operative effects has been suggested (*33*). Moreover, the 'cooperative unit' did not seem to exceed more than one repeat. Thus, this length-dependence structural transition may result from the (GAGAGS) block.

CONCLUSION

In the past decade, the design and synthesis of artificial proteins have been an emerging area of research with important implications for structural biology, materials science, and biomedical engineering. Significant progress has been reported in the design of fibrous proteins that adopt predictable secondary structures and have higher order protein folding. With virtually absolute control of sequence, chain length, and stereochemical purity, the artificial proteins can be designed to represent a new class of macromolecular materials, with properties potentially quite different from those of the synthetic polymers currently available and in widespread use. Tirrell's group (*11*, *34*) was interested in artificial fibrous proteins able to form crystals of defined thickness, and bearing functional groups at their surface. For instance, Glu residues were inserted in $(Ala\text{-}Gly)_n$ at intervals varying between 8 and 14 (*11*). The great size of Glu compared to Ala and Gly should prevent its insertion in the lamellar crystal. Therefore, when crystals form, the Glu residues must be found on the surface. This is facilitated by the fact that, among the 20 amino acids, Glu is the one that has the lowest tendency to form β-sheet. Cappello *et al.* (*10*, *35*) focused on the production of synthetic protein analogs of *B. mori* silk, and other well-known structural protein materials, and tried to reproduce defined properties of their natural counterparts. Also, by combining structural blocks from different natural proteins, they hoped to obtain new properties found in neither the synthetic homoblock polymers nor the natural proteins.

In this study, the conditions were optimized for the production of silk-like materials by using a genetic engineering method, which is expected to be a benchmark for the production of other silk-like materials. The structural characterization suggested that this recombinant protein *BcEn* captures the structural features of native *B. mori* silk fibroin. So the final hurdle is the development of an appropriate technology capable of converting these raw materials into useful forms. Recently, we developed a technology in which native silk fibroins are used for the preparation of non-woven nanofibers (*36*). It appears encouraging to apply this technique for the further processing of recombinant proteins produced in this study.

The authors are greatful to Mr. K. Ohgo for synthesis of the model peptides.

REFERENCES

1. Asakura, T. and Kaplan, D.L. (1994) Silk production and processing in *Encyclopedia of Agricultural Science* (Arutzen, C.J., ed.) Vol. 4, pp. 1–11, Academic Press, New York

2. Demura, M. and Asakura, T. (1991) Porous membrane of *Bombyx mori* silk fibroin: structure characterization, physical properties and application to glucose oxidase immobilization. *J. Membr. Sci*. **59**, 39–51
3. Sofia, S., McCarthy, M.B., Gronowicz, G., and Kaplan, D.L. (2001) Functionalized silk-based biomaterials for bone formation. *J. Biomed. Mater. Res.* **54**, 139–148
4. Altman, G.H., Diaz, F., Jakuba, C., Calabro, T., Horan, R.L., Chen, J., Lu, H., Richmond, J., and Kaplan, D.L. (2003) Silk-based biomaterials. *Biomaterials* **24**, 401–416
5. Mita, K., Ichimura, S., and James, T.C. (1994) Highly repetitive structure and its organization of the silk fibroin gene. *J. Mol. Evol.* **38**, 583–592
6. Zhou, C.-Z., Confalonieri, F., Medina, N., Zivanovic, Y., Esnault, C., Yang, T., Jacquet, M., Janin, J., Duguet, M., Perasso, R., and Li, Z.-G. (2000) Fine organization of *Bombyx mori* fibroin heavy chain gene. *Nucleic Acids Res.* **28**, 2413–2419
7. Strydom, D.J., Haylett, T., and Stead, R.H. (1977) The amino-terminal sequence of silk fibroin peptide CP—a reinvestigation. *Biochem. Biophys. Res. Commun.* **3**, 932–938
8. Asakura, T., Yao, J., Yamane, T., Umemura, K., and Ulrich, A.S. (2002) Heterogeneous structure of silk fibers from *Bombyx mori* resolved by ^{13}C solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **124**, 8794–8795
9. Asakura, T. and Yao, J. (2002) ^{13}C CP/MAS NMR study on structural heterogeneity in *Bombyx mori* silk fiber and their generation by stretching. *Protein Sci.* **11**, 2706–2713
10. Cappello, J. and Ferrari, F. (1994) Microbial production of structural protein polymers in *Plastics from Microbes* (Mobley, D.P., ed.) pp. 35–92, Hanser Publishers, New York
11. Krejchi, M.T., Atkins, E.D.T., Waddon, A.J., Fournier, M.J., Mason, T.L., and Tirrell, D.A. (1994) Chemical sequence control of β-sheet assembly in macromolecular crystals of periodic polypeptides. *Science* **265**, 1427–1432
12. Lazaris, A., Arcidiacono, S., Huang, Y., Zhou, J., Duguay, F., Chretien, N., Welsh, E.A., Soares, J.W., and Karatzas, C.N. (2002) Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. *Science* **295**, 472–476
13. Asakura, T., Kato, H., Yao, J., Kishore, R., and Shirai, M. (2002) Design, expression and structural characterization of hybrid proteins of *Samia cynthia ricini* and *Bombyx mori* silk fibroins. *Polymer J.* **34**, 936–943
14. Nicol, A., Gowda, D.C., and Urry, D.W. (1992) Cell adhesion and growth on synthetic elastomeric matrices containing Arg-Gly-Asp-Ser. *J. Biomed. Mater. Res.* **26**, 393–413
15. Urry, D.W., Hayes, L.C., Gowda, D.C., Peng, S.Q., and Jing, N. (1995) Electro-chemical transduction in elastic protein-based polymers: a model for an energy conversion step of oxidative phosphorylation. *Biochem. Biophys. Res. Commun.* **210**, 1031–1039
16. Sambrook, J., Frisch, E.F., and Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY
17. Smith, B.J. (1988) *Methods in Biology*, *New Protein Techniques*, p. 70, Humana Press, Clifton, NJ
18. Asakura, T., Ashida, J., Yamane, T., Kameda, T., Nakazawa, Y., Ohgo, K., and Komatsu, K. (2001) A repeated β-turn structure in poly (Ala-Gly) as a model for silk I of *Bombyx mori* silk fibroin studied with two-dimensional spin-diffusion NMR under off magic angle spinning and rotational echo double resonance. *J. Mol. Biol.* **306**, 291–305
19. Aota, S., Gojobori, T., Ishibashi, F., Maruyama, T., and Ikemura, T. (1988) Codon usage tabulated from the GenBank Genetic Sequence Data. *Nucleic Acids Res.* **16**, 315–402
20. Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89
21. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685
22. McGrath, K.P., Fournier, M.J., Mason, T.L., and Tirrell, D.A. (1992) Genetically directed syntheses of new polymeric materials. Expression of artificial genes encoding proteins with repeating-$(AlaGly)_3$ProGluGly-elements. *J. Am. Chem. Soc.* **114**, 727–733
23. Iizuka, E. and Yang, J.T. (1968) The disordered and beta conformations of silk fibroin in solution. *Biochemistry* **7**, 2218–2228
24. Venyaminov, S.Y. and Kalnin, N.N. (1990) Quantitative IR spectrophotometry of peptide compounds in water (H_2O) solutions. II. Amide absorption bands of polypeptides and fibrous proteins in alpha-, beta-, and random coil conformations. *Biopolymers* **30**, 1259–1271
25. Urry, D.W., Trapane, T.L., Sugano, H., and Prasad, K.U. (1981) Sequential polypeptide of elastin: Cyclic conformational correlates of the linear polypentapeptide. *J. Am. Chem. Soc.* **103**, 2080–2089
26. Asakura, T., Kuzuhara, A., Tabeta, R., and Saito, H. (1985) Conformational characterization of *B. mori* silk fibroin in the solid state by high-frequency ^{13}C cross polarization-magic angle spinning NMR, X-ray diffraction and infrared spectroscopy. *Macromolecules* **18**, 1841–1845
27. Yang, J.J., Buck, M., Pitkeathly, M., Haynie, D.T., Dobson, C.M., and Radford, S.E. (1995) Conformational properties of four peptides spanning the sequence of hen lysozyme. *J. Mol. Biol.* **252**, 483–491
28. Kemmink, J. and Creighton, T.E. (1995) Effects of trifluoroethanol on the conformations of peptides representing the entire sequence of bovine pancreatic trypsin inhibitor. *Biochemistry* **34**, 12630–12635
29. Cammers-Goodwin, A., Allen, T.J., Oslick, S.L., McClure, K.F., Lee, J.H., and Kemp, D.S. (1996) Mechanism of stabilization of helical conformations of polypeptides by water containing trifluoroethanol. *J. Am. Chem. Soc.* **118**, 3082–3090
30. Rajan, R., Awasthi, S.K., Bhattacharjya, S., and Balaram, P. (1997) "Teflon-coated peptides": hexafluoroacetone trihydrate as a structure stabilizer for peptides. *Biopolymers* **42**, 125–128
31. Yao, J., Masuda, H., Zhao, C., and Asakura, T. (2002) Artificial spinning and characterization of silk fiber from *Bombyx mori* silk fibroin in hexafluoroacetone hydrate. *Macromolecules* **35**, 6–9
32. Woody, R.W. (1995) Circular dichroism. *Methods Enzymol.* **246**, 34–71
33. Reiersen, H., Clarke, A.R., and Rees, A.R. (1998) Short elastin-like peptides exhibit the same temperature-induced structural transitions as elastin polymers: implications for protein engineering. *J. Mol. Biol.* **283**, 255–264
34. 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) Poly(L-alanylglycine): Multigram-scale biosynthesis, crystallization, and structural analysis of chain-folded lamellae. *Macromolecules* **30**, 42–49
35. Megeed, Z., Cappello, J., and Ghandehari, H. (2002) Genetically engineered silk-elastinlike protein polymers for controlled drug delivery. *Adv. Drug Delivery Rev*. **54**, 1075–1091
36. Ohgo, K., Zhao, C., Kobayashi, M., and Asakura, T. (2003) Preparation of non-woven nanofibers of *Bombyx mori* silk, *Samia cynthia ricini* silk and recombinant hybrid silk with electrospinning method. *Polymer* **44**, 841–846