J. Philip Anderson
Macromolecular Science and Engineering Center, H. H. Dow Building, College of Engineering, The University of Michigan, Ann Arbor, MI 48109-2136

Received 24 February 1997; accepted 3 October 1997

Morphology and Crystal Structure of a Recombinant Silk-Like Molecule, SLP4

Abstract: *Morphology and crystal structure of a recombinant silk-like molecule, SLP4, were studied. Wide angle x-ray scattering (WAXS) and electron diffraction revealed that SLP4 lyophilized powder and thin films were isomorphic with the silk I crystal structure. Transmission electron microscopy of SLP4 thin films demonstrated a morphology of flat, variable width, crystallites that may aggregate in an epitaxial manner. Theoretical diffraction patterns from silk I crystal structure models were critically compared with SLP4 WAXS data. The analysis concluded that while the crankshaft model is capable of describing details of the SLP4 structural data well, the out-of-register model does not explain the experimental results. In particular, the predicted intensities of the crystallographic reflections for the out-of-register model are inconsistent with the SLP4 WAXS data.* © *1998 John Wiley & Sons, Inc. Biopoly* ***45:*** *307–321, 1998*

Keywords: *morphology; crystal structure; recombinant silk-like molecule; silk I; wide angle x-ray scattering; electron diffraction*

INTRODUCTION

Polymeric materials derived from naturally occurring proteins have been created by the recombinant tools of molecular biology. One such protein, SLP4, is based solely upon a repetitive sequence found in the crystalline segment of *Bombyx mori* silk.[1] It is important to note that other *B. mori* sequences, not present in SLP4, may significantly influence the development of structure in silks, such as liquid crystals, which are critical to the natural fiber processing scheme. Previous work with other silk-like polymers has shown that these materials may assume a crystal structure referred to as silk I.[2,3] Silk I is a polymorph that appears in *B. mori* silk after certain processing conditions, such as the quiescent drying of silk glands.

While the natural mechanism of processing silk proteins has been of considerable interest, many structural aspects remain unclear. In particular, there is controversy about the crystal structure of silk I and its relationship to the natural mechanism of silk processing. The presence of this alternative crystal structure for *B. mori* silk proteins was first reported by Shimizu.[4] It was found that if the contents of the silk gland were allowed to dry without being mechanically disturbed, the resulting crystalline material yielded x-ray diffraction data that was inconsistent with the patterns obtained from fibrous silk. Kratky also found evidence for this polymorph by

Correspondence to: J. Philip Anderson
Contract grant sponsor: Protein Polymer Technologies, Inc.
Biopolymers, Vol. 45, 307–321 (1998)

CCC 0006-3525/98/040307-15

x-ray diffraction and designated it silk I.[5] Silk I has been previously referred to as α-silk.[4] The silk I notation is used here to avoid confusion with α-helical secondary structures.

Silk I can be obtained from unmodified natural silks or the chymotrypsin cleavage fragment of *B. mori* fibroin dissolved in aqueous lithium bromide, lithium sulfocyanide or cupriethylene diamine solutions and then dried.[6] Silk I can be reversibly disrupted by exposure to formic acid and with a small amount of mechanical deformation transformed to the well known extended β-sheet form associated with silk II.[7] Oriented samples are necessary to determine a unique crystal structure for silk I. However, such samples have been difficult to obtain as the silk II form is induced with mechanical alignment methods. Like other structural proteins, silks can be partially characterized by a repetitious primary structure. The crystalline segment of *B. mori* fibroin is well represented by a hexameric sequence, $(AGAGSG)_n$.[6] Synthetic polypeptides have been previously used as a model for silk I. When dialyzed from an aqueous LiBr solution, poly(AG) develops a crystal structure isomorphous with silk I and is referred to as poly(AG) II.[6]

Since the initial work by Shimizu and Kratky, information about the structure of silk I has been compiled by a variety of authors. These experiments have included wide angle x-ray scattering (WAXS), selected area electron diffraction (SAED), ir spectroscopy, and nmr. There have also been molecular simulations of candidate structures for silk I. Here, the salient features of these investigations are reviewed.

Wide Angle X-Ray Scattering

Although macroscopically orienting silk I samples has not been very successful, experimental data by WAXS has been generated by a variety of groups.[4,5,7] All of the silk I WAXS data has been consistent with that found originally by Shimizu and Kratky. Lotz and Keith studied the structure of poly(AG) II crystals by WAXS and found a crystal structure isomorphous with silk I. The WAXS patterns have several characteristic reflections including strong spacings near 0.45 and 0.72 nanometers (nm). The 0.72 nm peak occurs in a region of scattering space well removed from peaks found in the silk II structure. Therefore, any successful model for silk I must have a well-defined set of dense crystallographic planes at a spacing near 0.72 nm.

Selected Area Electron Diffraction

SAED has proven to be valuable in gaining detailed diffraction patterns from small crystals. Lotz and Keith studied the structure of poly(AG) II crystals by SAED.[6] They found a crystal structure isomorphous with silk I. This work led to the development of the crankshaft model, which is discussed in the following section.

IR Spectroscopy

Ambrose and colleagues identified a carbonyl stretching mode for a silk I sample at 1660 cm^{-1} that is suggestive of an α-helix.[8] Hayakawa and colleagues associated silk I with absorption frequencies at 1235 and 650 cm^{-1}.[9] These same bands had been assigned to amorphous silk by Miyazawa and colleagues.[10,11] Later Magoshi and colleagues supported those new findings and also associated the absorption band of 610 cm^{-1} with silk I.[12] More recent work on samples that are silk I by WAXS lack the absorption frequency at 610 cm^{-1}.[7] Due to these inconsistencies, the reliability of ir for silk I detection has been called into question.[7,13] However, ir dichroism experiments have indicated that hydrogen bonding occurs perpendicular to the molecular orientation axis.[14]

Nuclear Magnetic Resonance

Studies of the silk I chain conformation have been conducted by solid state ^{13}C-nmr cross-polarization–magic angle spinning (CP-MAS).[15] CP-MAS data collected from silk samples, supposedly verified as silk I by ir analysis, and poly(AG) II contained an alanine carbonyl signal at 177.1 ppm. It was argued that the crankshaft model failed to account for this feature. However, interpretation of ^{13}C-nmr CP-MAS data relies on comparison to accepted structural models.[16] There are no such model systems for silk I. Furthermore, the effect of hydrogen bonding perpendicular to the chain axis, as in sheet models for silk I, was not taken into account when the chemical shift displacement was derived from the structure.[15] Silk I and amorphous forms had the same chemical shifts and were only distinguishable by the extent of line broadening.[7,13] More recent data from Asakura and colleagues appears to clearly demonstrate the ability of CP-MAS nmr to distinguish between the silk I and silk II forms.[17] Although these researchers argue that an interpretation of data based on a comparison to globular proteins does not support the crankshaft model, other

interpretations of their nmr data are indeed reasonable. Apparently, the CP-MAS nmr data is still inconclusive.

Molecular Simulations

Recently a group simulated different configurations of silk I by modeling stacked sheets of poly(AG) using the Empirical Conformational Energy Program for Peptides.[18–21] Different starting conditions lead to either the out-of-register model or the crankshaft model. Ichimura and Okuyama found a poly(AG) structure consistent with the crankshaft model for silk I from a variety of starting conditions.[22] Others have modeled isolated chains of poly(AG) from helical starting conditions.[23]

SILK I CRYSTAL STRUCTURE MODELS

To facilitate direct comparison between models, unit cells here have been indexed similarly so that a, b, and c are the hydrogen bonding direction, sheet stacking direction, and chain direction, respectively.

Loose-Helix Model[15,23,24]

Orthorhombic cell, $a = 0.459$, $b = 0.720$, $c = 0.908$ nm.

Silk fibroin from *B. mori* was hydrolyzed with chymotrypsin, dissolved in cupriethylenediamine, and then dialyzed to produce silk I.[24] Although a clear fiber pattern was not obtained from x-ray diffraction, a model with one polypeptide chain per unit cell was proposed within which the molecule assumes a fourfold helical conformation with a residue translation of 0.227 nm. It has not been made clear how this single fourfold helix packs into an orthorhombic unit cell.

Comparison of the silk I fiber pattern with similar data from poly(AG) II showed that both samples had the same crystal structure.[25] The 0.227 nm reflection was indexed originally as the (004); however, better SAED data later found that this same reflection could be resolved into two arcs and therefore was off-meridional.[6] Hence, this reflection should be indexed as the (014) instead, which places the fiber repeat distance near 0.96 nm, significantly greater than the 0.91 nm proposed in the loose-helix model.

The conformation of the silk I chain was characterized by ^{13}C-nmr of silk fibroin in the solid state by CP-MAS.[15] Apparently, a loose-helix model with $(\varphi, \psi) = (-26°, 135°)$ is consistent with the nmr data even though these values deviate from low energy regions present on the Ramachandran conformational map. The samples used in this ^{13}C-nmr study were supposedly verified as silk I by ir. However, this same group later asserted that it is impossible to differentiate between silk I and amorphous silk with ir and by ^{13}C-nmr CP-MAS only in terms of line broadening.[7,13]

Crankshaft Model[6,25]

Orthorhombic unit cell, $a = 0.472$, $b = 1.44$, $c = 0.96$ nm.

From SAED and WAXS experiments on poly(AG) II, Lotz and Keith proposed the crankshaft model for silk I based on a unit cell with the space group $P2_12_12_1$. In their model for silk I, the polypeptide backbone is folded into a crankshaft conformation, yielding a hydrogen-bonded sheet form. Two types of chain conformation were proposed. Chain conformation type I represents the sequence poly(βAα_LG) [βA$(\varphi, \psi) = (-123°, 122°)$, α_LG$(\varphi, \psi) = (57°, 72°)$] while type II is poly($\beta$A$\alpha_D$G) [$\alpha_DG(\varphi, \psi) = (72°, 57°)$]. Antiparallel models based on either strictly conformation type I or II do not saturate the hydrogen-bonding potential of the structure and are of higher energy.

The crankshaft model is formed by sheets composed of antiparallel chains that regularly alternate in conformation between types I and II. As in the extended sheet forms, hydrogen-bonded sheets can stack to fill space. The face of the resulting solid, perpendicular to the chains, may be composed of regular hairpin turns, where the chains switch conformation type and direction, then proceed back through the crystal. This base crankshaft model was consistent with the diffraction data available to Lotz and Keith, except that certain (0kl) reflections were too intense, particularly on the first and second layer lines. It was proposed that this difficulty could be removed by using a model in which there is an alternative method of sheet stacking, described by an intersheet [001] axis offset of half the unit cell length [for poly(AG)].[6] Sheet packing that randomly incorporates this offset disrupts much of the periodicity responsible for the anomalous reflection intensities.

The data generated from ^{13}C-nmr of silk fibroin in the solid state by CP-MAS appears to contain an alanine carbonyl signal at 177.1 ppm.[15] This is claimed to be unaccounted for by the crankshaft model. However, it is conceded that the hydrogen-bonding perpendicular to the fiber axis proposed in

the crankshaft model and experimentally observed by ir dichroism might account for this discrepancy. In light of the fact that silk I and amorphous silk give rise to the same ^{13}C-nmr CP-MAS chemical shifts, conclusions from this data must be drawn cautiously.[7,13]

Out-of-Register Model[18]

Orthorhombic unit cell, $a = 0.894$, $b = 1.126$, $c = 0.646$ nm.

The out-of-register model was developed using the Empirical Conformational Energy Program for Peptides, which included an analytical expression for the first derivative of the potential energy with respect to rigid body variables, on an IBM 3060-600E computer.[19–21] Energy minimizations were carried out with the Secant Unconstrained Minimization Solver until a minimum delta energy of 10^{-3} kcal/mol was achieved.[26] Some 80% of the computations were performed in vector mode and sometimes run in parallel.

Three sheets, each composed of five antiparallel strands, six residues long [$(AG)_3$], were stacked. Strands projected side chains either strictly above or below the sheet plane in a regular alternating manner. Methyl groups of adjacent strands were on opposites sides of the sheet. Alanine residues were aligned with neighbor chain glycine residues and thus ''out-of-register.'' The conformational energy of the stack was minimized from eleven starting values of dihedral angles, in the range from 80° to 180° at 10° intervals, such that all residues initially had the same angles and $\varphi = -\psi$. All chains were allowed to move freely and the origin of space-fixed coordinates was set at the centroid of all atomic coordinates.

An additional starting point was added [$A(\varphi, \psi) = (-80°, 150°)$, $G(\varphi, \psi) = (-150°, 80°)$] after some minimizations converged near this conformation. This starting point yielded a conformational state with a calculated energy of -394 kcal/mol, referred to here as the out-of-register model. Other starting conditions lead to the crankshaft model, which was found to have a calculated energy of -365 kcal/mol. While the positions of certain reflections in the theoretical scattering pattern of the out-of-register model were compared with experimental data, the predicted diffraction intensities were not examined.

Asakura and colleagues have developed some experimental support for the out-of-register model from their CP-MAS nmr data.[17] However, their interpretation of the nmr data relies on a comparison to globular proteins that may present a significant error in this application to fibrous silk. Furthermore, for that nmr data to support the out-of-register model, a relatively large error of ± 1.5 ppm, nearly twice the standard deviation, was assumed.[17] A more detailed analysis of the nmr data requires considerations beyond the scope of this paper and thus further discussion here will focus on diffractional techniques.

EXPERIMENTAL

SLP4 was designed, manufactured, and supplied by Protein Polymer Technologies, Inc. (San Diego, CA) in a lyophilized form. The molecule consists of 1101 amino acids, corresponding to a theoretical molecular weight of 76 k daltons, and is essentially $(GAGAGS)_{168}$ with small head and tail features (totaling approximately 8% of the amino acids).[1] SLP4 was subject to transmission WAXS. A Rigaku Rotoflex θ-θ x-ray system equipped with a rotating anode and a 12 kW Cu Kα tube that emits x-ray radiation at a wavelength of 0.154 nm was used. The instrument has a curved crystal graphite monochromator and NaI scintillation type detector. A detailed WAXS scan was conducted at a rate of 0.1° per minute with sampling at 0.01° intervals, while operating at 100 mA and 40 kV. Pertinent data was obtained between 5° and 30° 2θ, in counts per second. Percent crystallinity was estimated by isolating the sharp peaks, clearly associated with material arraigned on a lattice, from the amorphous fraction.

Transmission electron microscopy (TEM) samples were prepared by dissolving 1 mg of SLP4 in 1 mL of formic acid and then adding 1 mL of deionized water to the solution. This solution was sprayed onto amorphous carbon coated mica and then immediately transferred to dry under vacuum. The amorphous carbon-coated mica, supporting SLP4 crystallites, was floated onto deionized water and collected with 300 mesh copper TEM grids. These SLP4 samples were imaged with a Philips 420 TEM at a magnification of 10.5 kX. Electron diffraction was performed on samples gold shadowed with a Denton evaporator (this is the same technique incorrectly described as SAED in the previous report).[3]

Poly(AG) models for silk I with published atomic coordinates were input into a molecular simulation software package, Cerius^2 1.5 (Molecular Simulations Inc.). These structures were then refined to energy minima while constrained to the proposed unit cells. The force field used for energy minimizations was Dreiding 2.21.[27] Lattice and coordinate energy minimizations were conducted using the Conjugate Gradient 200 algorithm in periodic boundary conditions and allowed to converge with a maximum step size of 0.2 nm, a maximum number of 10 substeps, and 10 atom steps per cell deformation.[28] These computations were terminated when the residual

rms force dropped below 1.0 kcal/mol/nm. After the structures were minimized, theoretical diffraction data were calculated with Cerius[2] for comparison with experimental WAXS patterns on SLP4. Pertinent information was collected between 5° and 30° 2θ.

$$F_{hkl} = \sum f_n \exp[2\pi i(hx_n + ky_n + lz_n)]$$

The diffracted intensities were calculated according to the above expression, which gives the form factor for the (hkl) reflection as the sum over the n different atoms in the unit cell at fractional coordinates (x_n, y_n, and z_n).[29] The form factor for atom n (f_n) is a function of the atomic number and the scattering angle. This expression is essentially the discrete Fourier transform of the unit cell contents evaluated at the positions of the reciprocal lattice. The diffracted intensity is given by the square of F_{hkl}. The positions of diffracted peaks depend on the size and characteristic angles of the unit cell, while the relative intensity of the different reflections are related to the distribution and symmetry of the molecules within the cell. Analysis of the expected diffraction intensities and comparison to experiment is a powerful means to determine the validity of a proposed structural model.

Loose-Helix Model

Although dihedral angles have been proposed for a loose helix model, they are based on data independent of packing into any unit cell.[15] It is not clear how the loose-helix is expected to fit into an orthorhombic unit cell.[15,24] This model was not sufficiently developed to be critically analyzed. Furthermore, ir dichroism and diffraction data are consistent with sheet-type structures.[6,14,25] Therefore, further attention here is given to the crankshaft and out-of-register models.

Crankshaft Model

Published atomic coordinates for poly(AG) in the chain I crankshaft conformation from Lotz and Keith were input into Cerius[2].[25] Chain II was generated by using the coordinates provided for chain I, performing a mirror plane transformation along the c axis, and then switching the positions of the carbonyl and amino groups. The antiparallel crankshaft model is of hydrogen-bonded sheets composed by regularly alternating chains I and II. However, previous work has shown that this base crankshaft model does not fully described the silk I crystal structure and gives rise to discrepancies with respect to experimental data in terms of some simulated $(0kl)$ diffraction intensities. As suggested by Lotz and colleagues, it is possible to consider an alternative method of sheet stacking, distinguished from the base crankshaft model by neighbor sheet translation of $+c/2$.[6] This alternative model, referred to here as the shifted crankshaft model, was also generated. These crankshaft models were minimized in unit cells constrained to the dimensions reported by Lotz and Keith (a = 0.472, b = 1.44, c = 0.96 nm). A simulated WAXS pattern was generated for a theoretical crankshaft model that randomly incorporated base and shifted stacking options, in equal probability, by using the Cerius[2] diffraction-faulted module which accesses the DIFFaX program.[30] The DIFFaX simulation was conducted with 2000 layers, using variable peak broadening via a Lorentzian function (u, v, and w were set to 0.0, −1.0, and 1.3, respectively).

Out-of-Register Model

The published out-of-register structure file, *pdb1slk.ent*, was obtained from the Protein Data Bank at the Brookhaven National Laboratory. The downloaded model for silk I consisted of a three-sheet stack. Each sheet was five chains by six residues long, which approximated a three-dimensionally periodic crystalline system. After loading the file into Cerius[2], a periodic unit cell was constructed from the repeating unit in the middle of the second sheet by eliminating the extra units. The resulting structure was energy minimized while constraining the dimensions of the unit cell to that given by the authors (a = 0.894, b = 1.126, c = 0.646 nm). Diffraction spacings and intensities for the out-of-register model were calculated in the Cerius[2] diffraction-crystal module. WAXS was simulated with Cu Kα radiation (λ = 0.145 nm), crystallite sizes of 5 nm in the a, b, and c directions, a polarization fraction of 0.5, 20 points per degree, no lattice strain, and no Debye–Waller factors.

RESULTS

The WAXS diffractometer data from SLP4 is shown as Figure 1. The SLP4 powder was estimated to be 29% crystalline. Several intense reflections are observable, most notably at 12.25° and 20.20° (2θ), which correspond to spacings at 0.725 and 0.448 nm. Using a Sherrer analysis, the peaks corresponding to spacings of 0.725 and 0.448 nm were found to arise from crystals with effective sizes of 18.1 and 29.2 nm, respectively. SLP4 WAXS spacings and relative intensities are listed in Table I and are very similar to previously published WAXS data from *B. mori* fibroin assigned the Silk I designation.[9]

Figure 2 illustrates crankshaft models projected in the chain direction, [001]. The [001] projection is oriented so that the hydrogen-bonding direction, [100], is horizontal and the sheet-to-sheet direction, [010], is vertical. The SLP4 WAXS reflection at 0.725 nm corresponds to the (020) crystallographic planes in the crankshaft model. It is clear that these planes are heavily populated and are likely to diffract intensely. Figures 3 and 4 show the [100] pro-

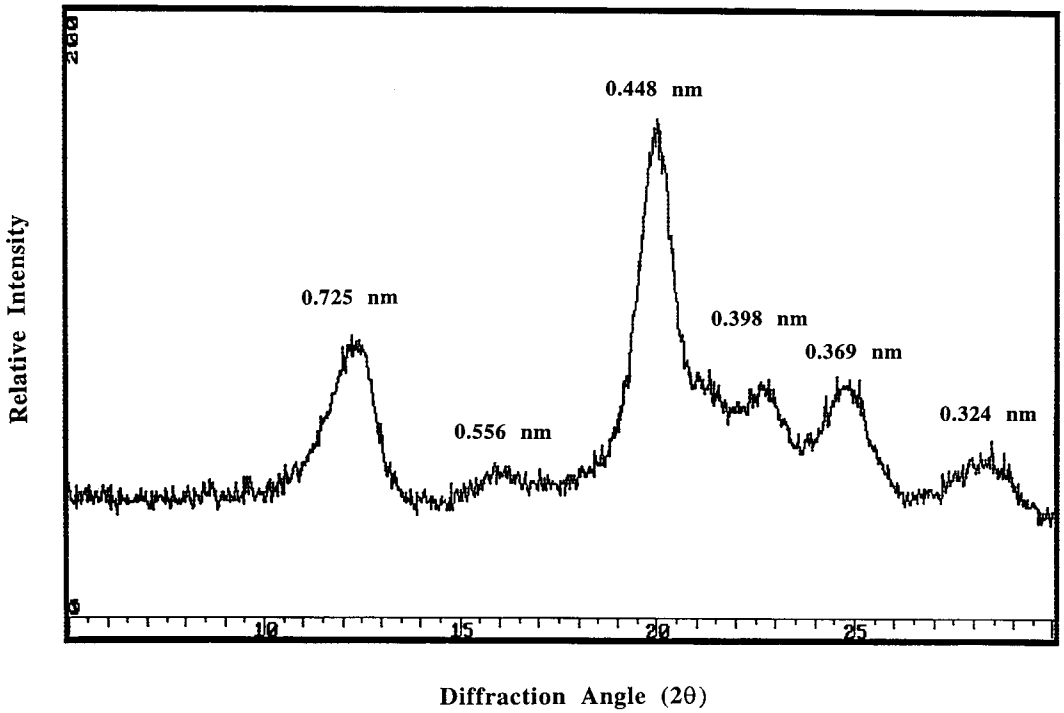

FIGURE 1 SLP4 lyophilized powder transmission WAXS.

jections of the base and shifted crankshaft models, respectively. These [100] projections are oriented so that the chain direction, [001], is horizontal and the sheet stacking direction, [010], is again vertical. The difference in sheet stacking between the base and shifted models in Figures 3 and 4, respectively, is readily apparent by observing relative locations of alanine methyl groups.

Figure 5 illustrates the out-of-register model projected in the chain direction, [001]. Again, the [001] projection is oriented so that the hydrogen-bonding direction, [100], is horizontal and the sheet stacking direction, [010], is vertical. Figure 6 shows the [100] projection of the out-of-register model. Again, this [100] projection is oriented so that the chain direction, [001], is horizontal and the sheet stacking direction, [010], is vertical. It is immediately evident that the conformation of the chains in the out-of-register model is very much different than in the crankshaft model. In specific, chains are sig-

Table I Comparison of Experimental Wide Angle X-Ray Scattering (WAXS) Data from SLP4 with Predicted Scattering from the Crankshaft and Out-of-Register Models for Poly(AG)[a]

SLP4 WAXS		Crankshaft Model			Out-of-Register Model	
Spacing (nm)	Relative Intensity	Spacing	Relative Intensity	Plane	Spacing	Relative Intensity
0.725	57	0.729	46	(020)	0.700	15
0.556	32	0.576	44	(021)	0.563	100
0.448	100	0.449	100	(110)	0.447	46
0.398	46	0.399	51	(120)	0.383	36
0.369	48	0.360	30	(040)	0.367	61
0.324	34	0.328	35	(013)	0.323	37

The crankshaft model shows agreement not only with the observed spacings but with the intensities as well. While certain reflections from the out-of-register model are observed experimentally, the calculated intensities of scattering do not match the experimental data. Further hindering the fit of the out-of-register model are several intense reflections expected to be present but not observed by experiment.

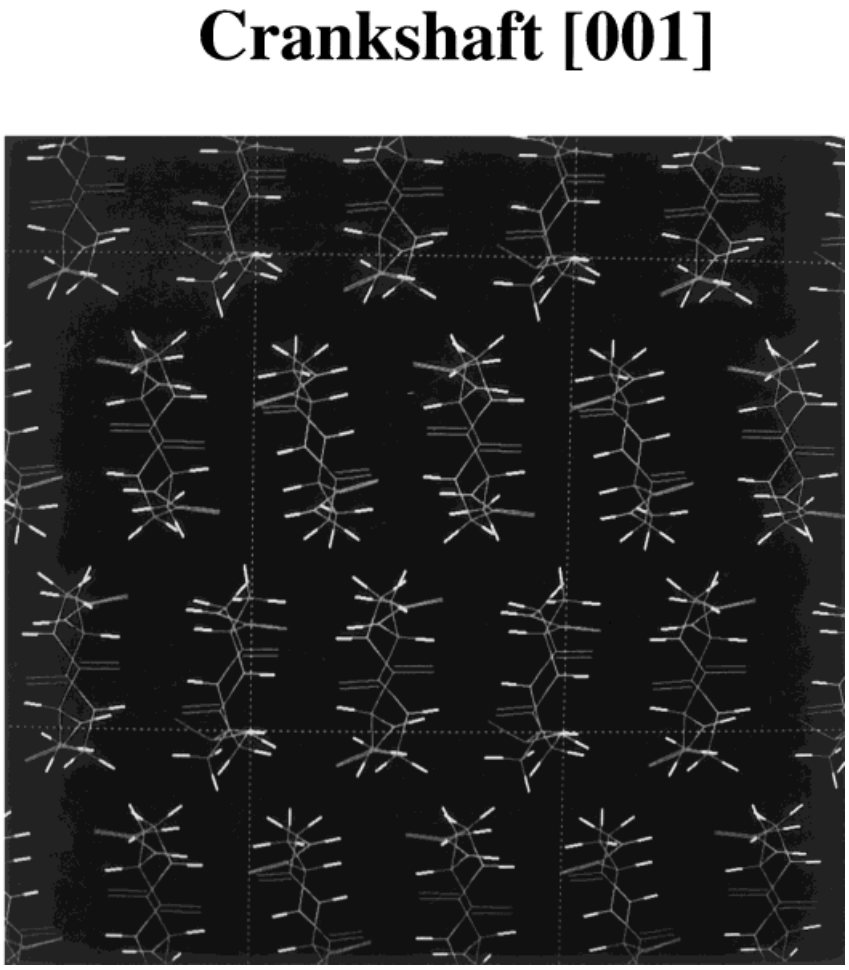

FIGURE 2 The crankshaft models projected in the chain direction, [001]. The hydrogen-bonding direction, [100], is horizontal and the sheet stacking direction, [010], is vertical.

nificantly more extended in the out-of-register model. Furthermore, unlike the crankshaft model, the out-of-register model lacks a readily apparent plane that repeats every 0.725 nm through the crystal.

Energy decompositions for the base and shifted crankshaft models, as well as the out-of-register model, are provided in Table II. It is clearly evident that hydrogen bonding provides a significant stabilizing force in all the models considered. The base and shifted crankshaft models are of similar energy in this simulation, while the out-of-register model is observed to be slightly lower in energy. However, it is not clear that this distinction is significant.

Figure 7 compares the DIFFaX theoretical diffraction pattern from the crankshaft model, composed of equally probable randomly alternating base and shifted sheet stacking, with the SLP4 WAXS data. It is clear from Figure 7 that this crankshaft model predicts not only the spacings but also the intensities of the SLP4 WAXS data. However, the spacing at 0.576 nm appear to be slightly more intense than the experimental data. Figure 8 compares the theoretical diffraction pattern from the out-of-register model with the SLP4 WAXS data. It is clear from Figure 8 that the out-of-register model does not describe the crystal structure of the SLP4 lyophilized powder.

Figure 9 is a TEM image of an SLP4 film edge. Whiskers of variable width and length are visible components of the organized film. These flat crystallites are clearly longer then wide. Some of the wider individual whiskers appear to fray at ends. SLP4 whiskers are observed to aggregate into sheaf-type structures with parallel or perpendicular orientation. Such perpendicular orientation, as observable in five distinct locations on a whisker aggregate shown in Figure 10, suggests epitaxial alignment

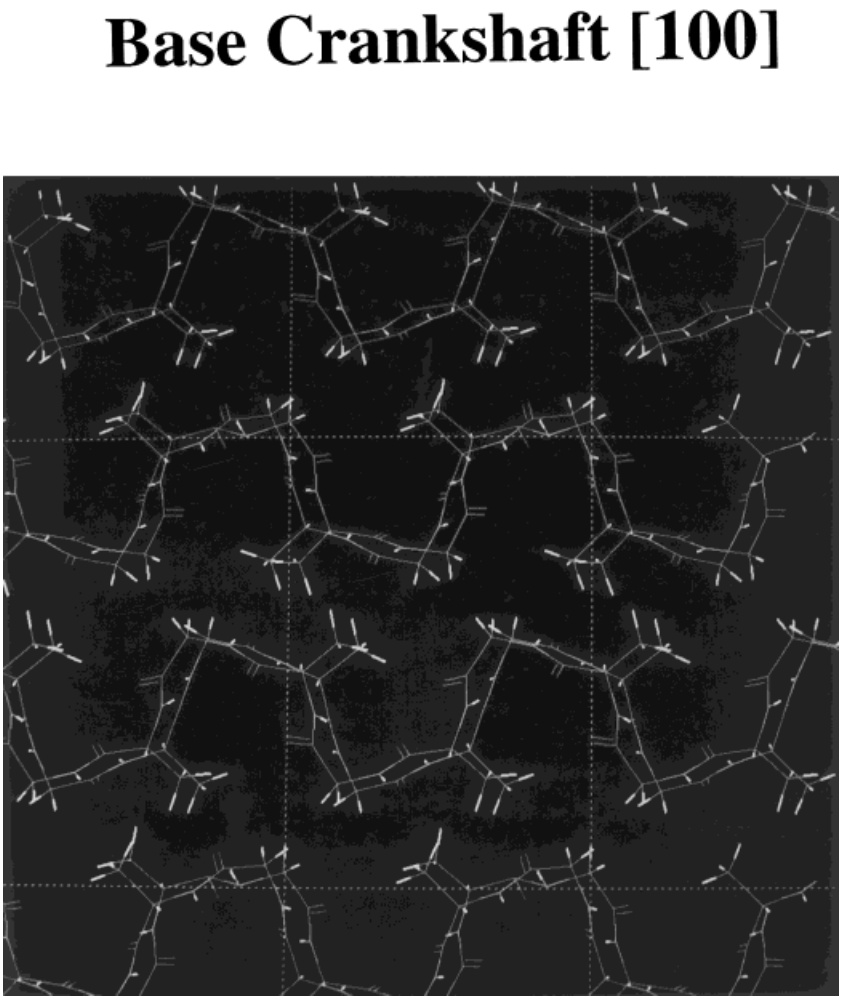

FIGURE 3 The [100] projections of the base crankshaft model. The chain direction, [001], is horizontal and the sheet stacking direction, [010], is vertical.

between SLP4 whiskers. Some whiskers appear to have density fluctuations parallel to the long axis. This phenomenon is particularly noticeable in whiskers which fray in Figure 10. Figure 11 is the SLP4 electron diffraction pattern that exhibits reflections at 0.44, 0.39, and 0.36 nm. While these spacings and their intensities are consistent with the silk I crystal structure, the electron diffraction pattern lacks a low angle spacing near 0.725 nm region, which is apparently masked by the inelastic electron scattering.

DISCUSSION

SLP4 appears to have a structure related to a previously studied silk-like material, SLPF.[3] Both SLPF and SLP4 assume the silk I crystal structure in the lyophilized format by WAXS and in thin films cast from formic acid based solutions by electron diffraction. The expected spacing near 0.725 nm appears to be masked by inelastic electron scattering in the SLP4 electron diffraction pattern (Figure 11). However, the previously generated SLPF electron diffraction pattern was found to contain a spacing at 0.74 nm by a tedious microdensitometer analysis. Furthermore, it is possible to distinguish between crystals in the silk I and silk II polymorphs solely on the basis of morphology. As reported by Lotz, Magoshi, and colleagues, silk crystals in the silk II form are systematically twisted, while silk I crystals are always flat.[31] The crystals shown in Figure 9 and 10 are flat. Therefore, these SLP4 crystals are in the silk I conformation by electron diffraction and morphological observation.

In Figure 2, the [001] projection of the crankshaft

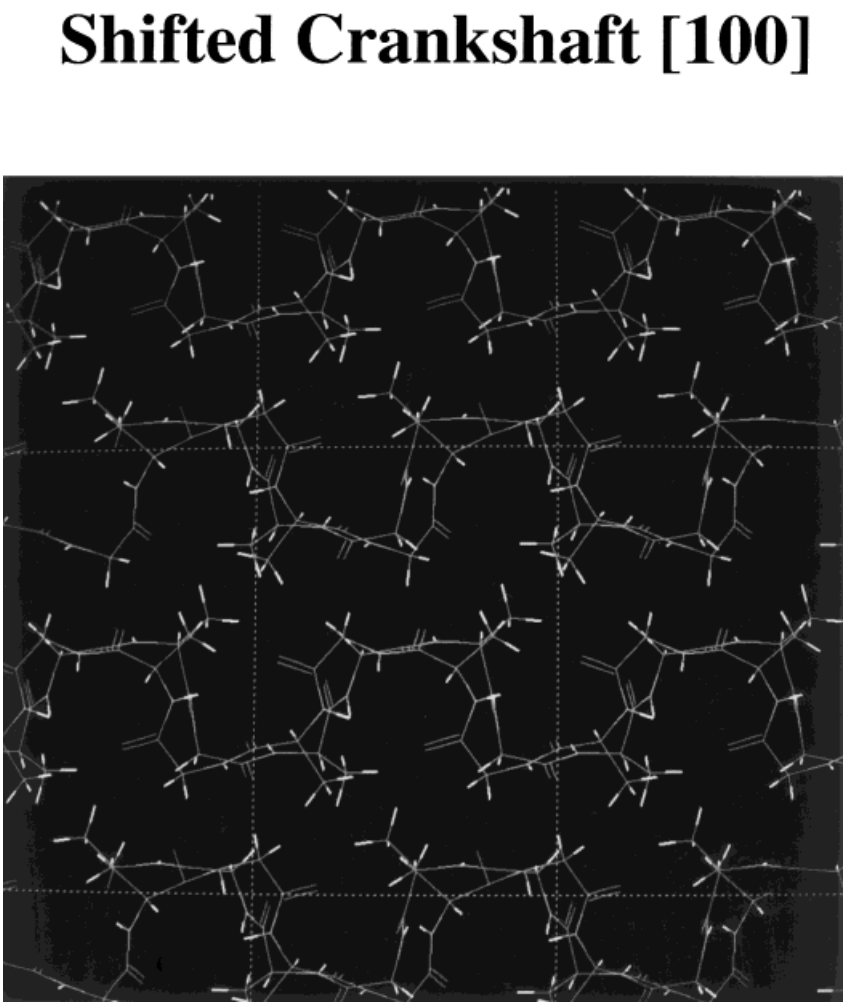

FIGURE 4 The [100] projections of the shifted crankshaft model. The chain direction, [001], is horizontal and the sheet stacking direction, [010], is vertical.

model, it is clear that the sheets are 0.729 nm apart. It is no surprise that the spacing from these dense planes reflects strongly in theoretical diffraction patterns. From the energy decomposition presented in Table II, it is apparent that [001] hydrogen bonding is a significant stabilizing force. This is consistent with ir dichroism experiments that have indicated that hydrogen bonding occurs perpendicular to the molecular orientation axis.[14] The base and shifted crankshaft models are of similar energy, also shown in Table II. This fact justifies their designation of being equally probable in the DIFFaX simulated diffraction pattern. However, comparison of the DIFFaX generated crankshaft diffraction pattern with SLP4 WAXS data, shown as Figure 6, reveals that certain (0*kl*) reflections in the crankshaft model reflect too intensely. This indicates that the crankshaft model considered here is not perfect. More disorder in sheet stacking, either by limited positional variability within the two alternative stacking options considered here (lattice strain) or additional well-defined sheet stacking alternatives, may resolve this issue.[6,24] However, sheet stacking alternatives, other than those represented by the base or shifted models, were not stable in the simulations conducted.

In Figures 5 and 6, the [001] and [100] projections of the out-of-register models, there are no blatantly dense planes that repeat at intervals near 0.725 nm through the crystal. Therefore, as demonstrated in Figure 8, the theoretical diffraction pattern from the out-of-register model does not display an intense spacing near the 0.725 nm spacing of the SLP4 WAXS pattern. Furthermore, the theoretical diffraction pattern from the out-of-register model is grossly inconsistent in terms of both spacings and

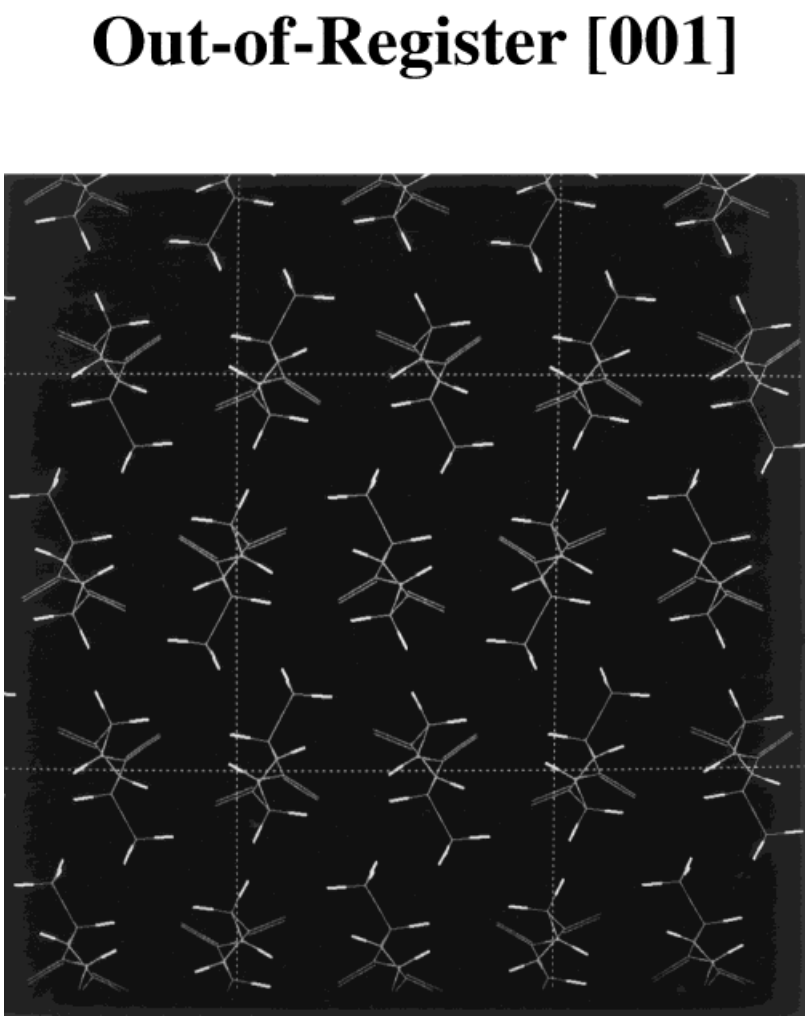

FIGURE 5 The out-of-register model projected in the [001], chain direction. The hydrogen-bonding direction, [100], is horizontal and the sheet stacking direction, [010], is vertical.

intensities with respect to the SLP4 WAXS pattern. Polypeptide chains in the out-of-register configuration are significantly more extended then those in the crankshaft conformation. If the SLPF whiskers previously studied were in the out-of-register conformation, one would expect them to be noticeably wider than the observed width, which agrees with the predictions of the crankshaft model.[3]

A comparison between the SLP4 WAXS data and the predicted spacings and intensities of the crankshaft and out-of-register models is presented in Table I. The crankshaft model is generally consistent with the experimental WAXS data; the strongest reflection is predicted to be the 0.449 nm spacing and the second strongest is at 0.729 nm. Both of these peaks are dominant in the experimental data. However, the out-of-register model predicts the strongest peak should be at 0.563 nm, which is only weak experimentally. Other peaks that are experimentally intense are predicted to be weak by the out-of-register model. Furthermore, the out-of-register model predicts some intense diffraction peaks that are not present in the data.

$$R = \int |I_p - I_d| \, d2\theta / \int I_d d2\theta$$

In order to quantitatively evaluate the fit of the models to the data, a crystallographic R factor was determined using the SLP4 WAXS data, and theoretical predictions generated with Cerius[2] and the DIFFaX program. The R factor was defined in the equation above where I_p and I_d are the diffracted intensities predicted by the model and determined by the data, respectively. For this analysis the WAXS patterns were converted to quantitative in-

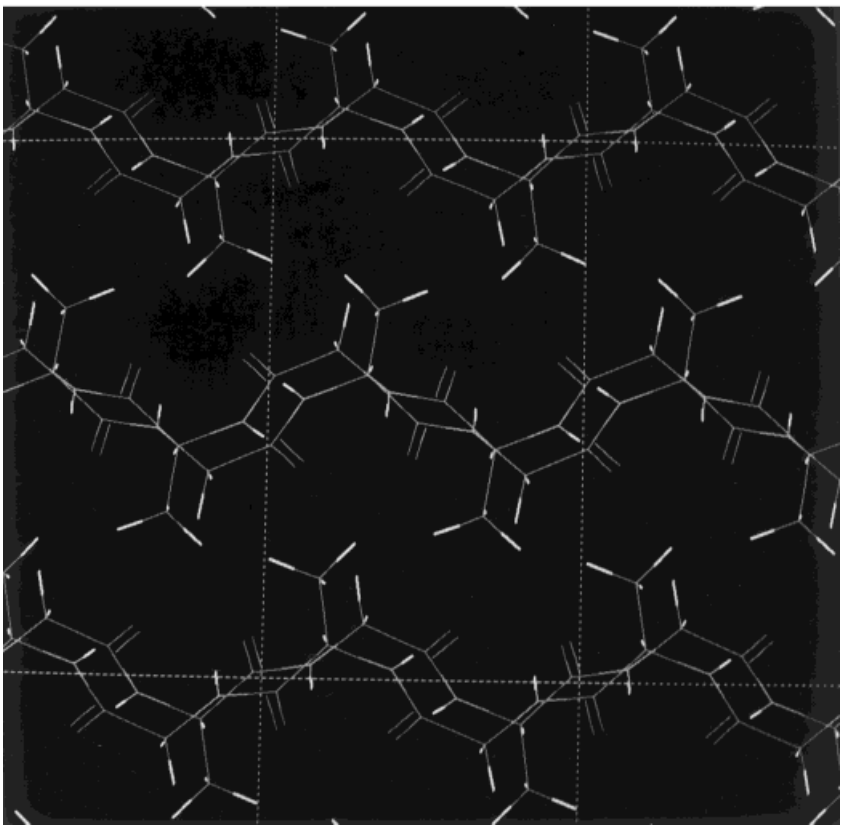

FIGURE 6 The [100] projection of the out-of-register model. The chain direction, [001], is horizontal and the sheet stacking direction, [010], is vertical.

tensities by normalizing the strongest peak to 100, while the integrals above were evaluated from 10° to 30° (2θ). This measurement gives R (crankshaft) = 0.26 and R (out-of-register) = 0.60. Clearly, the crankshaft model is a significantly better fit than the out-of-register model to the available x-ray diffraction data from SLP4 or silk I.

The morphology of SLP4 and SLPF thin films are also similar in that they both appear to be derived from flat whisker crystallites that may aggregate into sheaf structures. Lotz, Magoshi, and colleagues have identified whisker crystallites from natural silk proteins in the silk I and II conformations (also J. Magoshi, unpublished data).[31] Both SLP4 and SLPF whiskers are longer then wide—hence it seems very probable that the strong hydrogen bonding occurs parallel to the long axis. However, SLP4 whiskers appear to be of variable width, while the width of SLPF whiskers was found to be monodispersed. SLPF differs from SLP4 most notably in that the silk-like blocks of SLPF are periodically disrupted by a polar cell binding domain based on a sequence from human blood plasma fibronectin. It was hypothesized that the length of the silk-like block between these fibronectin based domains defined the whisker width in SLPF. Therefore, SLP4, which lacks the fibronectin sequence, is also observed to lack monodispersed whisker width.

However, it is more likely that SLP4 whisker crystals are orientated with chains normal to the carbon film. Synthetic polypeptide crystals have been reported to assume this orientation.[25,32,33] If this is the case, then the width of SLP4 crystals reflects the extent of sheet stacking. As shown in Figure 9, these SLP4 crystallites may fray at ends. Therefore this orientation is consistent with the preferential

Table II Energy Decompositions of Poly(AG) Silk I Models Minimized in Cerius[2] with the Drieding 2.21 Force Field Using Periodic Boundary Conditions

Silk I Models for Poly(AG)	Base Crankshaft Model	Shifted Crankshaft Model	Out-of-Register Model
Total Energy (kcal/mole/unit cell)	32.31	31.94	8.03
Bonds	5.29	7.20	3.77
Angles	21.82	23.40	7.80
Torsions	25.45	21.87	15.02
Inversions	0.87	1.13	0.85
Van der Waals	19.25	1.47	0.34
Hydrogen Bonds	−40.37	−23.13	−19.74

disruption of the weak van der Waals forces that are responsible for sheet stacking, as known to occur in the crankshaft model. As previous work on partially crystallized SLPF by WAXS has demonstrated that sheet formation precedes sheet stacking, it seems much more likely that formed sheets are incompletely stacked than the alternative explanation, which would involve merging sheets in the direction on the chain.[34] Furthermore, SLP4 whiskers are observed to preferentially aggregate by perpendicular organization, shown in Figure 10, which suggests epitaxial alignment. The α-helices have been reported to organize in an epitaxial manner as coiled coils.[35] As illustrated in Figure 12, epitaxial alignment may occur by interdigitation of hairpin turns of overlapping whiskers. It seems unlikely that perpendicular epitaxial whisker alignment would be driven by perpendicular sheet stacking. Furthermore, SLP4 whiskers have not been observed to perpendicularly align themselves.[2,3] Density fluctuations are noticeable in fraying whiskers shown in Figure 10. It is possible that this represents the effect of the SLP4 sequence permitting sheets of variable width (in the chain direction) to be nucleated and that later imperfectly aggregate.

CONCLUSIONS

The spacings and intensity of the SLP4 transmission WAXS pattern are synonymous with the silk I crystal structure. Thin films of SLP4 cast from formic acid were observed to assume the silk I crystal structure by electron diffraction. The morphology of SLP4 is based on a crystalline whisker unit that aggregates into sheath-type structures. The width of SLP4 whiskers is variable and apparently influenced by processing conditions. Given that SLP4 whiskers

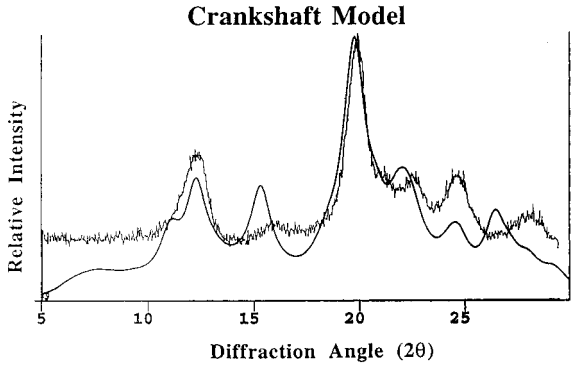

FIGURE 7 The DIFFaX theoretical diffraction pattern from the crankshaft model, composed of equally probable randomly alternating base and shifted sheet stacking, compared with the SLP4 WAXS data.

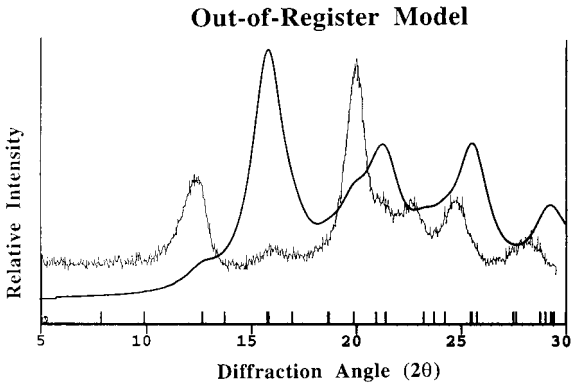

FIGURE 8 The theoretical diffraction pattern from the out-of-register model compared with the SLP4 WAXS data. It is clear that the out-of-register model does not describe the crystal structure of the SLP4 lyophilized powder.

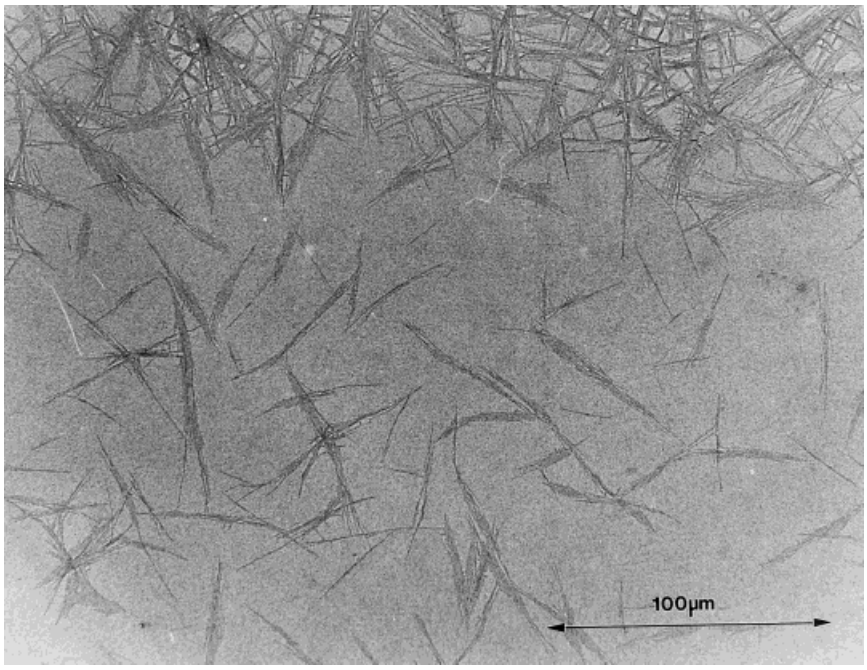

FIGURE 9 TEM image of an SLP4 film edge. Note whiskers of variable width, which may fray at ends or aggregate in a perpendicular manner.

may fray, it is consistent with the crankshaft model and previous work elucidating steps in silk I crystallization that chains are normal to the carbon film plane.[34] Furthermore, SLP4 whiskers are observed to aggregate perpendicularly, suggesting epitaxial alignment that is most likely to be driven by the interdigitation of hairpin turns at chain ends.

The crankshaft model is capable of reproducing spacings and most of the intensities from the WAXS data for SLP4. However, it is clear that the current

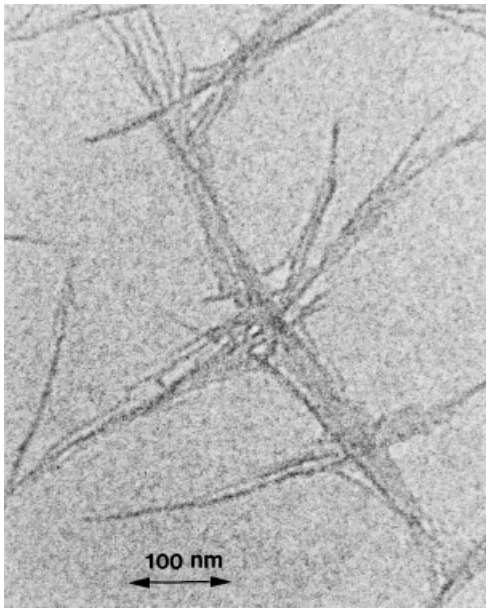

FIGURE 10 TEM image of SLP4 whiskers perpendicularly oriented.

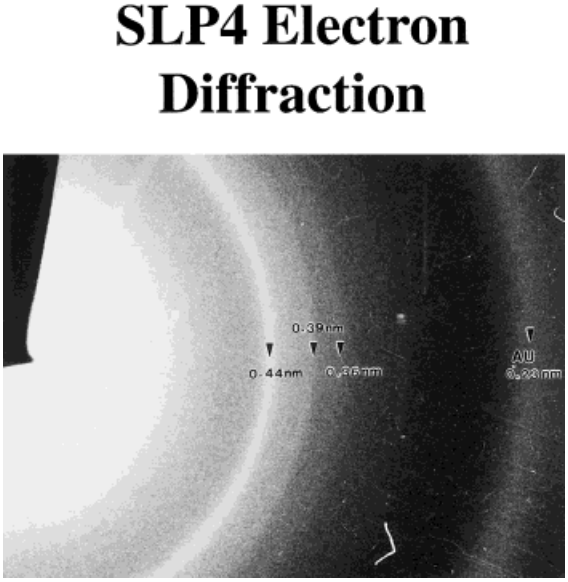

FIGURE 11 SLP4 electron diffraction. The low angle spacing expected near 0.725 nm is apparently masked by the inelastic electron scattering.

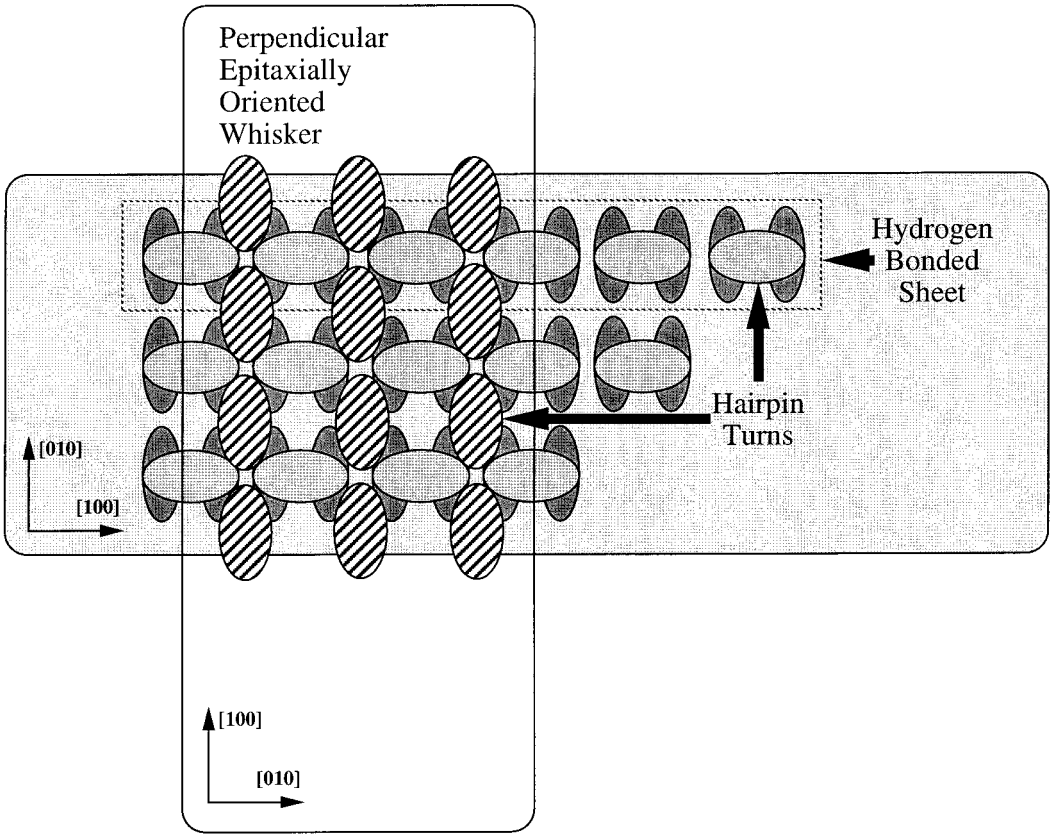

FIGURE 12 Epitaxial model of SLP4 whisker orientation. Chains are normal to the figure plane.

state of the crankshaft model is not perfect and may need to specify more intersheet disorder as observed in SLP4 TEM images. While the out-of-register model demonstrates some reflections that have similar spacings, the intensities of these reflections are inconsistent with that found in experiment. Therefore, it can be concluded that the out-of-register model does not describe the molecular packing of SLP4 or silk I crystals. This result suggests that molecular simulations alone are not yet reliable enough to predict crystal structures. To be effective such models need to be analyzed and critically compared with all of the experimental data available.

Silk fiber formation involves subtle changes in structure during the phase transitions from aqueous solutions to a liquid crystal and finally an insoluble solid.[36,37] It has been suggested that the silk I structure is involved in processing, and perhaps contributes to the formation of the observed liquid crystalline phase.[38,39] The out-of-register model in part stems from a need to better understand the structure of processing intermediates on the pathway to natural silk fiber, which is composed of silk II crystallites. Solid state transformations from silk I to silk II have been readily observed.[36] It is much easier to imagine such a transformation from the out-of-register model to the silk II structure than that for the crankshaft model. This is simply because the out-of-register model is the silk II model with regular disorder. Every other chain in the out-of-register model is rotated 180° and translated one residue, otherwise it is the silk II model. Furthermore, when the out-of-register model unit cell was not dimensionally constrained during energy minimization, the crystal structure was transformed to the silk II model. Therefore, unlike the crankshaft model, the out-of-register model is not a local energy minima for poly(AG).

Magoshi and colleagues have generated some diffractional information from natural silk solutions.[36] Their work demonstrates peaks corresponding to spacings at 0.620 and 0.435 nm. Perhaps the out-of-register model somehow relates to a silk fiber processing intermediate, but it certainly does not describe the solid state structure referred to as silk I.

This work has been supported in part by Protein Polymer Technologies, Inc., of San Diego, CA. The author would like to thank Lynn W. Jelinski (Cornell) and Tetsuo Asakura (Tokyo University of Agriculture and Technology) for stimulating discussions. Research was performed in the laboratory of David C. Martin (University of Michigan), who was coauthor on a related paper previously submitted to *Biopolymers* and who contributed to this work.

REFERENCES

1. Cappello, J., Crissman, J. W., Dorman, M., Mikolajczak, M., Textor, G., Marquet, M. & Ferrari, F. A. (1990) *Mat. Res. Soc. Symp. Proc.* **174,** 267–276.
2. Anderson, J. P. (1997) in *Protein-Based Materials*, Kaplan, D. & McGrath, K., Eds., Birkhuaser, Boston, pp. 369–421.
3. Anderson, J. P., Cappello, J. & Martin, D. C. (1994) *Biopolymers* **34,** 1049–1058.
4. Shimizu, M. (1941) *Bull. Imp. Sericult. Expt. Sta. Japan* **10,** 475.
5. Kratky, O., Schauenstein, E. & Sekora, A. (1950) *Nature* **165,** 319–320.
6. Lotz, B. & Cesari, F. C. (1979) *Biochemie* **61,** 205–214.
7. Asakura, T., Kuzuhara, A., Tabeta, A. & Saito, H. (1985) *Macromolecules* **18,** 1841–1845.
8. Ambrose, E. J., Bamford, C. H., Elliot, A. & Hanby, W. E. (1951) *Nature* **167,** 264–265.
9. Hayakawa, T., Kondo, D., Yamamoto, S. & Noguchi, J. (1970) *Kobunshi Kagaku* **27,** 300.
10. Miyazawa, T., Shimanouchi, T. & Mizushima, S. (1958) *J. Chem Phys.* **29,** 611–616.
11. Miyazawa, T., Masuda, Y. & Fukushima, S. (1962) *J. Polym. Sci.* **62,** S62.
12. Magoshi, J., Mizuida, M., Magoshi, Y., Takahashi, K., Kubo, M. & Nakamura, S. (1979) *J. Polym. Sci. Polym. Phys. Ed.* **17,** 515–520.
13. Ishida, M., Asakura, T., Yokoi, M. & Saito, H. (1990) *Macromolecules* **23,** 88–94.
14. Hirabayashi, H., Ishikawa, Kakudo, M. & Go, Y. (1968) *Sen-I Gakkaishi* **24,** 397.
15. Saito, H., Tabeta, R., Asakura, T., Iwanaga, Y., Shoji, A., Ozaki, T. & Ando, I. (1984) *Macromolecules* **17,** 1405–1412.
16. Ando, I., Saito, H., Tabeta, R., Shoji, A., Shoji, A. & Ozaki, T. (1984) *Macromolecules* **17,** 457–461.
17. Asakura, T., Demura, M., Date, T., Ogawa, K. & Williamson, M. P. (1997) *Biopolymers* **41,** 193–203.
18. Fossey, S. A., Nemethy, G., Gibson, K. D. & Scheraga, H. A. (1991) *Biopolymers* **31,** 1529–1541.
19. Momany, F. A., McGuire, F. F., Burgess, A. W. & Scheraga, H. A. (1975) *J. Phys. Chem.* **79,** 2361–2381.
20. Nemethy, G., Pottle, M. S. & Scheraga, H. A. (1983) *J. Phys. Chem.* **87,** 1883–1887.
21. Sippl, M. J., Nemethy, G. & Scheraga, H. A. (1984) *J. Phys. Chem.* **88,** 6231–6233.
22. Ichimura, S. & Okuyama, K. (1989) *Polym. Prep. (Japan)* **38,** 487.
23. Oka, M., Baba, Y., Kagemoto, A. & Nakajima, A. (1990) *Polym. J. Japan* **4,** 416–425.
24. Konishi, T. & Kurokawa, M. (1968) *Sen-I Gakkaishi* **24,** 550–554.
25. Lotz, B. & Keith, H. D. (1971) *J. Mol. Biol.* **61,** 201–215.
26. Gay, D. M. (1983) *ACM Trans. Math Software* **9,** 503–512.
27. Mayo, S. L., Olafson, B. D. & Goddard, W. A., III (1990) *J. Phys. Chem.* **84,** 8897–8909.
28. Fletcher, R. & Reeves, C. M. (1964) *Comput. J.* **7,** 149–154.
29. Alexander, L. E. (1969) *X-ray Diffraction Methods in Polymer Science*, Robert E. Krieger Publishing Company, Malabar, FL, p. 458.
30. Treacy, M. M. J., Newsam, J. M. & Deem, M. W. (1991) *Proc. R. Soc. Lond. A* **443,** 499–520.
31. Lotz, B., Gonthier-Vassal, A., Brack, A. & Magoshi, J. (1982) *J. Mol. Biol.* **156,** 345–357.
32. Padden, F. J., Jr. & Keith, H. D. (1965) *J. Appl. Phys.* **36,** 2987–2995.
33. Keith, H. D., Giannoni, G. & Padden, F. J., Jr. (1969) *Biopolymers* **7,** 775–792.
34. Anderson, J. P., Steven-Hassard, M. & Martin, D. C. (1994) in *Silk Polymers: Materials Science and Biotechnology*, Vol. 544, Kaplan, D., Adams, W. W., Farmer, B. & Viney, C., Eds., American Chemical Society, Washington, DC, pp. 135–137.
35. Chothia, C., Levitt, M. & Richardson, D. (1981) *J. Mol. Biol.* **145,** 215–250.
36. Magoshi, J., Magoshi, Y. & Nakamura, S. (1985) *J. Appl. Polym. Sci. Appl. Polym. Sym.* **41,** 187–204.
37. Kerkam, K., Viney, C., Kaplan, D. & Lombardi, S. (1991) *Nature* **349,** 596–598.
38. Kaplan, D. J., Lombardi, S. J., Muller, W. S. & Fossey, S. A. (1991) in *Biomaterials: Novel Materials from Biological Sources*, Byrom, D., Ed., Stockton Press, New York, pp. 2–53.
39. Kaplan, D. L., Fossey, S., Viney, C. & Muller, W. (1992) in *Hierarchically Structured Materials*, (Materials Research Society Symposium Proceedings), Vol. 255, Aksay, I. A., Baer, E., Sarikaya, M. & Tirrell, D., Eds., Materials Research Society, Pittsburgh, PA, pp. 19–30.