Trigonal Crystal Structure of *Bombyx mori* Silk Incorporating a Threefold Helical Chain Conformation Found at the Air−Water Interface

Regina Valluzzi, Samuel P. Gido,* and Weiping Zhang

Department of Polymer Science and Engineering, University of Massachusetts−Amherst, Amherst, Massachusetts 01003

Wayne S. Muller and David L. Kaplan†

United States Army Natick Research, Development, and Engineering Center, Natick, Massachusetts 01760

Received November 29, 1995; Revised Manuscript Received August 16, 1996[⊗]

ABSTRACT: A new crystalline polymorph of *Bombyx mori* silk, which forms specifically at the air−water interface, has been characterized. This new polymorph has a trigonal crystal structure and is distinctly different from the two previously observed silk crystal structures, silks I and II. Our identification of this new silk polymorph is based on evidence from transmission electron microscopy and electron diffraction, coupled with molecular modeling. Electron diffraction indicates that the crystal structure has a trigonal unit cell. This structure consists of a hexagonal packing of chains, each of which assumes a three-fold helical conformation. The resulting crystal structure is found to be similar to that observed for polyglycine II. The sterics of the alanine and serine residues in the crystallizable segments of silk fibroin strongly favor a left-handed 3/2 helix over a right-handed 3/1 helix. Electron diffraction from unoriented samples (powder-type diffraction) provides quantitative support for a left-handed polyglycine II type of 3/2 helical conformation for the silk chains in the crystals of this new polymorph. Single-crystal diffraction patterns and patterns from uniaxially oriented samples are consistent with the proposed crystal structure.

1. Introduction

The relationship between protein chemical structure and conformation has been an active topic of study for decades. The majority of studies have focused on the formation of secondary structure and subsequent tertiary structure selection in the globular proteins. The fibrous proteins are in some ways simpler than the globular proteins, often featuring regular repeating motifs in their amino acid sequences. These fibrous, or structural, proteins tend to crystallize in regular helical forms, often exhibiting crystalline polymorphism. An understanding of polymorphism and conformation selection in fibrous proteins will aid our understanding of the conditions that influence conformation selection, and hence secondary structure formation, in proteins in general. One of the most thoroughly studied fibrous proteins is *Bombyx mori* silk. Crystalline dimorphism in silk was observed in the 1950s[1−4] and has been an active topic of study ever since.

There are two previously known crystalline forms of *B. mori* silk, silks I and II. The conditions under which these two polymorphs are formed have been well documented in previous studies.[2,3,5−9] Both the silk I and silk II polymorphs can be obtained from natural silk solutions, which exist in the *B. mori* silk gland, or from regenerated aqueous silk fibroin solutions, where LiBr or LiSCN is used to solubilize the fibroin from silk cocoons, and the resulting solution is dialyzed to remove these salts. Natural silk fibers are semicrystalline, consisting of the silk II crystalline form as well as highly aligned amorphous regions. Silk II has a monoclinic unit cell with the protein chains in a pleated β-sheet (two-fold "zig-zag") conformation,[4,7] which forms when silk is sheared during crystallization.[10] The structure of silk I[1−3] is not well characterized. This crystalline form is an unstable lattice easily destroyed by shear or by addition of precipitating agents to an aqueous silk solution.

Here we report the discovery of a new crystal structure of *B. mori* silk which has been observed in ultrathin films formed at the air−water interface. This new structure has trigonal symmetry, suggesting a three-fold helical polyglycine II type of conformation for the silk chains. The silk three-fold helix is left-handed, corresponding to a 3_2 crystallographic screw axis. Relative intensities from diffraction data support a 3/2 helical structure. The new three-fold helical silk polymorph was first observed, in our initial experiments,[11] in surface excess layers picked up off of a trough by the Langmuir−Blodgett (LB) dipping method. In these experiments, the silk films formed at the surface of a subphase consisting of a ∼0.1% aqueous solution of LiBr-regenerated silk fibroin. A hexagonally packed crystal structure with a three-fold helical protein chain conformation results when the surface layer is allowed to form and is then compressed. Several different surface pressures were studied, ranging from 16.7 to 34 mN/m.[12,13] All of the resulting films contain needle-like crystallites of the new helical silk structure. In addition, larger lamellar crystallites are observed at 16.7 mN/m, and thus we have concentrated our studies at this surface pressure. These semicrystalline silk films possess a uniaxial oriented texture. Longer times at 16.7 mN/m increase the crystallinity of the films and the orientation. Free-standing films cast from dilute aqueous silk solution by dropping solution directly onto TEM grids without a supporting substrate have the same trigonal crystal structure as the LB films but

* To whom correspondence should be addressed.
† Present Address: Biotechnology Center, Department of Chemical Engineering, Tufts University, Medford, Massachusetts 02155.
⊗ Abstract published in *Advance ACS Abstracts,* December 1, 1996.

S0024-9297(95)01775-X CCC: $12.00

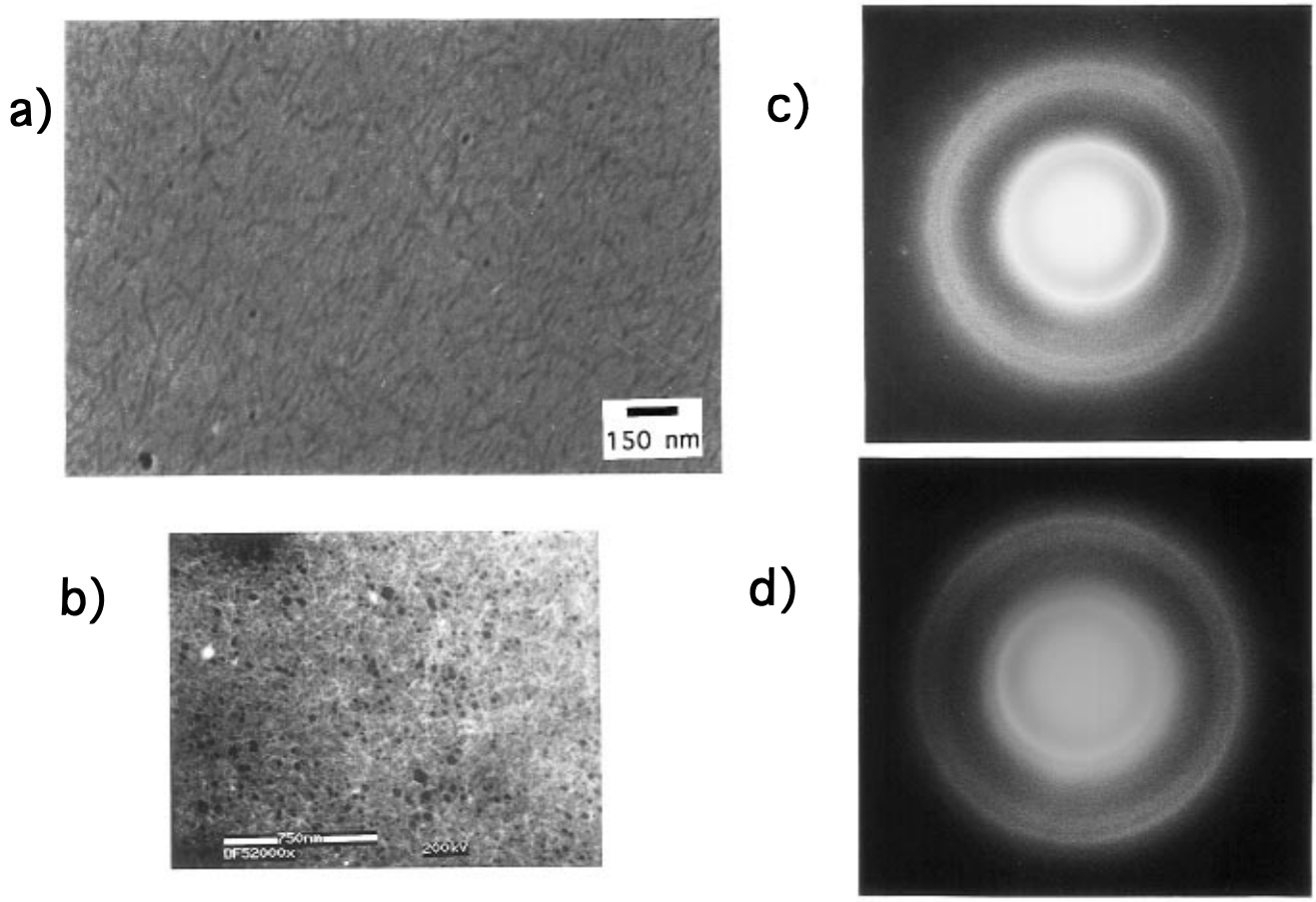

Figure 1. Semicrystalline film morphology and electron diffraction. (a) Bright-field TEM micrograph of an LB silk film. (b) Dark-field STEM image showing needle-like crystallites. These appear bright. (c) Typical pattern of electron diffraction rings for an LB film at 16.7 mN/m. (d) Typical diffraction pattern for an unoriented cast silk film.

possess a powder-like orientation. We refer to these samples as "cast films". We have investigated the morphologies and crystal structures formed by silk proteins in these films using TEM and electron diffraction. Figure 1 shows the type of semicrystalline film morphology we observe and typical electron diffraction patterns for the LB and cast films. Following the nomenclature of previously discovered polymorphs, we will call this structure silk III. The name silk III has been used previously[14] for another reported structure of silk fibroin also displaying pseudohexagonal symmetry, which nevertheless appears to be different from the structure presented here.

B. mori silk consists of two major protein components, sericin and fibroin. The sericin protein acts as a binder, holding the strong semicrystalline fibroin protein filaments together to form a cocoon.[15] The crystallizable sequence of silk fibroin can essentially be considered as a repeating hexapeptide: (Gly-Ala-Gly-Ala-Gly-Ser)$_x$.[7,16−19] Many silklike polypeptides possess a crystalline polymorph that incorporates a 3/2 helical, or left-handed three-fold helical, conformation. This polymorph occurs for copolypeptides with more than 50% glycine and appears to be most common in poly*tri*peptides such as poly(Gly-Gly-Ala).[20−24] *B. mori* silk has a more complex sequence of alanine, glycine, and serine residues in its crystallizable segments. *B. mori* silk contains almost exactly 50% glycine in its crystallizable segments and is thus poorer in glycine than the polytripeptides exhibiting 3/2 helical crystalline chain conformations. Nevertheless, the strict alternation of glycine with the larger residues, serine and alanine,

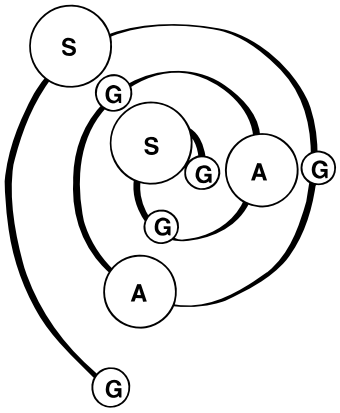

Figure 2. Diagram of silk fibroin in three-fold helical conformation.

results in favorable sterics if the silk takes on a 3/2 helical conformation. This is illustrated in Figure 2 by diagramming the placement of the different residues as seen looking down the chain axis. In the 3/2 (polyglycine II) helical conformation, the intramolecular nearest neighbors to serine (the largest residue) are all glycine residues. In contrast, the four- and six-fold helices place large serine and alanine residues in close spatial proximity. The two-fold (β-sheet) conformation also places all of the larger functional groups on one side of the chain backbone, but in this instance the chain is extended enough to accommodate them easily.

2. Experimental Section

Sample Preparation. Pure silk fibroin protein was obtained by removing the sericin from silkworm silk by boiling

the cocoons in water, $CaCO_3$, and sodium dodecyl sulfate (SDS). This process is known as degumming. Fibroin samples degummed without the addition of SDS, using only hot water or hot water and sodium carbonate to remove sericin, were also examined as controls. The resulting fibroin protein was then dissolved in aqueous LiBr or LiSCN solution. The fibroin−Li salt solution was dialyzed against distilled water for 1 week using a 12 000−14 000 molecular weight cutoff dialysis membrane to remove the lithium salts and any other impurities. The dialyzed solution was put into a Langmuir trough containing an aqueous subphase; this resulted in a protein solution concentration in the trough of about 0.1%. A surface excess layer of silk then formed at the air−water interface. Ultrathin silk films were deposited onto TEM grids using the LB dipping technique. The films were allowed to air dry. Nickel TEM grids without a supporting substrate were used for the LB samples. Other samples were prepared by placing droplets of dilute aqueous silk solution onto uncoated nylon or gold TEM grids. The solution dried to form an unsupported silk film spanning the holes in the TEM grid. These samples are referred to as *cast films*. Two sets of TEM samples were prepared from these solutions: (1) LB silk films which remained on the trough for either 1 h or 24 h at 25 °C and a surface pressure of 16.7 mN/m before being deposited onto TEM grids and (2) unsupported cast films prepared from a 3 wt % aqueous silk solution.

Characterization. The samples of LB silk and cast silk films were studied using a JEOL 2000FX TEM operated at 200 kV. TEM imaging, selected area electron diffraction, and defocused diffraction techniques were used to observe the morphology of the silk films and to determine the crystal structures of the crystalline silk within the films. Low dosage techniques were used to minimize beam damage. Electron microscopy and diffraction were performed at cryogenic temperatures of approximately −160 °C. The crystals displaying the silk III helical structure are very beam-sensitive, requiring only 0.05−0.10 C/cm^2 at 200 kV for the diffraction spot pattern to completely disappear at cryogenic (−160 °C) temperatures.

To determine the presence of a preferred crystalline orientation in the LB films, and to demonstrate powder (random) orientation of the crystallites in the cast films, pairs of diffraction patterns were taken at two different tilt orientations between the sample and the incident beam. Such a *diffraction tilt pair* was obtained by recording two diffraction patterns of the same exact sample area, separated by about 25° of goniometer tilt. Since this technique involves collecting two diffraction patterns in succession from the same region of the sample, we were concerned about the effects of accumulated beam damage, especially on the second pattern of each pair. To distinguish variations in the recorded diffraction patterns due to tilting from beam damage effects, the order of the pair collection was alternated. Ten diffraction tilt pairs were taken from different areas of LB films (sample 1) at tilts of 0° and 25°. Five tilt pairs were recorded from the cast films (sample 2) in order to demonstrate a consistent powder-like orientation.

Diffraction Data Analysis. Diffraction patterns, recorded on X-ray film, were digitized using a microdensitometer with a pixel size of 100 μm to obtain plots of integrated intensity, I, vs the inverse spacing, $1/d$. An exponential background and constant baseline were fit to each I vs $1/d$ plot. Integrated intensities of individual reflections were obtained using a Voigt profile to fit the peaks.[25,26] Experimentally, we see a small but noticeable Lorentzian component to the diffraction peak profiles which is convoluted with a larger Gaussian component. In all cases, the amorphous scattering can be modeled with a broad Lorentzian peak. Gold, sputtered directly onto the silk films, provided an internal standard for the determination of diffraction camera lengths and thus lattice spacings. Interplanar spacings and integrated relative intensities were analyzed for seven powder patterns from cast films.

Analysis of Possible Contaminants. Assurance was needed that the observed diffraction ascribed to the new silk structure came from silk and not residual Li salts or other contaminants. In a cryogenic study, ice is always a concern, but there are several reasons why the crystalline structure observed cannot be ice. Patterns can be obtained at room temperature with the same largest spacings as the cryogenic diffraction patterns, indicating that the same crystal structure is being observed at both temperatures. The largest spacing observed for our silk structure is 4.55 Å, while all three forms of ice (cubic, amorphous, and hexagonal) have a largest d-spacing of 3.67−3.7 Å.[27,28] Hexagonal ice has been observed as an impurity in the cryogenic diffraction studies, and its presence is accounted for in our analyses of the diffraction patterns.

A LiBr hydrate has been observed in some of the LB silk films, even though a great deal of care was taken to remove as much salt as possible by dialysis. The LiBr hydrate that is observed is scarce and occurs as separate crystallites embedded in the amorphous part of the silk film. To distinguish diffraction by LiBr hydrate crystallites from silk, electron diffraction was performed on LiBr hydrate crystals obtained by casting LiBr solution onto TEM grids in much the same way that the cast silk films are prepared. These crystallites can easily be distinguished from the silk crystallites by their diffraction patterns, their size and shape, and spatially resolved elemental analysis techniques such as energy dispersive X-ray spectroscopy (EDS) and parallel electron energy loss spectroscopy (PEELS). Films on copper grids were found to contain significant cuprous impurities, but nickel, gold, and nylon grids did not introduce impurities. All of the quantitative diffraction data were obtained using nickel or nylon grids. EDS compositional analysis obtained from fibrous silk films of the type shown in Figure 1 revealed no significant inorganic impurities. The only distinctive morphological features were the needle-shaped silk crystals, and PEELS indicated that these regions were, indeed, organic material with a protein-like elemental composition, i.e., nitrogen, oxygen, and carbon in reasonable proportions. The protein-like elemental composition observed also ensures that we have obtained films of silk protein and not silk wax or other non-protein impurities present in cocoons that could conceivably segregate to the air−water interface. TEM images and electron diffraction data correlate sample regions consisting only of needle-like crystallites to diffraction patterns containing no detectable salt.

The diffraction patterns observed for the films prepared from LiSCN-solubilized silk are identical to the patterns from LiBr solubilized silk. The fact that the difference in anions does not affect the silk structure suggests that the silk III structure does not involve the salt. The possibility of anions from the LiBr or LiSCN salt being included in the crystalline unit cell is also precluded because inclusion of a Br or SCN anion would make the crystals unrealistically dense. The unit cell is not large enough to easily accommodate a bromide or thiocyanate anion in addition to the silk helix already present. The density calculated for the silk III unit cell is 1.56 g/cm^3, which is very similar to the density of 1.6 g/cm^3 measured for polyglycine II.[29] The silk III unit cell volume is 190 Å^3, and a bromide ion has a volume of 32 Å^3. Thus, a single bromide ion would occupy 17% of the unit cell by itself, resulting a density of 2.1 g/cm^3 when one protein chain is also accommodated. The LiBr hydrate diffraction peaks, when present, could be identified as such in the diffraction data from regenerated silk samples. All of the LB film and cast film data used in quantitative diffraction analysis were from LiBr-solubilized silk fibroin films. Samples prepared with LiSCN were used to compare the effects of the different salt anions.

The degummed silk fibroin and the dialyzed aqueous solutions were analyzed for amino acid content to ensure that the sericin from the cocoons had been removed. Sericin and fibroin have similar proportions of most of the amino acids[30,31] but differ significantly in their aspartic acid and glutamic acid contents. Fibroin contains approximately 1.39 mol % aspartic acid and 1.06 mol % glutamic acid, whereas all of the sericins contain between 20 and 30% of these residues. The degummed silk prepared using only boiling water to remove sericin contained 3.1% aspartic acid and 2.2% glutamic acid, indicating that there was some residual sericin in these samples. These samples also contained crystallites with spacings typical of the two high molecular weight sericins, sericins III and IV. The sericin crystallites had habits and diffraction patterns that

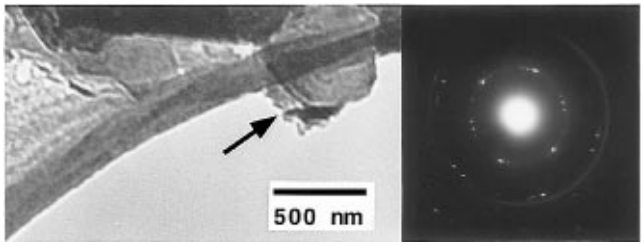

Figure 3. Bright-field TEM image and electron diffraction pattern of a hexagonal crystallite of three-fold helical silk III from a 16.7mN/m LB silk film.

were easily distinguished from those of the fibroin in the films. The number of sericin crystallites also increased when the films were exposed to moisture. The amorphous-to-β transition in sericin occurs as humidity increases, corroborating our identification of the sericin in the films degummed with boiling water only. Solutions degummed using boiling water, $CaCO_3$, and SDS along with solutions prepared using boiling water and $CaCO_3$ without SDS had aspartic and glutamic acid contents that were identical, within experimental accuracy, to those reported for fibroin: 1.2 mol % aspartic acid and 1.0 mol % glutamic acid.[30] In addition, LB films were collected from the surface of the trough, and amino acid analysis was performed to determine the composition of the LB films. Comparison of amino acid compositions of LB films produced from silk degummed using $CaCO_3$ and SDS with published values for fibroin and sericin[30,31] indicate that the films are pure fibroin to within the limits of the measurement. No sericin crystallites were observed in these films using TEM and electron diffraction. The silk III structure was clearly observed in all of the LB films, regardless of the degumming method used, but solutions degummed using SDS and $CaCO_3$ in addition to boiling water were used in our crystal structure analyses in order to avoid artifacts due to sericin contamination.

3. Results and Discussion

Figure 1a shows a bright-field TEM micrograph of a silk LB film (surface pressure, 16.7 mN/m) from sample 1, indicating that the films consisted of very small needle-like crystallites embedded in a featureless amorphous matrix. The dark-field STEM image in Figure 1b confirms the crystalline nature of the needles, which appear bright in the figure due to diffraction contrast. In general, the films with longer residence times at 16.7 mN/m appeared to contain more needle-like crystallites than those which spent less time compressed on the surface at this pressure. These mats of needle-shaped crystallites result in rings of diffracted intensity, as shown in Figure 1c and d for LB films and cast films, respectively. Additionally, faint images of very thin triangular lamellar crystals were sometimes visible in the LB films. Figure 3 shows a bright-field TEM image containing several lamellar crystals; the corresponding six-fold diffraction pattern from these crystals is also shown. In the figure, the edges of one of the crystallites can be clearly seen. This crystallite is indicated with an arrow and has a hexagonal shape, consistent with the proposed crystal structure and the six-fold diffraction seen in the inset pattern. The slight asymmetry in the intensities of the innermost set of diffraction peaks is due to a small tilt of the crystallite out of the plane perpendicular to the incident electron beam. The two rings in the diffraction pattern are due to the gold standard. Figure 4 shows an additional spot pattern taken from a single lamellar crystal at room temperature. These single-crystal patterns are frequently superimposed on rings, as can be seen in the figure, because the lamellar crystal invariably rests on a mat

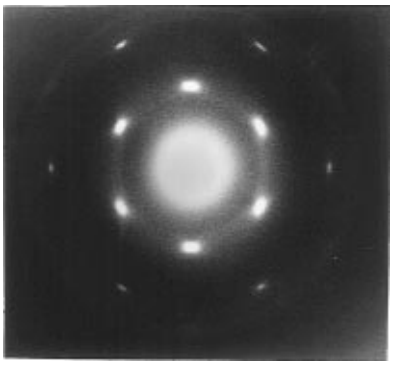

Figure 4. Hexagonal single-crystal pattern from a lamellar crystal taken at room temperature showing superimposed rings from surrounding small crystallites.

Table 1. Experimental and Model *d*-Spacings and Relative Intensities ($R = 0.11$)[a]

	experimental diffraction				predicted model diffraction	
index	*d* (Å)	SD *d*	*I* (rel)	SD[a] *I*	*d* (Å)	*I* (rel)
(100)	4.56	0.025	100	0.0	4.42	100
(101)	3.93	0.023	47	8.2	4.00	45
(102)	3.29	0.060	28	9.3	3.22	43
(003)	2.89	0.011	4	0.5	3.14	2
(103)	2.67	0.058	11	1.9	2.56	11
(110)	2.48	0.007	3	1.2	2.55	3
(111)	2.36	0.006	5	2.5	2.46	2
(112)	2.23	0.080	4	0.7	2.24	2
(201)	2.15	0.090	2	1.0	2.15	1
(104), (113), (202)	2.0−2.1	0.016	9	1.5	2.0−2.1	9

[a] SD = standard deviation.

of needles within the region from which diffraction data are collected. Cast films produced similar needle-like crystalline morphologies to the LB films which resulted in ring diffraction patterns with identical *d*-spacings to those observed for the LB films. The apparent differences in the ring patterns obtained from cast and LB films are due to different crystalline orientations, or textures, within the two types of film. As will be discussed in more detail later, the cast films are unoriented, yielding powder diffraction patterns, whereas the LB films have a uniaxial orientation. The powder diffraction data obtained from the cast films were used for quantitative comparison of observed diffracted intensity to the intensities predicted by structural models. Table 1 lists the *d*-spacings and integrated intensities for the 10 strongest reflections averaged over the seven cast film (powder) patterns analyzed.

The lattice spacings, symmetry, and peak intensities in the silk III diffraction patterns approximately match those found previously by Keith and Padden for polyglycine II.[29] Thus, we have constructed a model unit cell for silk III using polyglycine II as a structural motif. This model incorporates a three-fold helical chain conformation and trigonal symmetry, resulting in a hexagonally packed arrangement of chains. When the unoriented diffraction patterns are indexed using a trigonal unit cell, a weak 003 reflection is observed, but the 001 and 002 reflections are absent. The systematic absence of all 00*l* reflections except when *l* is a multiple of 3 is indicative of a three-fold screw axis, resulting from a three-fold helical chain conformation. In polyglycine II, each polypeptide chain adopts a three-fold helical conformation,[29] resulting in the following unit cell dimensions: $a = b = 4.8$ Å, $c = 9.3$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$. The silk III unit cell, as determined by electron diffraction, has a similar shape with dimensions

$a = b = 5.1$ Å, $c = 9.42$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$. The Miller index assignments of experimentally observed silk III diffraction maxima based on this unit cell are listed in Table 1. The spacings and symmetry of the six-fold pattern of spots in the single-crystal silk III diffraction pattern (Figure 4) indicates a similar hexagonally packed structure, analogous to that found for polyglycine II.

Unit Cell Modeling. The silk III crystal structure was modeled using the Cerius software package from Molecular Simulations Inc., running on a Silicon Graphics Indigo workstation. Based on the similarity of silk III diffraction patterns to those of polyglycine II, crystal structures with a three-fold helical chain conformation were investigated. In polyglycine II, the glycine residues comprising the helix are nonchiral, and the helix is not expected to have a preferred handedness. In other proteins, the chirality of the residues causes each specific helix type to favor a particular handedness. The fact that all known examples of three-fold helices in proteins, besides the degenerate case of polyglycine, occur in the left-handed form leads us to suspect that left-handed helices are likely for silk as well. The preponderance of left-handed structures when a protein made up of L-amino acids adopts a three-fold or nearly three-fold helical conformation can be understood by finding the left- and right-handed three-fold helical conformations on a Ramachandran map,[32,33] where protein and polypeptide chains with left-handed three-fold helical conformations are clustered in the same region as the 3.3-fold collagen helix, with $\phi = -60$ and $\psi = 150$ or $(-60, 150)$.[33−35] All of these left-handed three-fold conformations fall well within the allowable "extended chain" region of the map. The corresponding right-handed helices are all in the forbidden region around $(60, -150)$. The left-handed three-fold silk III helix, a helical structure based on the polyglycine II conformation with slight changes in ϕ and ψ to match the larger *c* dimension found for silk III, has angles of $(-77, 142)$, placing it well within the same allowable region as collagen, whereas a right-handed three-fold silk helix would fall in a sterically forbidden region.

Molecular mechanics energy calculations on single chains with the Dreiding II force field were employed to compare the sterics of left- and right-handed versions of a three-fold helix with the Gly-Ala-Gly-Ala-Gly-Ser repeat sequence. In these energy minimization calculations, helices were constrained to a dimension of 9.4 Å per turn to correspond with the *c*-dimension of the observed unit cell. These calculations indicate that the left-handed (3/2) helix is about 4 times more stable than the right-handed helical conformation. The left-handed (3/2) helical conformation is observed in numerous other proteins and synthetic polypeptides, providing a precedent and basis for comparison. Molecules of poly(L-proline), poly(Ala-Gly-Gly), poly(Ala-Gly-Gly-Gly), and poly(Gly-Ser-*Pro*) form three-fold left-handed helices, while collagen forms a left-handed helix that is also close to a three-fold helical conformation.[20,34−40]

A model unit cell for helical silk III was constructed using polyglycine II as a general structural motif. Ramachandran et al.[41,42] proposed a modification to the Rich and Crick polyglycine II structure which improves the H-bonding interaction for polyglycine II. However, the Ramachandran polyglycine II structure does not easily accommodate the side groups present in silk. The original Rich and Crick polyglycine II structure does provide a sterically reasonable model when the residues present in silk are incorporated into a left-handed version of the helix, and thus this structure was used as a basis to construct model silk III structures.

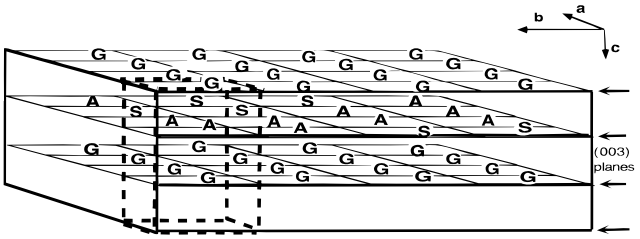

Figure 5. Arrangement of residues in the (003) planes of a silk III crystal where planes containing both alanine and serine alternate with planes containing only glycine.

There are several distinguishable ways of packing 3/2 helices of the crystallizable silk sequence while preserving intermolecular H-bonds with reasonable lengths and angles. The distances between a helix and its six nearest neighbors must be equal in order to maximize the stabilization due to H-bonding. In addition, the residues must all lie in the same planes of the crystal, the (003) planes, and the H-bonds must be perpendicular to the helical axis (*c* axis). The different packing possibilities are distinguished by the identities and arrangement of residues in each (003) plane of the crystal, as is illustrated in Figure 5. For example, the packing arrangement in Figure 5 results in planes of glycine which alternate with planes containing both alanine and serine. One can also devise arrangements in which each (003) plane contains all three types of residues arranged in various regular patterns. Or, alternatively, it is possible for each (003) plane to contain all three residue types in a random arrangement. An ordered packing of residues in the (003) planes would lead to a slight distortion of the unit cell to accommodate the different residue volumes. Diffraction evidence of perfect hexagonal packing of the silk helices implies a large degree of randomization of the sequential pattern of residues within each of the (003) planes in the lateral (*a* and *b*) directions. Sufficient randomization in the sequence of residues in the (003) planes would result in diffraction that "sees" an apparent average residue identity at each residue position. This results in a primitive unit cell containing one full turn of the 3/2 helix, or three average residues in the *c* direction, a 3_2 screw axis, and dimensions very similar to those of the polyglycine II unit cell. Because the crystallizable portion of silk fibroin is essentially a repeating hexapeptide sequence, the arrangement of residues in the (003) planes would have a six-residue periodicity in the *c* direction if the packing of residues was nonrandom. This longer periodicity would lead to a unit cell that is twice as long in the *c* direction, incorporating two helical turns (six residues) in the *c* direction. Our experimental *c* direction lattice constant indicates that only one turn of the 3/2 helix can be accommodated in the silk III unit cell, supporting a random arrangement of residues in the (003) planes.

Hydrogen-bonding between residues of adjacent chains is preserved regardless of the relative up−down orientation of the chains.[29] A regular arrangement of up and down chain directions would result in a unit cell containing at least two chains, but the unit cell dimensions from our diffraction experiments on silk III indicate only one chain per unit cell. Thus, we conclude that the up−down arrangement of chains is most likely random. A random arrangement of up and down chains

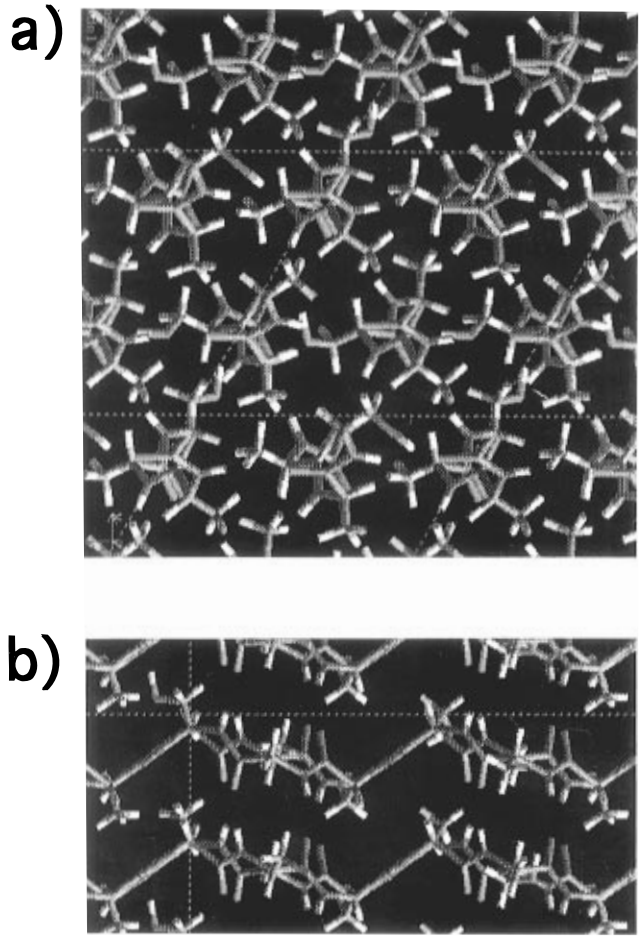

Figure 6. Superlattice structure, or "supercell", for the silk III model, consisting of two primitive unit cell repeats in each crystallographic direction. There are a total of eight primitive cells in the supercell used to construct the average primitive unit cell. Oxygen atoms are red, carbon atoms are gray, nitrogen atoms are blue, and hydrogen atoms are white. (a) View down the *c* axis (b) View down the *a* axis.

in the silk III structure was simulated by constructing a 2 × 2 × 2 supercell, as shown in Figure 6. This supercell contains two repeats of the primitive cell in each of the three crystallographic directions. This results in a group of eight subcells containing four chains (two helical turns of each chain) which can be arranged in any combination of *up* and *down* (parallel and antiparallel) directions. The eight cells in the supercell can then be superimposed in space to produce a primitive unit cell with an electron density distribution that is an average of that of the eight helical repeats. In the resulting average primitive unit cell, the superimposed chains include all of the possible arrangements of alanine, glycine, and serine in appropriate proportions in (003) planes and both helical *up*−*down* orientations in equal proportions. Unequal weighting of the various possible configurations can be used to simulate more complicated crystal structures.

A series of models using 2 × 2 × 2 supercells (superimposed to give average primitive cells as explained above) were used to model the proportions of parallel and antiparallel packing. The registry and probable degree of disorder for the packing of residues in the (003) planes was also examined. In the *a* and *b* directions, an evenly spaced packing of molecular helices was compared to a slightly disordered packing. Any minimization used to modify unit cells was carried out using the Dreiding II force field with H-bonds and van

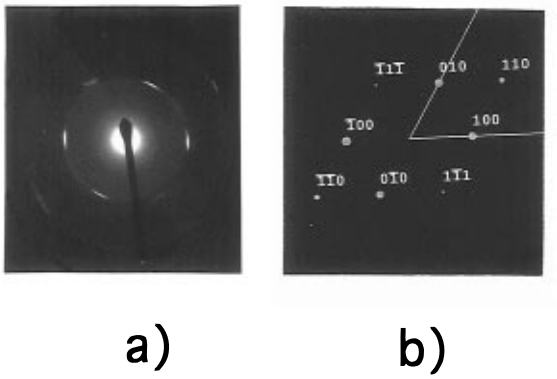

Figure 7. (a) Tilted, room temperature, single-crystal diffraction at approximately 20° sample tilt. (b) Simulated model single-crystal diffraction pattern at 16° tilt.

der Waals forces considered in the energy calculation. Dreiding II is a general force field and is not accurate enough to distinguish between closely related crystal structures; thus, comparison to experimental diffraction data, rather than energy minimization, was used as a criterion in selecting the best-fit unit cell. Quantitative comparisons were carried out using data from the unoriented, cast films. Because of the possibility of beam damage, the highest angle portions of the diffraction patterns, which are most sensitive to these effects, were not used. The smallest *d*-spacing used in analysis was ∼2.0 Å, which is sufficient for an intermediate resolution description of the crystal structure. While this description will not allow us to locate hydrogen atoms and distinguish between closely related models, it is sufficient to elucidate backbone conformation. In fact, since there appear to be numerous random elements in the packing and relative arrangements of the chains in the structure, attempts at a more resolved description may be meaningless.

Diffraction data from the unoriented cast films and simulated diffraction from the trigonal models were compared, and values of the residual, R,[43,44] were calculated to determine which models best matched the data. As indicated in Table 1, the highest angle peak is compound, consisting of the 104, 113, and 202 reflections. In our calculations of R, the experimentally determined and modeled intensities for these three overlapping peaks are lumped together. Once the best models ($R < 0.3$) were identified with preliminary calculations using no temperature factor, simulated diffraction from these models was recalculated using a series of temperature factors. This refinement of the models was carried out because a low, but non-zero, temperature factor is expected for an organic (polymeric) crystal at liquid nitrogen temperatures. The closest fit to experimental diffraction data was obtained for a model with an average primitive cell that had an equal weight for up and down chain directions, evenly spaced helices in the *a* and *b* directions, highly aligned H-bonds, and a random distribution of alanine, glycine, and serine units in the (003) planes. The lowest value of R obtained was 0.11, with a temperature factor, B_{iso}, of 7.9 Å^2 (corresponding to a value of 0.1 for U_{iso}). The supercell shown in Figure 8 was the best model. Artificial periodicities in the supercell model are necessary to ensure a correctly weighted representation of all possible structural elements, i.e., up−down chain direction and (003) plane residue arrangements. These periodicities are lost in the superposition process used to create the average primitive cell.

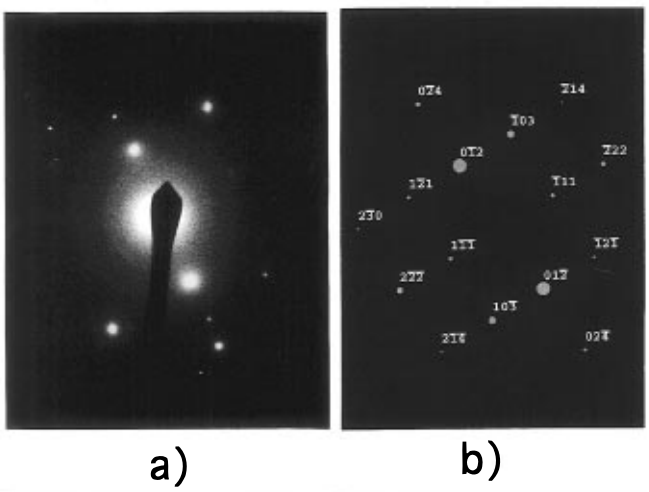

Figure 8. (a) Cryogenic electron diffraction pattern obtained for a crystal of silk III near the [321] zone axis. (b) Simulated model diffraction for a single silk III crystal near the [321] zone axis.

Packing energies and total energies were calculated for the silk III model that provided the closest match to the electron diffraction data. Energies were also calculated for silks I and II, in order to provide a basis for comparison. Energies for the three structures were similar regardless of the methods employed to constrain and minimize the structures. We believe that the energies obtained are essentially similar for the three structures within the limitations of the force field and of the molecular mechanics algorithms employed to analyze the structures. Because only the backbone is specified when a polyglycine II helical structure is used to build the three-fold silk helix, initial minimization of the silk III structure was carried out with the backbone conformation fixed. The pendant hydrogen atoms and functional groups were allowed to move in the minimization. The silk I and silk II models were constructed using Lotz's and Marsh's coordinates, respectively.[45,46] All unspecified atoms, such as hydrogen and serine side groups, were allowed to move in the minimization, unit cell parameters were constrained, and the *Q equilibrium* algorithm, in Cerius, was used to calculate charges. The values obtained for the total energy of the three structures were all of the same order when normalized per residue; 15.5 kcal/mol·residue for silk I, 51.0 kcal/mol·residue for silk II, and 25.5 kcal/mol·residue for silk III. When all atoms were allowed to move in an unconstrained minimization, an energy of around −5.5 kcal/mol·residue was obtained for silk II, and −2.3 kcal/mol·residue was obtained for silk III. In the unconstrained minimization, the imposed unit cell periodicity provided the only conformational constraint, and the two-fold silk II β-sheet and three-fold silk III helix both retained their original helical character. Energies for the silk I crankshaft crystal structure could not be obtained from unconstrained (except for crystalline periodicity) minimization because gross conformational changes occurred when the minimization was attempted. These calculations indicate that the proposed silk III structure is feasible from an energetic standpoint.

The silk III hydrogen-bonded distances were examined for the minimized structure obtained in this way. The distance between the amide hydrogen (H-bond donor) and the carbonyl oxygen (H-bond acceptor) averages 2.32 Å for the interchain hydrogen bonds, with a range of 1.9−2.8 Å. The reported average distance for H-bonds in similar bonded environments taken across a variety of organic crystalline structures is 1.93 Å.[47,48] For pleated β-sheet proteins and polypeptides, the average distance is 1.96 Å, but the NH---O and C═O---H angles are 160° and 150°, respectively.[48,49] Measurements of minimized polyglycine II structures (in Cerius) with constrained unit cell parameters indicate an average H---N distance of 2.08 Å and NH---O and C═O---H angles of 153° and 151°, respectively. In the silk III structure, the corresponding angles are 161° and 156°. If α carbons are fixed during minimization of the silk III structure, constraining the three-fold conformation more rigidly, the distance between the amide hydrogen (H-bond donor) and the carbonyl oxygen (H-bond acceptor) averages 2.16 Å for the interchain hydrogen bonds, with a range of 1.9−2.4 Å. The H-bonded angles then become 168° and 156°. While the H-bonded distances calculated for our model structure are longer than the average, the H-bonding angles are more favorable than the averages reported for β-sheets and polyglycine II. Also, the silk III H-bonded distances are shorter than the sum of the van der Waals radii for hydrogen and oxygen. Hence, a stabilization of the model structure due to hydrogen-bonding is expected. Additional hydrogen bonds involving serine can be formed between neighboring chains when serine units happen to be in close spatial proximity, leading to additional structural stabilization.

Consistency in electron diffraction data is of greater concern than in the typical X-ray diffraction study. Electron diffraction characterizes small sample volumes, and a representative population of crystals is thus harder to obtain than in the much larger samples characterized in X-ray diffraction studies. Beam damage can affect electron diffraction results in proteins and polymers, while it is not a major concern in an X-ray diffraction study. To evaluate the consistency of our experimental diffraction data as well as the statistical relevance of the *R* value for our best model, we calculated statistical dispersions of our experimental structure factors. The dispersion, *D*, is calculated analogously to *R*:

$$D = \frac{\sum_{hkl} |F_{av} - sF_{exp}|}{\sum_{hkl} F_{av}}, \quad s = \frac{\sum_{hkl} F_{av}}{\sum_{hkl} F_{exp}}$$

Here, F_{av} is the average experimental structure factor over all of the diffraction patterns for a particular *hkl*. Thus, the dispersion is a measure of the deviation of a single experimental set of structure factor components (from a single diffraction pattern) from the set of average experimental structure factor components. The dispersions for our unoriented diffraction data ranged from 0.06 to 0.12. These dispersion values indicate that an *R* value of about 0.12 is the best that can be obtained given the level of uncertainty in our experimental data. The *R* value of 0.11 which we report for our best model is right at this limit and is thus within experimental error. Models with lower *R* values cannot improve on this fit due to the uncertainty in the experimental data.

Diffraction from Different Zone Axes of the Silk III Structure. The hexagonal diffraction patterns shown in Figures 3 and 4 were each taken from the [001] zone of a trigonal silk III crystal. Because of the beam sensitivity of our samples, we were rarely able to obtain more than one zone from the same crystal.

Figure 7a shows a pattern that was obtained by tilting approximately 20° from the hexagonal [001] zone direction. Single-crystal patterns of silk III were simulated at tilt angle increments of 1° over a tilt range of 10−25°. A simulated pattern at 16° to the [001] axis (Figure 7b) displays a very similar pattern, both in arrangement of the spots and in their relative intensities. Additional single-crystal diffraction patterns, from zone axes other than the [001] axis, were obtained, and *d*-spacings were observed identical to those of the corresponding reflections from the previously discussed powder patterns. Because of beam damage and changes in the relative intensities of the reflections when a single-crystal pattern is tilted even slightly, the single-crystal patterns obtained were not used to quantitatively compare intensities with model diffraction. However, in all cases, a very good qualitative match could be obtained between experimental data and simulated patterns on the basis of the silk III structure. An electron diffraction pattern from near the [321] zone axis, seen in Figure 8a, can be compared with the corresponding simulated diffraction from the silk III structural model shown in Figure 8b. Again, in addition to having the predicted *d*-spacings, the symmetry, arrangement, and relative intensities of the simulated diffraction spots are consistent with the data.

Uniaxial Orientation of Silk LB Films. Diffraction patterns from the unoriented cast films (sample 2) were similar to patterns obtained from the LB films (sample 1). However, there were visible differences in the relative intensities of the reflections, as well as a number of absent reflections in the LB film patterns. This can be appreciated by comparing the diffraction patterns in Figure 1c (unoriented film) and Figure 1d (LB film). The peak centers are at the same *d*-spacings, but in the unoriented patterns there are differences in intensity and the appearance of additional peaks, the moderately strong 101 reflection and one very weak peak, indexed as the 102 reflection. The discrepancies in relative intensity between the diffraction from cast films and LB films indicate different average crystallite orientations for the two types of samples.

If no preferred orientation exists in a sample, then its diffraction will not change as the sample is tilted relative to the electron beam. Diffraction data obtained for the cast films revealed no change when the films were tilted, indicating that they have a powder orientation. The LB silk films were found to have a uniaxial orientation. The uniaxial orientation observed in the LB film samples is reminiscent of the orientation of polymer crystallites in a sedimented mat with the orientation axis perpendicular to the sample film and parallel to the incident electron beam at 0° goniometer tilt. In this 0° tilt geometry, only rings corresponding to *hk*0 reflections should be observed. Tilting the sample will result in *hkl* reflections, with non-zero *l*, appearing as arcs in the diffraction patterns.

Experimentally observed diffraction patterns from uniaxially oriented LB silk films at tilts of 0° and 25° are shown in Figure 9. Simulations of these tilted patterns generated in Cerius using the silk III structural model can accurately reproduce the spacings, relative intensities, and distribution of arcs observed. Figure 10a,b shows schematic diagrams of the uniaxially oriented LB film diffraction patterns at 0° and 25° of goniometer tilt. These schematics correspond to the experimental data in Figure 9a, and the labeled reflections correspond to experimentally observed diffraction

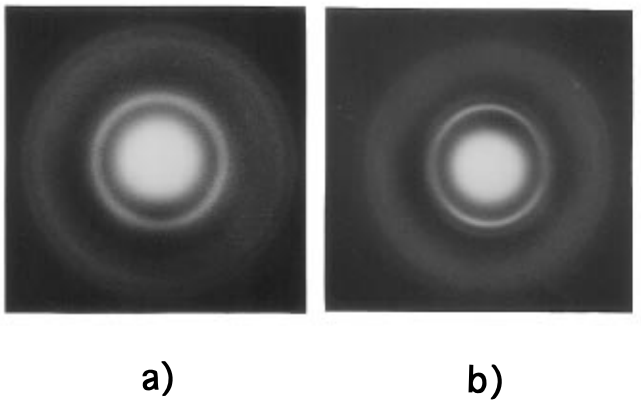

Figure 9. Diffraction patterns obtained from a uniaxially oriented (16.7 mN/m) LB film at (a) 0° sample tilt and (b) 25° sample tilt.

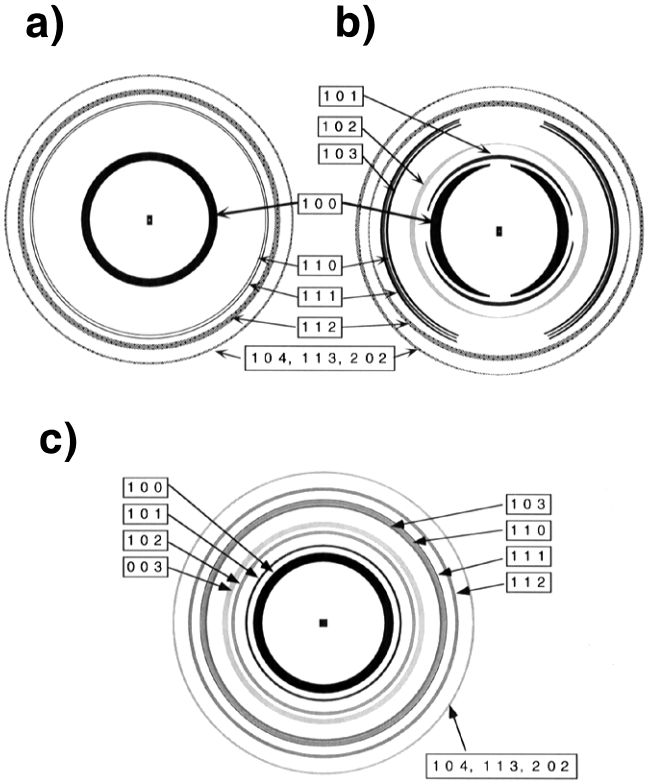

Figure 10. Schematic diagrams of the diffraction rings that result from uniaxially oriented silk III LB films and unoriented silk III cast films in the electron diffraction geometry. (a) Uniaxially oriented LB film, 0° sample tilt. (b) LB film, 25° sample tilt. (c) Unoriented cast film.

maxima which are all present on the negatives, but some of which are too faint to reproduce well in photographic prints. For comparison, a schematic diagram of an unoriented powder pattern is given in Figure 10c, showing all of the experimentally observed reflections. If the uniaxial orientation in the LB films is perfect, then Figures 9a and 10a should contain only *hk*0 reflections. The presence in this pattern of *hkl* reflections, with non-zero *l*, (103, 111, 112, etc.) is due to a combination of imperfect crystallite alignment and peak broadening due to small crystallite size. These *hkl* reflections are, however, substantially reduced in intensity relative to their appearance in the powder pattern (Figure 1c and 10c) and in the tilted LB pattern (Figure 9b and 10b).

The 101 and 102 reflections are present in both the powder patterns and tilted diffraction patterns from the

uniaxially oriented LB films (incident electron beam tilted off the normal to the film), but these reflections are absent in untilted diffraction patterns of the LB films. A sequence of diffraction patterns from different, but adjoining, areas of an LB silk film were taken at tilt angles from 0° to 25° in 5° increments. The 101 reflection first appears at 15° tilt; based on the silk III structure, it is expected to appear at 11°. The 102 reflection is expected to appear at 21° and is first observed at 20° tilt. These results are consistent with the silk III structure and partial uniaxial alignment.

4. Conclusions

A new crystal structure for silk fibroin protein, with 3/2 helical chains arranged in a hexagonally packed array, was found to occur in free-standing thin films of silk cast from dilute solution and in films of silk prepared using the LB method. The structure, silk III, is supported by six-fold single-crystal electron diffraction patterns, uniaxially oriented electron diffraction ring patterns taken at several different tilts, and unoriented "powder-like" patterns. The dimensions measured for the unit cell of silk III, when compared to those of polyglycine II, show a small increase in the *a* dimension to accommodate the larger alanine and serine residues found in the silk fibroin protein. The *c* dimensions are quite similar for the two structures, consistent with similar helical conformations. The density calculated for the silk III structure from the measured unit cell parameters is the same as that for polyglycine II within experimental error, supporting the structural model which was proposed on the basis of diffraction evidence.

Acknowledgment. This work was supported by the NSF Materials Research Science and Engineering Center (MRSEC) at the University of Massachusetts. We also acknowledge the use of the facilities of the W. M. Keck Polymer Morphology Laboratory at the University of Massachusetts—Amherst and of MRSEC central experimental facilities. We thank Dr. S. L. Hsu for providing access to the Cerius software package used in our simulations. Additional support was provided by the U.S. Army Natick Research, Development and Engineering Center. Significant discussions of this work with the following individuals are acknowledged: E. D. T. Atkins and S. Fossey.

References and Notes

(1) Kratky, O.; Schauenstein, E.; Sekora, A. *Nature* **1950**, *165*, 319.
(2) Kratky, O. *Monatsh. Chem.* **1956**, *87*, 269.
(3) Ambrose, E. J.; Bamford, C. H.; Eliiott, A.; Hanby, W. E. *Nature* **1951**, *167*, 264.
(4) Warwicker, J. O. *J. Mol. Biol.* **1960**, *2*, 350.
(5) Asakura, T.; Kuzuhara, A.; Tabeta, R.; Saito, H. *Macromolecules* **1985**, *18*, 1841.
(6) Bhat, N. V.; Ahirrao, S. M. *J. Polym. Sci., Polym. Chem. Ed.* **1983**, *21*, 1273.
(7) Lotz, B.; Cesari, C. *Biochimie* **1979**, *61*, 205.
(8) Magoshi, J.; Mizuide, M.; Magoshi, Y.; Takahashi, K.; Kubo, M.; Nakamura, S. *J. Polym. Sci., Polym. Phys. Ed.* **1979**, *17*, 515.
(9) Magoshi, J.; Kamiyama, S.; Nakamura, S. *Proceedings of the 7th International Wool Textile Research Conference*; Tokyo, 1985; pp 337.
(10) Kricheldorf, H. R.; Muller, D.; Ziegler, K. *Polym. Bull.* **1983**, *9*, 284.
(11) Zhang, W.; Gido, S. P.; Muller, W. S.; Fossey, S. A.; Kaplan, D. L. *Proc.-Annu. Meet., Electron Microsc. Soc. Am.* **1993**, 1216.
(12) Muller, W. S.; Samuelson, L. A.; Fossey, S. A.; Kaplan, D. L. *Langmuir* **1993**, *9*, 1857.
(13) Muller, W. S.; Samuelson, L. A.; Fossey, S. A.; Kaplan, D. In *Silk Polymers: Materials Science and Biotechnology*; Kaplan, D., Adams, W. W., Farmer, B., Viney, C., Eds.; American Chemical Society: Washington, DC, 1994; p 342.
(14) Kanishi; Sakabe; Hiei. *ISF-85 Proceedings*; Hakone, Japan, 1985; p 134.
(15) Magoshi, J.; Magoshi, Y.; Nakamura, S. In *Silk Polymers: Materials Science and Biotechnology*; Kaplan, D., Adams, W. W., Farmer, B., Viney, C., Eds.; American Chemical Society: Washington, DC, 1994; p 292.
(16) Shimura, K.; Kikuchi, A.; Ohtomo, K.; Katagata, Y.; Hyodo, A. *J. Biochem.* **1976**, *80*, 693.
(17) Manning, R. F.; Gage, L. P. *J. Biol. Chem.* **1978**, *253*, 2044.
(18) Ohshoma, Y.; Suzuki, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5363.
(19) Strydom, D. J.; Haylett, T.; Stead, R. H. *Biochem. Biophys. Res. Commun.* **1977**, *79*, 932.
(20) Traub, W. *J. Mol. Biol.* **1969**, *43*, 479.
(21) Rippon, W. B.; Chen, H. H.; Anderson, J. M.; Walton, A. G. *Biopolymers* **1972**, *11*, 1411.
(22) Rippon, W. B.; Walton, A. G. *J. Am. Chem. Soc.* **1972**, *94*, 4319.
(23) Rippon, W. B.; Lowbridge, J.; Walton, A. G. *Biopolymers* **1977**, 1139.
(24) Brack, A.; Spach, G. *Biopolymers* **1972**, *11*, 563.
(25) David, W. I. F.; Matthewman, J. C. *J. Appl. Crystallogr.* **1985**, *18*, 461.
(26) Greaves, C. *J. Appl. Crystallogr.* **1985**, *18*, 48.
(27) Falls, A. H.; Wellinghoff, S. T.; Talmon, Y.; Thomas, E. L. *J. Mater. Sci.* **1983**, *18*, 2752.
(28) Dowell, L. G.; Rinfret, A. P. *Nature* **1960**, *188*, 1144.
(29) Crick, A.; Rich, F. H. C. *Nature* **1955**, *176*, 780.
(30) Sprague, K. U. *Biochemistry* **1975**, *14*, 925.
(31) Gamo, T.; Inokuchi, T.; Laufer, H. *Insect Biochem.* **1977**, *7*, 285.
(32) Schulz, G. E.; Schirmer, R. H. *Principles of Protein Structure*; Springer-Verlag Inc.: New York, 1979; p 314.
(33) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95.
(34) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *175*, 593.
(35) Rich, A.; Crick, F. H. C. *Nature* **1955**, *176*, 915.
(36) Cowan, P. M.; McGavin, S. *Nature* **1955**, *178*, 501.
(37) Lotz, B.; Keith, H. D. *J. Mol. Biol.* **1971**, *61*, 195.
(38) Shibnev, V. A.; Khalikov, S. K.; M. P, F.; Poroshin, K. T. *Izv. Akad. Nauk SSSR, Ser. Khim.* **1970**, *1970*, 2822.
(39) Shibnev, V. A.; Chuvaeva, T. P.; Khalikov, S. K.; Poroshin, K. T. *Izv. Akad. Nauk SSSR, Ser Khim.* **1970**, *1970*, 2566.
(40) Tormo, J.; Puiggali, J.; Vives, J.; Fita, I.; Lloveras, J.; Bella, J.; Aymami, J.; Subirana, J. *Biopolymers* **1992**, *32*, 643.
(41) Ramachandran, G. N.; Sasisekharan, V.; Ramakrishnan, C. *Biochim. Biophys. Acta* **1966**, *112*, 168.
(42) Ramachandran, G. N.; Ramakrishnan, C.; Venkatachalam, C. M. In *Conformation of Biopolymers*, 1st ed.; Ramachandran, G. N., Ed.; Academic Press: New York, New York, 1967; Vol. 2, pp 785.
(43) Atkins, T. A. In *Polymer Characterization*; Booth, C., Price, C., Eds.; Pergamon Press: Elmsford, NY, 1989; Vol. 1, p 613.
(44) IUCr. *International Tables for Crystallography*; The Kynoch Press: Birmingham, England, 1962; Vol. III.
(45) Lotz, B.; Keith, H. D. *J. Mol. Biol.* **1971**, *61*, 201.
(46) Marsh, R. E.; Corey, R. B.; Pauling, L. *Biochim. Biophys. Acta* **1955**, *16*, 1.
(47) Taylor, R.; Kennard, O.; Versichel, W. *Acta Crystallogr.* **1984**, *B40*, 280.
(48) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1994; p 570.
(49) Baker, E. N.; Hubbard, R. E. *Progr. Biophys. Mol. Biol.* **1984**, *44*, 97.

MA9517759