1-1-1998

The interfacial behavior of Bombyx mori silk fibroin.

Regina, Valluzzi
University of Massachusetts Amherst
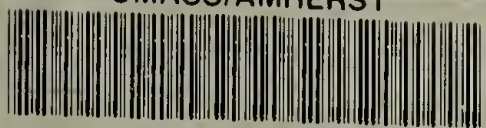
Follow this and additional works at: https://scholarworks.umass.edu/dissertations_1

THE INTERFACIAL BEHAVIOR OF *BOMBYX MORI*
SILK FIBROIN

A Dissertation Presented

by

REGINA VALLUZZI

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial
fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 1998

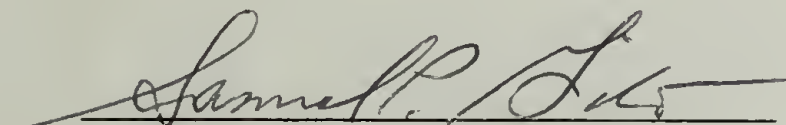Department of Polymer Science and Engineering

THE INTERFACIAL BEHAVIOR OF *BOMBYX MORI*

SILK FIBROIN

A Dissertation Presented

by

REGINA VALLUZZI

Approved as to style and content by:

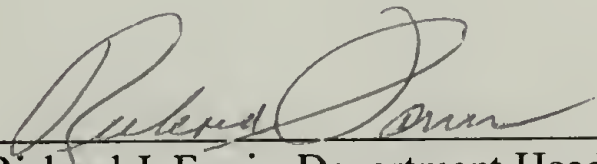_____
Samuel P. Gido, Chair

_____
Shaw-Ling Hsu, Member

_____
David A. Tirrell, Member

_____
Richard J. Farris, Department Head
Polymer Science and Engineering

ACKNOWLEDGEMENTS

ABSTRACT

THE INTERFACIAL BEHAVIOR OF *BOMBYX MORI* SILK FIBROIN

SEPTEMBER 1998

REGINA VALLUZZI, B.S. MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Ph. D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Samuel Patrick Gido

A new crystal structure has been observed for Bombyx mori silk fibroin at air-water interfaces. This structure, silk III, incorporates a left-handed three-fold polyglycine II conformation and an approximately hexagonal lattice. Detailed crystallographic studies using electron diffraction data have been used to characterize the silk III crystal structure. There is a hexapeptide repetitive sequence found in the crystallizable portions of silk fibroin. When this sequence, Gly-Ala-Gly-Ala-Gly-Ser, is in a threefold helical conformation, a row of alternating glycine and serine residues parallel to the helical axis results. One third of the helix thus becomes slightly hydrophilic, whereas the other two thirds consist of glycine and hydrophobic alanine residues. The data indicate that the helices are arranged so that the serine residues pack preferentially in serine-rich sheets in the (110) planes of the crystal. The result is a monoclinic crystal structure, where the basal plane angle γ is 116° rather than the 120° expected for perfect hexagonal packing, due to the distortion in nearest neighbor interhelical packing distances that results when the serine residues have a preferred packing. The separation of hydrophobic and hydrophilic residues in a threefold helical conformation of fibroin suggests that the air-

water interface may stabilize the threefold conformation because this conformation allows the fibroin to behave as a surfactant at the interface. The sheet-like arrangement of serine residues deduced for the monoclinic silk III crystallites also supports a role for surfactancy in stabilizing the silk III structure at the air-water interface. If the crystallizable portions of fibroin are behaving as a surfactant, then the three-fold helical silk III structure should be oriented with the axes of the three-fold fibroin helices in the plane of the interface. This orientation is observed in uncompressed fibroin films which were picked up onto TEM grids. In LB films compressed to 16.7 mN/meter on the trough prior to being picked up onto TEM grids a uniaxial orientation is observed for silk III, with the helical axes perpendicular to the plane of the sample film. A surface compression of 34 mN/meter results in films containing silk II crystallites with the same uniaxial orientation, placing the helical axes perpendicular to the plane of the film. In addition to air-water interface experiments several experiments were carried out using aqueous-organic interfaces. A hydrated crystal structure incorporating a left-handed 6/2 helix which is still roughly three-fold, is observed at the water-hexane and water-chloroform interfaces. Large lamellar crystallites possessing a hexagonal habit were observed at both of these interfaces. A banded cholesteric mesophase which in some regions crystallizes into the same silk III hydrate structure as the lamellae was also observed and characterized.

TABLE OF CONTENTS

Page

Chapter

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

BACKGROUND

The conformation (secondary structure) of proteins has a great influence over their biological activity and over their ability to self assemble into larger structures. In the case of the globular proteins, secondary structure formation may influence the subsequent folding of the protein into its native state (tertiary structure). In contrast, fibrous proteins do not form globules, but rather assemble into fibrils or crystals containing many molecules. In a sense, the longer - range assemblies and crystals formed by fibrous proteins can be considered analogous to the tertiary structure of globular proteins, because both represent the next level of order, packing, and interaction after secondary structure is formed.

There has been a great deal of interest in understanding structure formation in proteins associated with cellular processes such as transmembrane proteins and many of the globular proteins. Interest in extracellular proteins, which include many fibrous or structural proteins has been growing recently, as evidence has grown to support an active role for extracellular structures. Previously they were believed to be passive support structures.
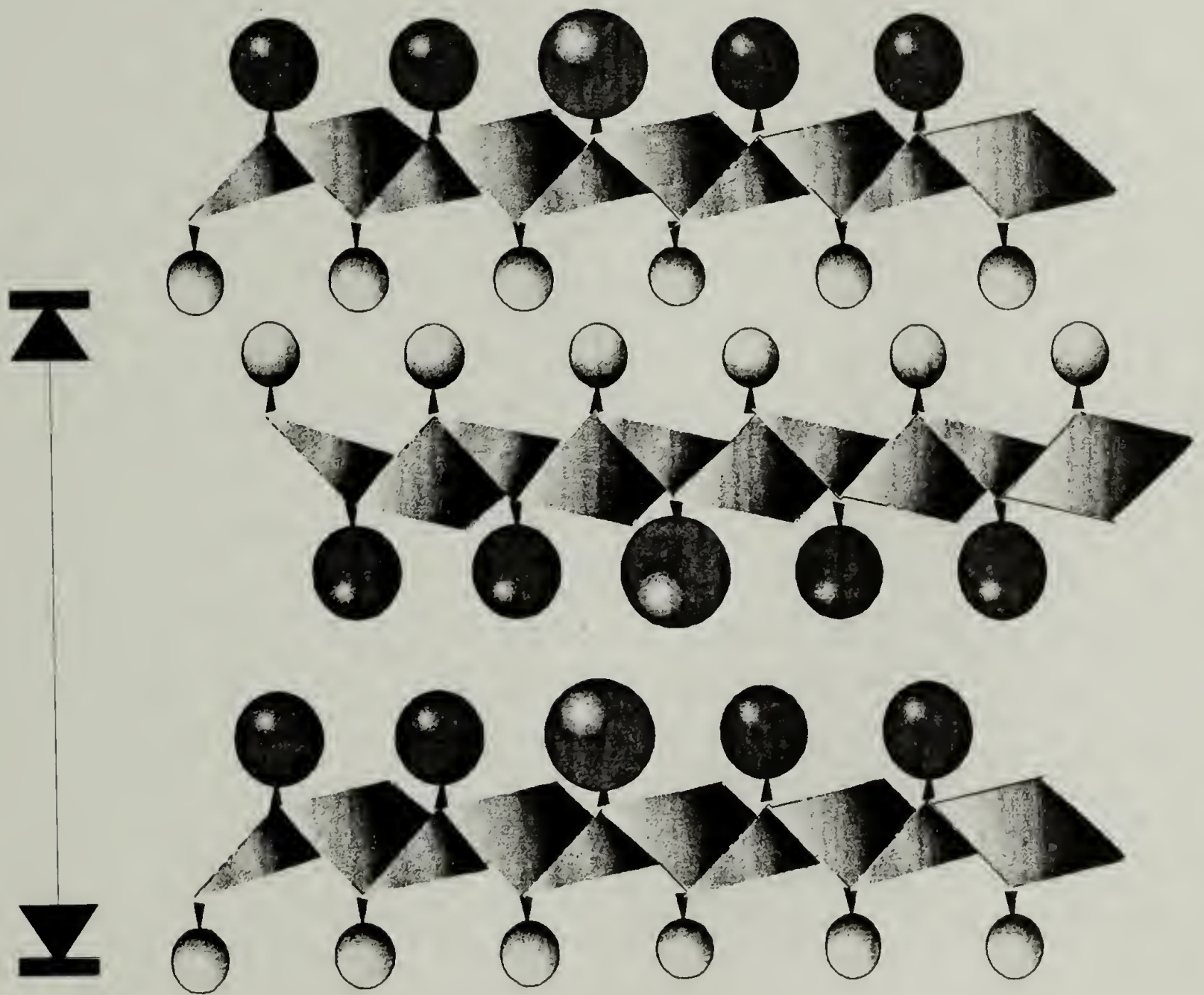
Figure 1: Asymmetric β-sheet packing. The predominance of different residue sizes on different sides of the β-sheet results in two different intersheet spacings when the sheets are packed asymmetrically. The resulting crystalline unit cell incorporates two sheets in the intersheet direction.

Fibrous proteins, which often feature repeated amino acid sequences, can in some ways be viewed as simplified models for structure formation in globular proteins. In a globular protein, regions of secondary structure interact with other regions from the same protein because the folded state puts them in close proximity. The secondary structures formed may influence the formation of tertiary structure. Conversely the environment provided by the tertiary folded structure may stabilize a particular secondary structure through interactions such as hydrogen bonding. In a fibrous protein crystal, the unit cell usually contains one or several protein helices with the same secondary structure or conformation. The interactions between helices will stabilize the conformation in analogy with the globular protein case. The crystal structure can be considered somewhat analogous to the tertiary structure of a globular protein because it results from the packing of protein regions into an ordered structure. In a fibrous protein crystal there is generally one conformation and in many cases there is a repetitive crystallizable sequence comprising the helix-forming regions. These crystalline proteins thus provide a simplified system for studying the interaction between sequence, secondary structure, and formation of larger-scale structures in proteins in general.

In addition to aiding in the understanding of the physical phenomena underlying the structure of globular proteins, fibrous proteins with a fairly straightforward and well-known functions, such as silks, can be used as models for more complicated fibrous proteins. They may share general structural features and behavior with other fibrous proteins that play a more active and complicated biological role such as collagen.

1.1 Silks

Silks have been the subject of numerous protein structural studies. *Bombyx mori* silk consists of two major protein components, sericin and fibroin.[1-14] The sericin protein acts as a binder, holding the strong highly crystalline fibroin protein filaments together to form a cocoon. In natural silk fibers, the fibroin is semicrystalline, containing both crystalline and amorphous material.[1,5,10-12,15-17] Studies on the chemical composition of silk fibroin have revealed that the crystalline portions contain only four residues: glycine, alanine, serine, and a small amount of tyrosine.[1,2,4,6,8,9,18-23] These residues are arranged in a very regular sequence where tyrosine marks the beginning and end of a crystallizable segment, and most of the crystalline sequence is built up from the hexapeptide repeat Gly-Ala-Gly-Ala-Gly-Ser.[1] The sequence is currently believed to be Gly-Ala-Gly-Ala-Gly-Ser-Gly-*Ala-Ala*-Gly-[Ser-Gly-(Ala-Gly)$_n$]$_8$Tyr which incorporates strict alternation of glycine with alanine or serine except at the two sequential alanine residues in italics.[20-22] Since n is usually two, the crystallizable sequence can essentially be considered as a repeating hexapeptide: (Gly-Ala-Gly-Ala-Gly-Ser)$_x$.

1.1.1 Silk Biology

Silk proteins are generated by a wide range of organisms, including silkworms and spiders. Silkworm cocoon silks from *Bombyx mori* are well characterized and a good understanding of their protein composition, structure, genetics and processing is available.[2,4,6,9,14,17-36] These cocoon silks are composed of two main structural proteins (fibroins) and a family of glue-like proteins (sericins). The two fibroins consist of a heavy chain (~325,000 Da) and a light chain (~25,000 Da). The fibroin heavy chain is characterized by a highly repetitive core region consisting of Gly-Ala-Gly-Ala-Gly-Ser repeats in large blocks; these blocks form the crystallizable portion of the protein.[14,19,25-27,33-35,37]

1.1.1.1 Silk Spinning

The biological silk spinning process produces protein fibers with a reasonably high crystallinity and orientation using processing stresses available to a silkworm and aqueous conditions.[1,8,10-13,15,17,23,24,28,29,36,58] The silk fibroin protein starts in a water soluble state where it may be a random coil, but in the spun fiber it is an insoluble crystalline β-sheet. Under gentle crystallization conditions where the solution conformation might be expected to be preserved, the silk I crystalline polymorph is obtained. Spectroscopic evidence, notably from IR and NMR studies,[16,53-55,59-63] indicates that many of the conformational features of silk I are similar to those of fibroin in the amorphous solid state and in aqueous solution.[53,54,60,62-64] The crystalline unit cell observed for silk I is small, and clearly does not represent a random coil, but rather a very regular structure. However features such as the torsional angles along the fibroin chain backbone and the relative predominance of either intrachain or interchain hydrogen bonds are probably very similar between silk I, amorphous silk, and fibroin in solution. An understanding of the nature of the silk I conformation will aid elucidation of the conformation of silk fibroin in aqueous solution. Any new insight regarding the conformation of water soluble silk can be used to help understand how the biological silk spinning process produces strong insoluble fibers using an aqueous processing environment. The concentration of silk in the B. mori silk gland is believed to vary from 15 - 30% fibroin as the silk progresses towards the spinneret.[5,10,11,39]
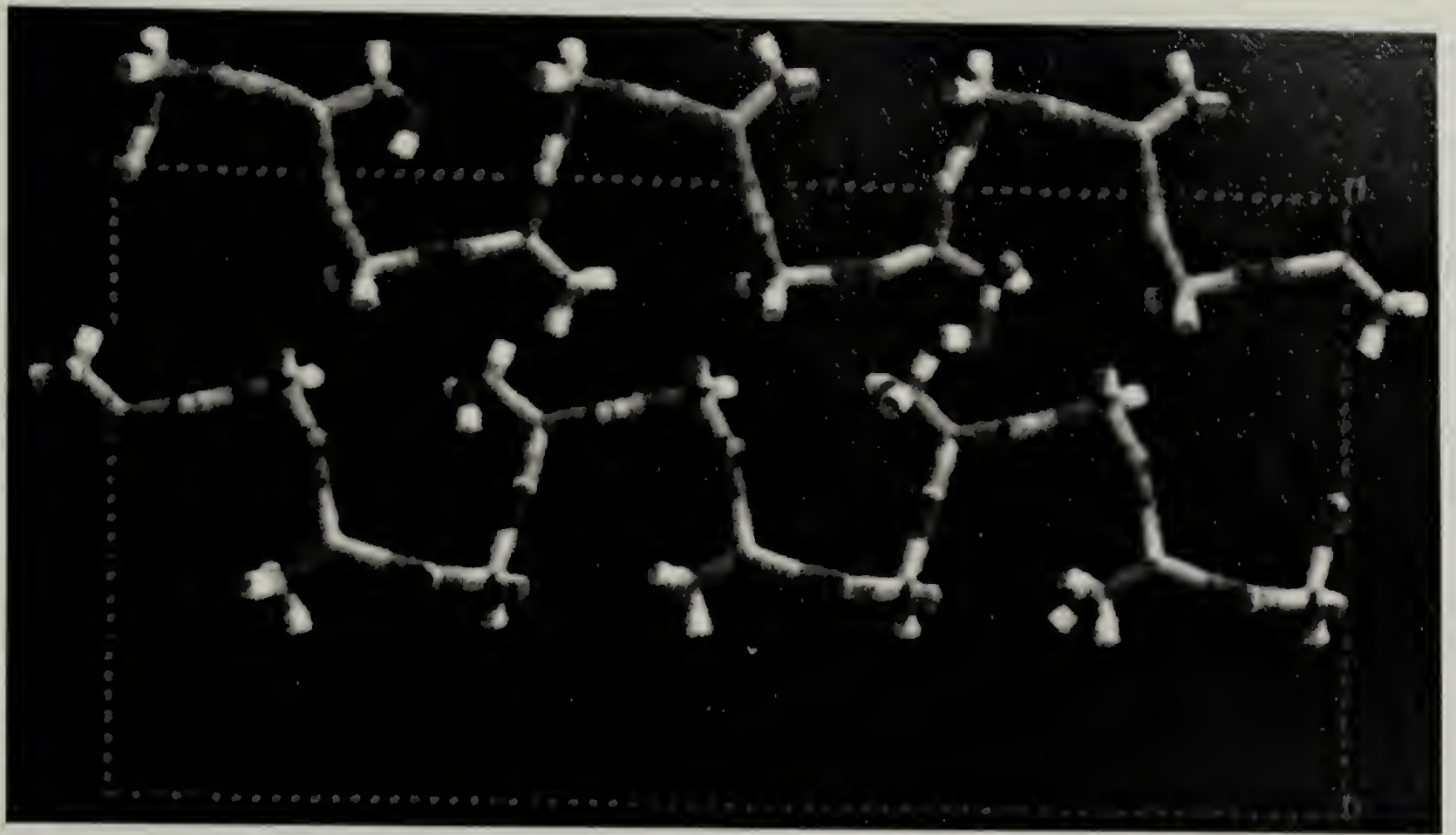
Figure 2: The crankshaft conformation proposed for silk I (Created using Cerius)
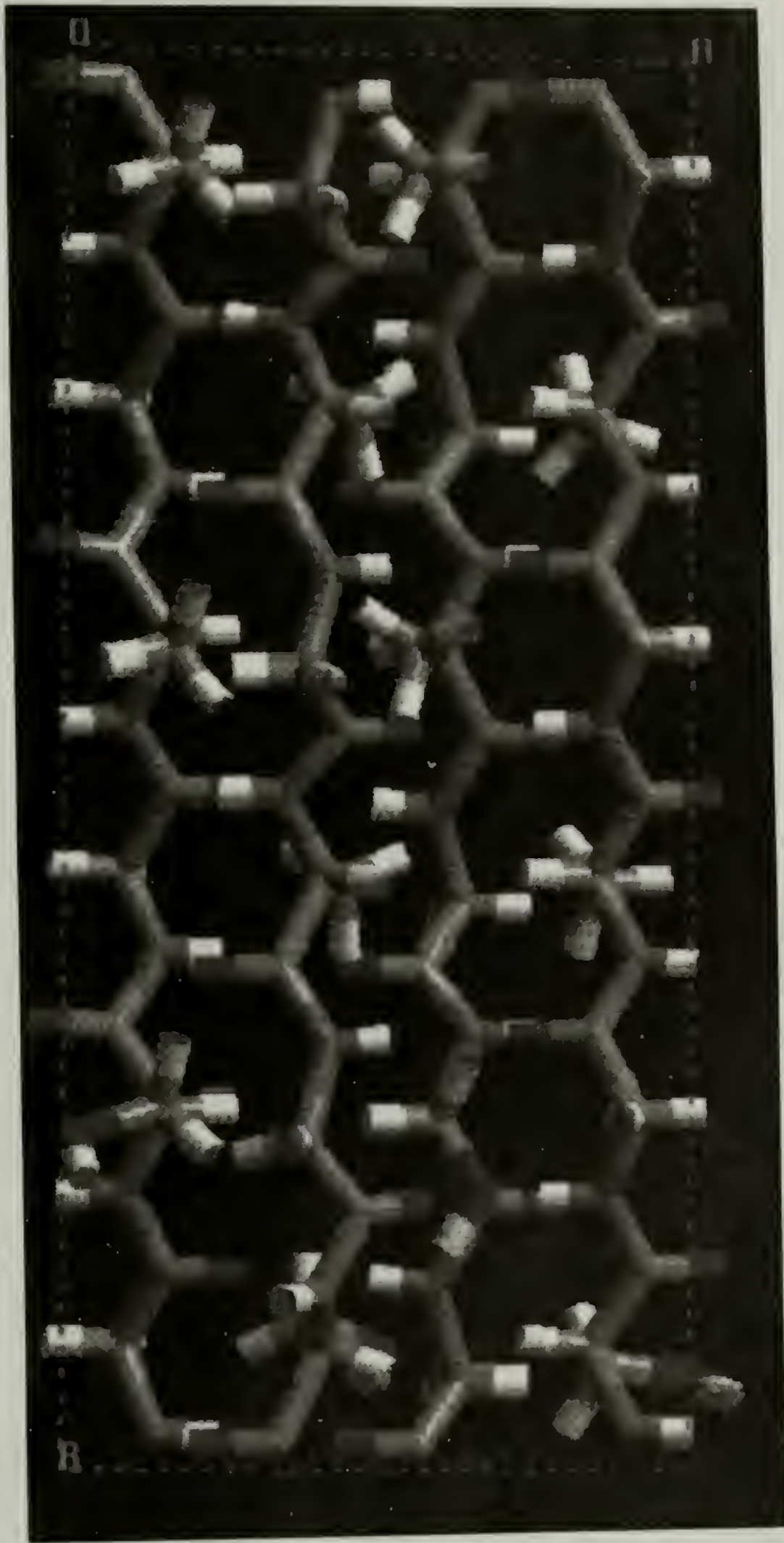
Figure 3: The silk II β–sheet conformation. (Created using Cerius)

The other major component of silk cocoons (in addition to fibroin) is sericin.[65] Sericin was named for its high serine content and is now known to consist of a mixture of serine-rich proteins with a range of molecular weights. Different fractionation studies have reported different numbers of component proteins in serine, but in all cases there are a group of lower molecular weight sericins (< 50,000 Da) and one or a few higher molecular weight components (>100,000 Da).[1,20,22,23,25,33-35] In the *B. mori* silk gland, the fibroin and sericin do not appear to be mixed. The sericin resides in a shell surrounding a fibroin core. In most studies on silk glands, when the gland is cut open to remove the silk proteins, the sericin and fibroin appear to be a gel or very viscous liquid. These observations may be indicative of the properties of silk solutions *in vivo*, or may be an artifact of the dissection process.[5,7,15,28,66,67]

The lower molecular weight sericin components are more water soluble than fibroin than the higher molecular weight sericin components. They are more easily removed when fibroin is purified from cocoons. The higher molecular weight sericins are less water soluble than low molecular weight sericins and can be difficult to separate from fibroin.[66-70] Epitaxial crystallization of high molecular weight sericin onto β−sheet crystalline fibroin has been reported.[65,68,69,71] The sericin takes on a β−sheet conformation and crystallizes in an orthorhombic unit cell that is different from those reported for silk I and silk II, but has spacings in the intersheet direction that are commensurate with the need to pack serine in the intersheet regions. Crystallization of

sericin appears to be enhanced by moisture or humidity, and precipitation of sericin has been reported for aqueous salt solutions of sericin that are cooled.[65-69,71]

The pleated β–sheet conformation crystallizes into a unit cell that contains two chains. The individual chains in the extended two-fold helical conformation associate into pleated sheets through a two-dimensional network of hydrogen bonds (between amide nitrogens and carbonyl oxygens) which are perpendicular to the chain axis. This interchain distance is fairly constant. The pleated sheets formed by the hydrogen bonded chains in the β conformation then stack up to form a three dimensional crystal. The distance between the sheets varies depending on the pendent groups from the residues comprising the protein. Depending on the primary structure of the protein, all the intersheet distances can be the same, or the sheets can stack so that intersheet regions containing predominantly large residues alternate with intersheet regions containing predominantly smaller residues. In the latter case there are two different intersheet distances and the crystalline unit cell encompasses two sheets. This case is shown in Figure 1, and the arrow shows the spatially periodic repeat encompassing two sheets.

1.1.1.2 Silk Crystallization

Silkworm cocoon silks are semicrystalline materials with amorphous flexible chains reinforced by strong stiff crystals.[38] Most silks assume a range of different secondary structures during processing from soluble protein in the glands to insoluble spun fibers. Marsh et al. first described the crystalline structure of spun silk as an

antiparallel hydrogen bonded β-sheet based on characterization of *B. mori* fibroin.[1,3,5,8,10,11,39-48]

The existence of two different crystal structures, silk I and silk II, was first observed in the 1950's [36,48,49,56] and these two bulk crystal structures and the relationship between them has been an active topic of study. Aqueous solutions of fibroin can be crystallized to form bulk crystalline solids that are composed primarily of silk I or silk II depending on the crystallization conditions. In interfacial studies involving fibroin, we have observed a new structure, silk III, which incorporates a threefold helical polyglycine II conformation. Dimorphism involving a β-sheet structure and another structure (either similar to silk I or similar to polyglycine II) has been observed in silk-like and collagen-like model polypeptides as well as in bulk studies of fibroin. The subject of conformation selection and its influence on crystal growth and polymorphism in fibroin may thus have more general implications. The existence of a specific interfacial conformation and the structures that result from this conformation can be used as a springboard for studies of conformation study. Variation of the interfacial environment can provide insight into specific factors that can influence conformation and aggregation in fibroin, and may be applicable to the wider area of fibrous extracellular proteins in general.

1.1.1.3 Silk I

The Silk I polymorph appears similar to amorphous silk fibroin when observed using spectroscopic techniques and may reflect the solution conformation of fibroin. Its structure is not well characterized. The silk I structure is unstable, and upon shearing, mechanical drawing, heating, spinning, or exposure in an electric field or to polar solvents such as methanol or acetone converts to silk II.[40,41,45,49,53-57] This structure is formed when the silk inside a silk worm gland is allowed to air dry in the absence of mechanical disturbance, and is believed to have an orthorhombic unit cell. Both the Silk I and Silk II polymorphs can be obtained from regenerated aqueous silk fibroin solutions, where LiBr or LiSCN is used to solubilize silk fibroin and the resulting solution is dialyzed to remove these salts.

In order to avoid formation of the Silk II β–sheet polymorph, precipitants, changes in solvent quality, or temperature changes are introduced very gradually. The same crystallization conditions are used to obtain polyglycine II, the threefold conformation of polyglycine shown in Figure 4. However, the conformation adopted by silk and poly(Ala-Gly) under these conditions does not appear to be threefold, but is not well understood.[60,72] It has not been possible to obtain oriented samples, to better characterize the Silk I crystal structure, because any attempts at orientation destroy the delicate Silk I crystallites. The exact structure of Silk I and its poly(Ala-Gly) analog may involve a "crankshaft" chain conformation, shown in Figure 2.[1,60,72-74]

1.1.1.4 Silk II

Silk II is found in the crystalline regions of silk fibers and has an monoclinic unit cell with the protein chains in a pleated β– sheet (two fold "zigzag") conformation.[1,36] Silk II forms when silk is sheared during crystallization.[61] The silk II structure is shown in Figure 3. Stroking a fibroin solution to induce orientation results in the formation of silk II crystals. Silk II is also observed when silk is crystallized from water using dehydratants and under conditions similar to those used to crystallize polyglycine I, a similar β–sheet form of polyglycine.[14-21] Note that the numbers I and II, indicating polymorphs, are reversed for polyglycine versus silk. Thus polyglycine I and silk II are both β–sheet structures while polyglycine II and silk I are not. Polymorphs of model silk-like polypeptides follow the silk nomenclature.

The silk II crystal structure was originally believed to be orthorhombic, but as refinements to the structure became possible the lattice type was found to be monoclinic.[1,8,23,36,40,49-52] The dimensions of the silk II unit cell, shown in Figure 3 are: 9.4 Å (a, interchain), 6.97 Å (b, fiber axis), 9.2 Å (c, intersheet). These unit cell dimensions are consistent with an antiparallel β-sheet crystal structure. Interchain hydrogen bonds perpendicular to the chain axis between carbonyl and amine groups stabilize the individual β-sheets, and Van der Waals forces stabilize intersheet interactions. The β-sheets, consisting of the glycine-alanine/serine crystalline repeats in the silkworm fiber are asymmetric (see Figure 1), with one surface primarily projecting alanyl methyl groups and seryl oxymethyl groups and the other surface of the same sheet

containing hydrogen atoms from the glycine residues. Physical shear generated during spinning of the soluble silk appears in a large part responsible for conversion from the soluble silk I-like solution conformation to the insoluble silk II β-sheet in the natural spinning process.[5,10,11,39]

A transition between the silk I crystal structure and the silk II structure has been reported and is of interest because of its possible bearing on the silk spinning process. When fibroin in the silk I form is highly swollen with aqueous salt, it no longer has a crystalline diffraction pattern in an X-ray experiment. When the water is removed again, the silk I diffraction pattern reappears. If the swollen silk I is subjected to stress, the silk II crystal structure results. Unfortunately, experiments on silk I and its transformation to silk II have utilized polycrystalline samples which were characterized by bulk techniques such as x-ray diffraction. In the absence of data on single crystals it is impossible to rule out alternatives to a silk I to silk II transition that could yield similar scattering data. One such possibility would be transformation of amorphous material to crystalline silk II accompanied by destruction of delicate silk I crystallites.
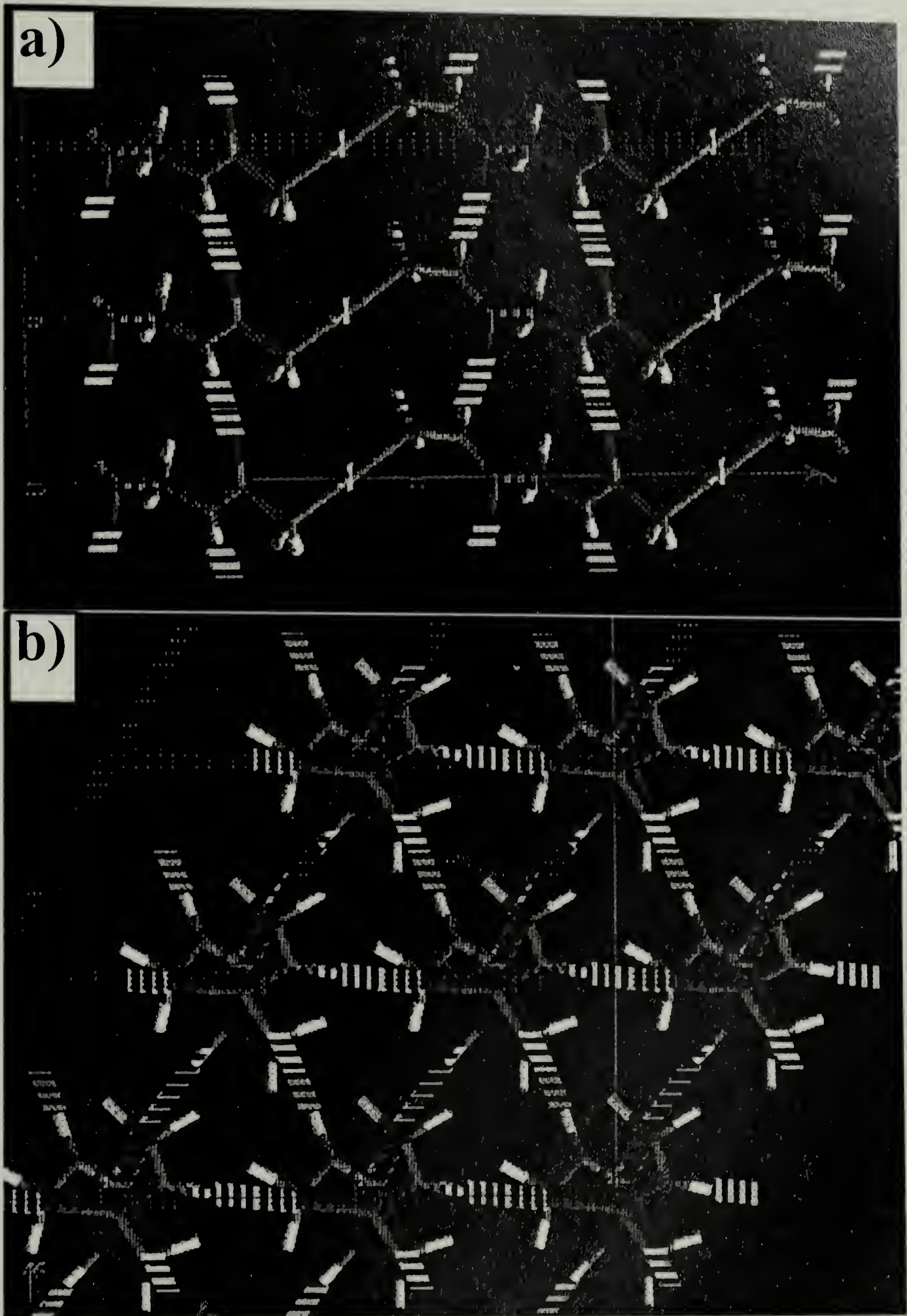
Figure 4: The polyglycine II structure. (a) Side view (b) Top view, looking down the chain axes. The hexagonal packing of polyglycine II threefold helices can be seen.

Interestingly, while polyglycine does not appear to possess a crystalline polymorph similar to silk I, a transition analogous to the silk I to silk II transition is observed in films of polyglycine II. Recall that polyglycine II is the threefold helical dimorph of polyglycine. In these films, stroking of the film or other mild mechanical treatment, converts the polyglycine II into polyglycine I, the extended β-sheet dimorph. The transformation of polyglycine II into polyglycine I is believed to involve rotation of the peptide groups to the polyglycine I hydrogen bonding positions without gross translational motion that would disrupt crystallinity. A similar rearrangement is difficult to envision for the controversial crankshaft structure proposed for silk I. Because of its similarity to the polyglycine II crystal structure, the silk III interfacial crystal structure may be useful in clarifying behavior attributed to silk I.

1.1.1.5 Model polypeptides

Many studies of fibrous protein conformational behavior have focused on synthetic polypeptide models. Once the repeating sequence of interest in a fibrous protein has been identified, a polypeptide can often be synthesized to incorporate only that sequence, and the effects of different permutations of residues can be isolated. The conformational behavior of the crystallizable fibroin and its crystalline polymorphs should be analogous to similar polyhexapeptides or closely related polycopeptides. The crystallizable portion of silk fibroin should be similar to either the polydipeptide,

16

poly(Ala-Gly), or to the polyhexapeptide poly(Gly-Ala-Gly-Ala-Gly-Ser). Because of its high glycine content silk fibroin also bears some chemical similarity to the polytripeptides and other polypeptides used to model collagen, another fibrous protein. Theoretical conformational analysis of a silk model, poly(Ala-Gly), indicates a rich variety of sterically allowable regular helical conformations,[74] and it is plausible that variables such as exact repeating residue sequence or solvent, pH, temperature, pressure, and dimensionality of the molecule's local environment will all play a strong role in determining the conformation adopted by fibroin and its subsequent crystalline form.

In some of the polypeptides that are chemically similar to silk fibroin, several conformational changes can be observed depending sensitively on variables such as concentration of the polymer, pH, and solvent quality.[75-77] One of these models, poly(Ala-Gly), mimics the bulk crystalline dimorphism observed in silk fibroin. Under conditions where fibroin would be expected to form the silk I structure, poly(Ala-Gly) forms a structure that appears closely analogous, poly(Ala-Gly) I. Under physical or chemical stress poly(Ala-Gly) II forms. The poly(Ala-Gly) II crystalline polymorph incorporates a pleated β-sheet conformation, similar to that observed in the fibroin crystal structure, silk II.

Many of these polypeptides possess a crystalline polymorph that incorporates a threefold helical conformation. This polymorph occurs for copolypeptides with more than 50% glycine and appears to be more common in poly*tri*peptides.[78-83] *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 threefold helical crystalline conformations. In the polypeptide studies, the effect of a strict alternation of glycine with another residue was thoroughly studied only for poly(Ala-Gly) [52,72,83] The more complicated polyhexapeptide poly(Ala-Gly-Ala-Gly-Ser-Gly) should accurately mimic silk, but we were only able to find studies using conditions analogous to those used to obtain Silk II.[50]

1.2 Greater Implications

1.2.1 Fibrous Proteins

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 better known fibrous proteins is *Bombyx mori* silk.

While there have been numerous studies on the chemical structure and physical chemistry of silk proteins, especially fibroin, there are still unanswered questions regarding the structural behavior of this protein. The occurrence of two crystalline polymorphs in the bulk raises questions regarding the conformational diversity of the fibroin molecule. Since a third structure is observed exclusively at interfaces the possibility of interfacial chemistry influencing the conformation of a flexible protein must be considered. The supermolecular assembly of fibrous proteins into larger scale ordered structures is still not fully understood. The structures formed are not necessarily crystalline, and may in some cases be mesophases. Changes in the overall morphology of silk fibroin interfacial films as a result of changes in the interfacial environment may be relevant to the larger area of fibrous protein self-assembly.

1.2.2 Proteins at Interfaces

The fibrous proteins can be considered to be a special class of proteins that exist outside of cells and serve a structural function in addition to any biological activity they might possess. In living organisms these proteins are often found in thin layers, sandwiched between other extracellular biomaterials. When studying the physical chemistry of extracellular fibrous proteins *in vitro*, dimensional restriction, such as the use of a two dimensional thin film or interfacial environment, will help the proteins self-assemble more

efficiently. The limited dimensionality of the environment provides far less "room" or volume to explore than a three-dimensional bulk environment. There is thus a great potential benefit in studying the synergistic interaction between the behavior of fibrous proteins in dimensionally restricted environments (such as thin films or two-dimensional layers) and the generation of structure and long range order through self-assembly.

There have been a number of studies on the adsorption of proteins to air-water interfaces. Generally, when a protein is studied at an air-water interface, a combination of Langmuir surface pressure measurements and IR spectroscopic examination of Langmuir-Blodgett films picked up off of the interface is used to determine the interfacial behavior and structure of the protein film.[84-90] Much of the published data focus on globular proteins. It is found that some membrane proteins self-assemble readily at a two-dimensional interface.[91] Many proteins, including hemoglobin and pepsin, will denature at an air-water interface, complicating a Langmuir trough experiment. Denaturation at the air water interface can be a fairly slow process, taking several minutes, and is related to the amount of spreading and expansion of the film observed.[84,85,87,88,92-95] The formation of water insoluble aggregates of denatured proteins has been observed for solutions shaken to create a high air-water interfacial surface area.[93] Studies of proteins at the air-water interface are complicated further by the fact that natured proteins tend to be water-soluble. Simultaneous adsorption, desorption, denaturation, and agglomeration essentially occlude the straightforward two dimensional pressure-area phase behavior that would be expected for a synthetic polymer

or small molecule Langmuir film. While spread monolayers of protein do not exhibit the sharp Π-A transitions typical of synthetic polymers, reproducible results can be obtained and at least in the high pressure regions, the isotherms appear to be representative of a colligative property.[84,85,87,88,92-94]

Different conformations can be stabilized by the interface, such as an extended chain β-sheet conformation which maximizes the protein's surface area and spreading. If the protein or model polypeptide has hydrophobic side chains, and can readily take on a stable α-helical conformation, α-helices will be stable at the interface.[84-88,92-101] If the predominant interfacial conformation incorporates intrachain hydrogen bonds, like an a-helix, or other features that cause the protein to behave as a stiff or wormlike chain, liquid crystalline aggregates or micelles can form. Liquid crystallinity will greatly affect the behavior of a protein interfacial layer, and surface pressure can be used to orient a liquid crystalline adsorbed protein layer. The influence of side chain character in stabilizing an interfacial conformation suggests that hydropathicity can be used as a determinant for interfacial conformation. Carrying this idea further, if a sequence of residues results in particular conformations that could exhibit surfactant behavior, these conformations should be stabilized at an interface. In our work with *B. mori* silk fibroin, we find that a threefold helical polyglycine II or polyproline II type of conformation is stabilized by the interface, even though it is not observed in bulk.[102-104]

As a consequence of the difficulties entailed in attempting detailed surface measurements at a liquid-liquid interface, there have been few studies on the behavior of proteins at these interface to date. Murray and Nelson, working with a novel trough design,[94,95] have published intriguing results on the comparative behavior of protein films at air-water and oil-water interfaces that appear consistent with structural results obtained for fibrous proteins at air-water and oil-water interfaces. They find that films at the oil-water interface are more expanded, and also more expansible and compressible than corresponding films at the air-water interface. This is believed to be due to a reduction in agglomeration. The increased solubility of the hydrophobic groups in oil as opposed to air is cited as a reason for the greater stability of films at the oil-water interface. Shchipunov has studied phospholipids at an oil water interface,[105] and observed that the presence of the amphiphiles results in more oil on the water side of the interface and more water on the oil side. The amphiphile compatibilizes the two liquids forming the interface, and in the process, the interface thickens. Both the compatibilization effect observed for the phospholipids and the film stability observed for the protein films suggest that there is oil and water closely interacting with the side chains of the protein. Side chain - side chain interactions would thus be expected to be screened.[87,94,95,98,101,105,106] In fact, in studies of fibroin at oil water interfaces where a good solvent for oils such as hexane is used to form the oil phase, behavior is observed which strongly suggests screening of some intermolecular aggregation interactions, as is described in Chapter 5. Whereas only fibrous agglomerates and lamellar crystals are observed at the air-water interface, at the hexane-water interface a cholesteric liquid

22

crystalline surface is observed, and fibrous aggregates are not in evidence. While crystallinity is observed, the evidence suggests that the crystallinity progressed from the cholesteric molecular liquid crystal, and no large lamellae are observed.

1.2.3 Globular Proteins

Studies of globular proteins in aqueous and nonaqueous solvents clearly indicate that the role of the surrounding environment in controlling protein structure and activity is profound. Globular proteins partition in water, such that hydrophilic amino acid residues reside in or near aqueous contact while the hydrophobic residues become occluded within the protein and shielded from the water phase. This phenomenon is critical in protein folding to attain appropriate catalytic activity for enzymes.

Protein folding typically involves nonrepetitive amino acid sequences and complex three dimensional geometries, thus it is not surprising that these phenomena are so difficult to understand. By using fibrous proteins the complexity of the amino acid sequence reduced. By confining self-assembly to an essentially two dimensional interfacial region the geometrical complexity is also reduced. Within this simplified context fundamental understanding of the interplay between chemical segregation effects and chain sterics is possible, by observing the structures (chain conformations) formed as a function of residue sequence and interface characteristics.

1.3 Objectives

In these experiments, the crystal structure, crystalline texture (orientation), and in some cases mesophase structure and orientation, were characterized for regenerated silk fibroin processed using a variety of interfacial conditions. A new crystal structure, which we call silk III, was found to predominate at the interface. This structure incorporates a left-handed threefold helical conformation of fibroin. Langmuir Blodgett films from surface compression experiments were examined to determine the effect of surface pressure on film structure, including the overall morphology of the films, the crystal structure of the crystalline fibroin present, and the orientation of any fibroin crystals found in the films. In order to further probe the influence of interfacial environment on the structure of fibroin, aqueous-organic liquid-liquid interfaces were also used to create surface excess films of fibroin.

Through these experiments a better understanding of the interfacial structure of fibroin is expected. The conformation present in silk III, the interfacial crystal structure of silk fibroin, will be elucidated as will the details of the crystal structure. The data from the different interfacial environments, when viewed together, will be used to assess reasonable explanations for the presence of a threefold helical conformation of fibroin as the interfacial preferred conformation.

CHAPTER 2

SILK III, A NOVEL CRYSTAL STRUCTURE

2.1 Summary

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 than the two previously observed silk crystal structures, silk I and silk 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 threefold 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 threefold 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.

2.2 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 1950's, [36,48,49,56] and has been an active topic of study ever since.

There are two previously known crystalline forms of *B. mori* silk, silk I and silk II. The conditions under which these two polymorphs are formed have been well documented in previous studies.[1,36,40,41,48,49,55,56,58,64] 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, shown in Figure 5, has a monoclinic unit cell with the protein chains in a pleated β–sheet (two fold "zigzag") conformation [1,36] which forms when silk is sheared during crystallization.[61] The structure of silk I [48,49,56] 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.
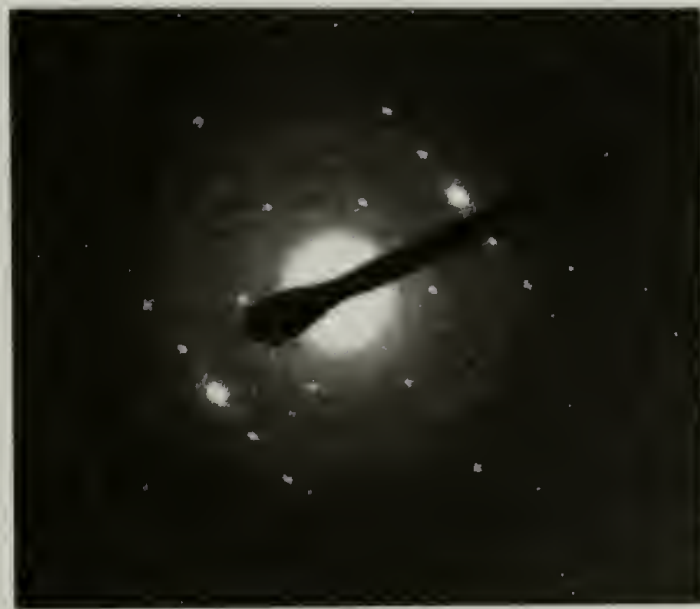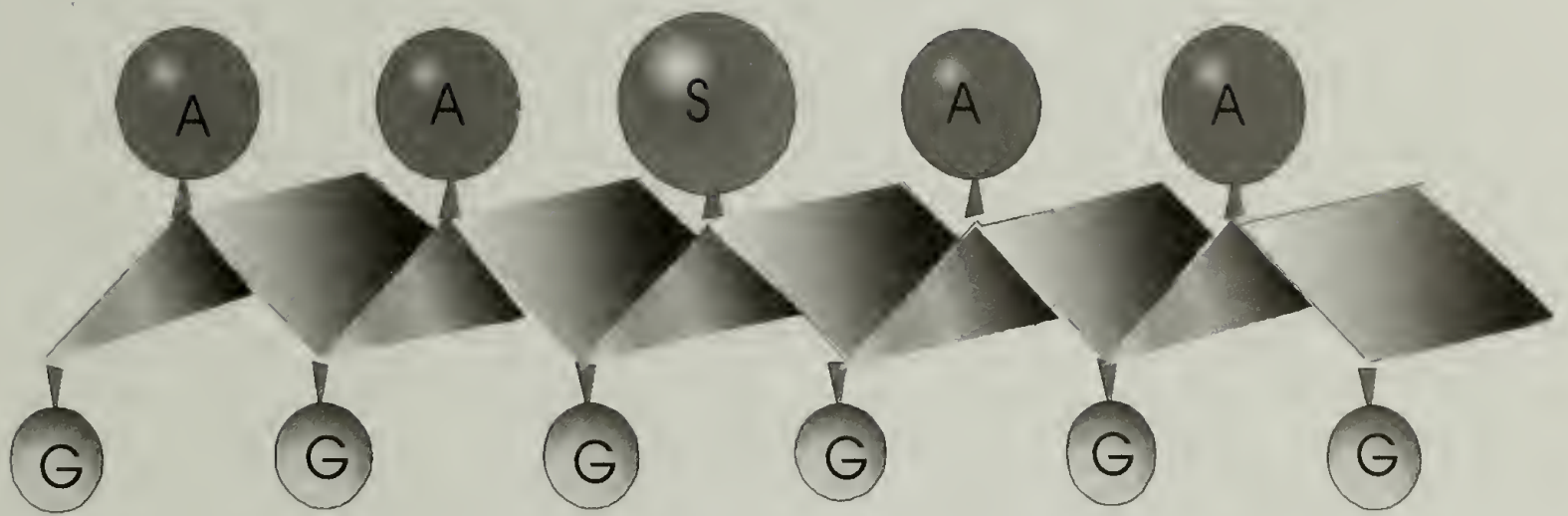
Figure 5: Silk II. *Top:* The β-sheet conformation. *Bottom:* Single crystal diffraction
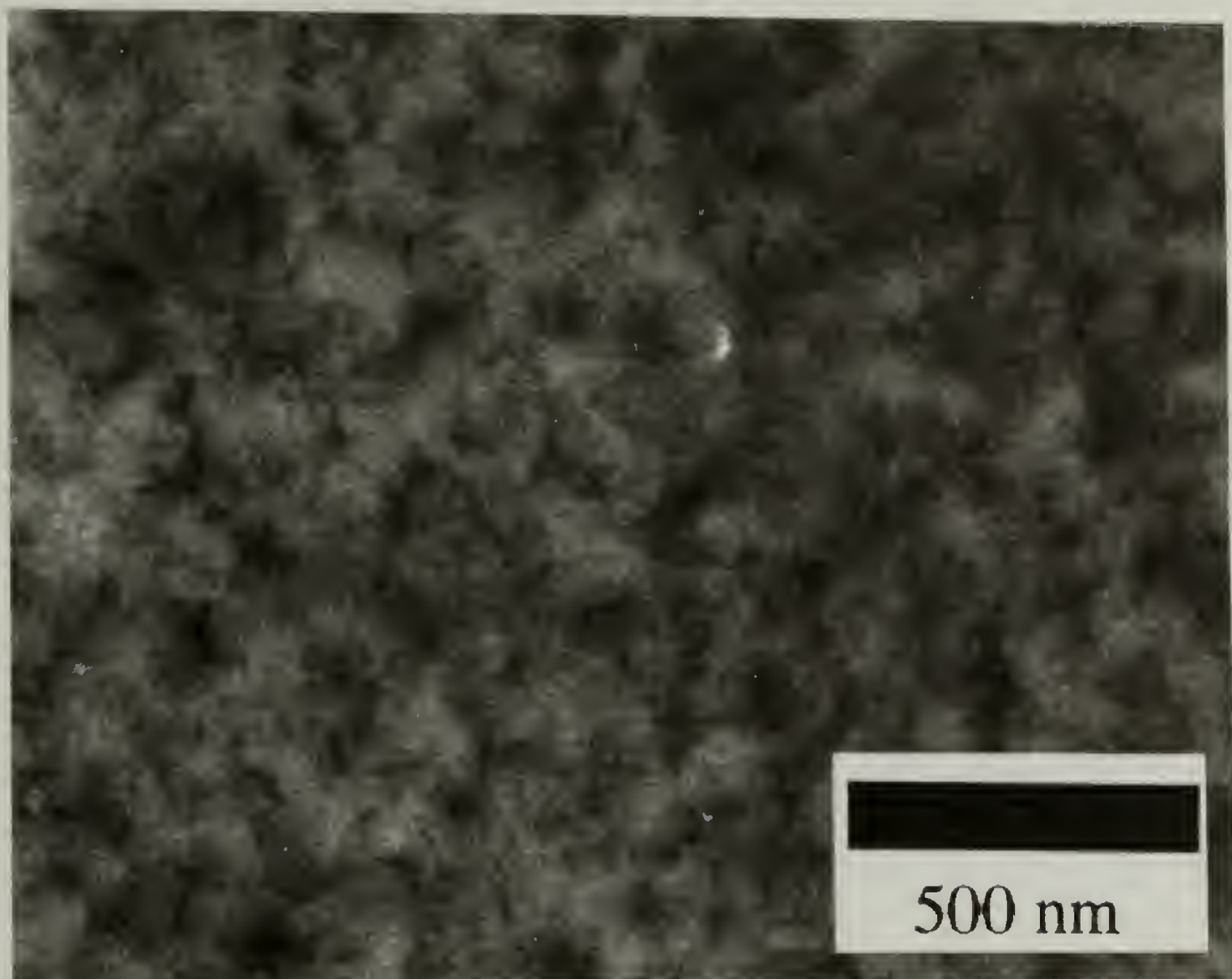
Figure 6: Brightfield TEM morphology image of a typical film region. The dark needle-like crystallites diffract polycrystalline rings.
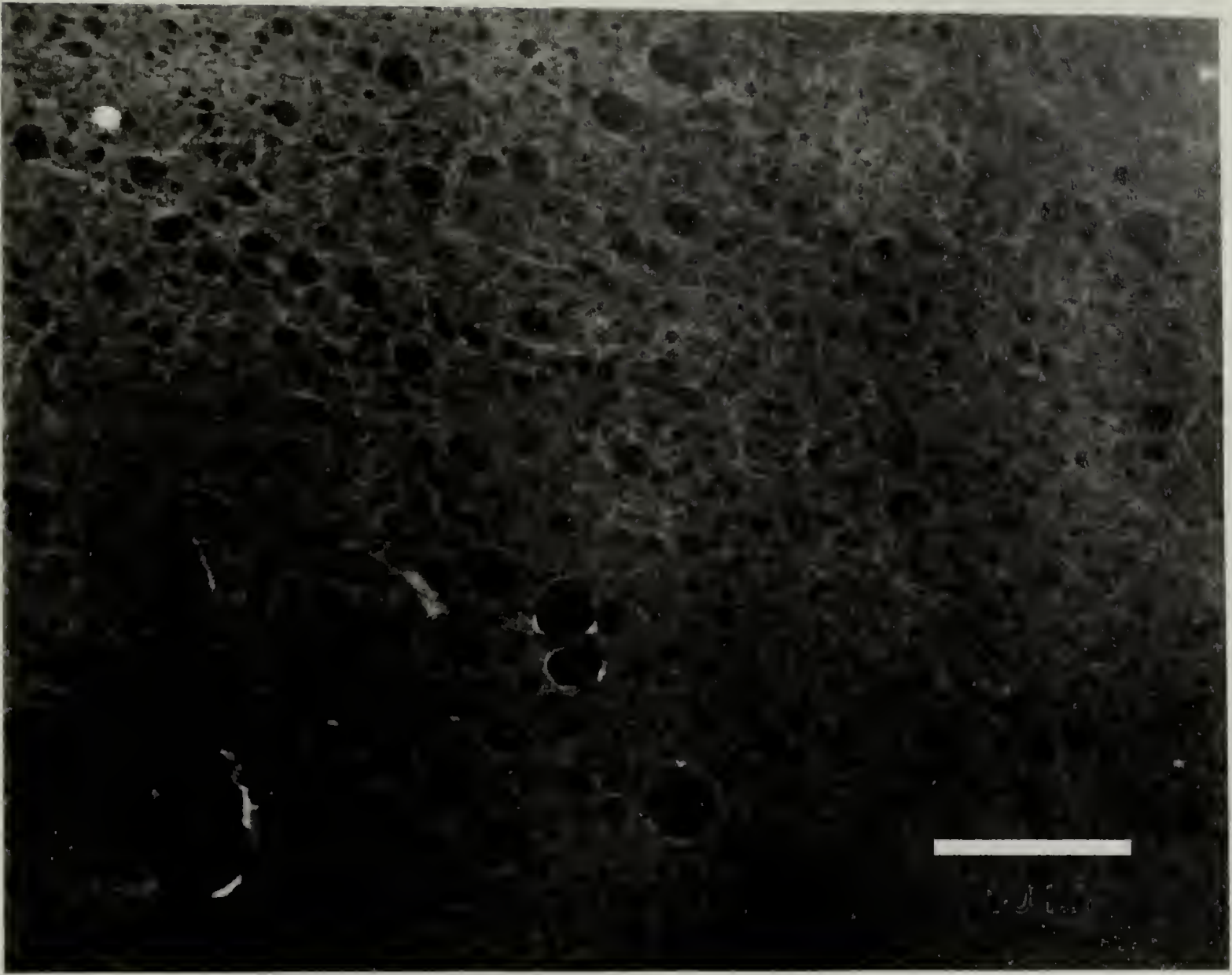
Figure 7: Darkfield high angle annular STEM image of a typical LB film morphology. The needle-like crystallites appear bright; this is evidence of crystallinity.
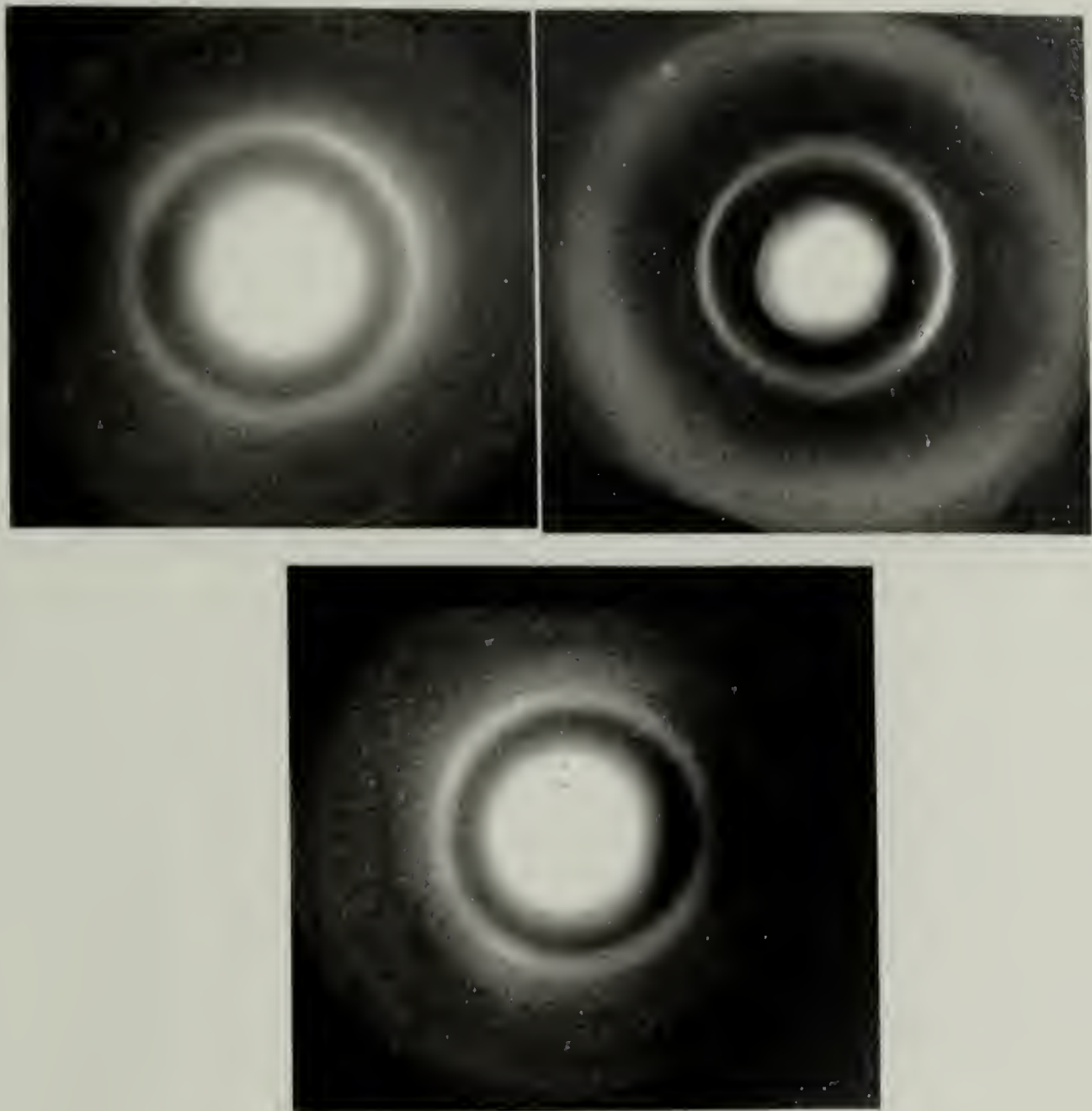
Figure 8: Polycrystalline Silk III diffraction. (a, b) Uniaxially oriented LB film
(a) 0° sample tilt (b) 25° sample tilt (c) Unoriented cast film yields "powder-like" diffraction
patterns
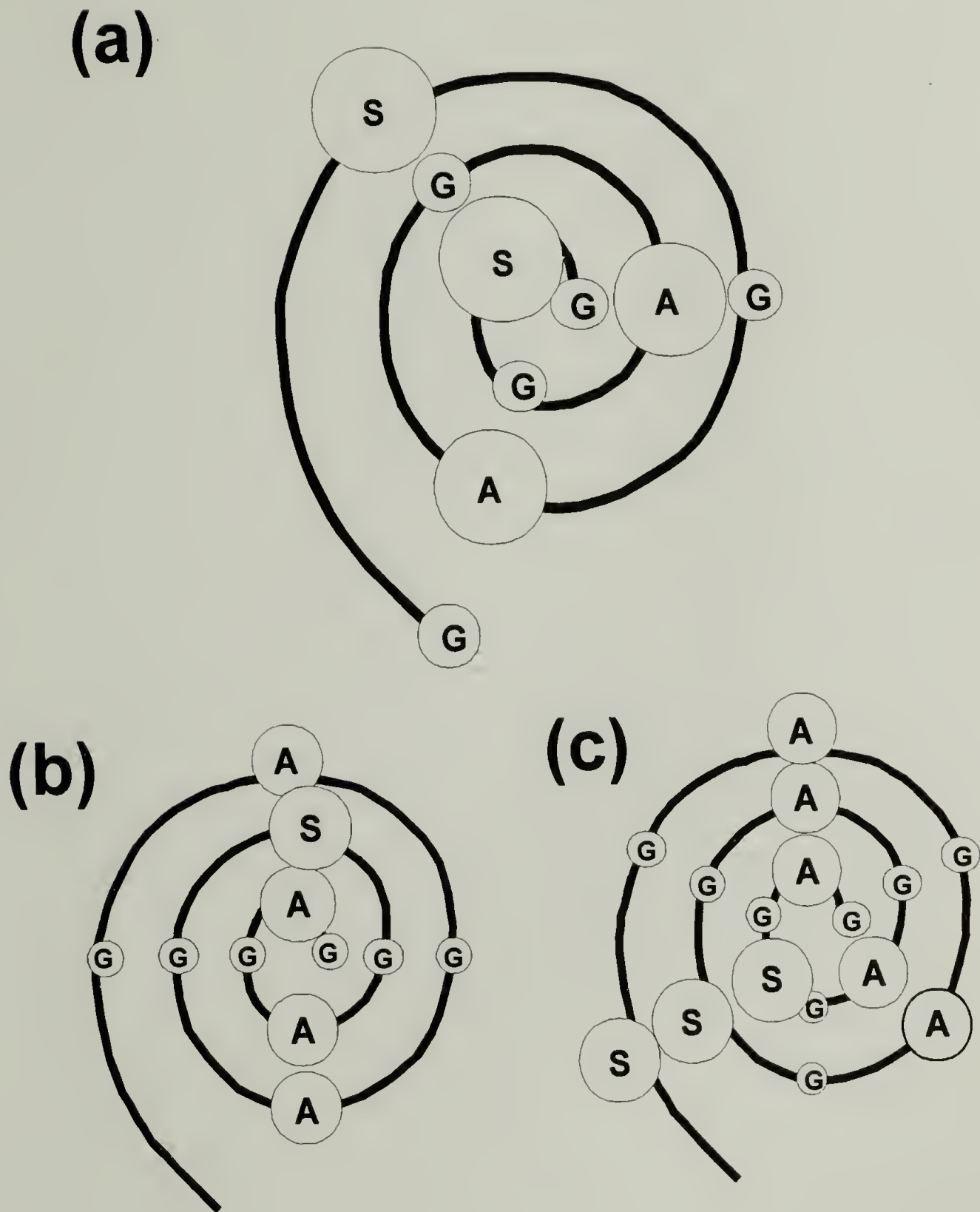
Figure 9: (a) In the threefold conformation the glycine residues in the hexapeptide fibroin crystallizable repeat (Gly-Ala-Gly-Ala-Gly-Ser) alternate with the larger residues. Thus sterical crowding is alleviated in a threefold conformation because the glycine residues act as spacers. In a fourfold (b) or sixfold (c) conformation, crowding is evident.

Here we report the discovery of a new crystal structure of *Bombyx 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 threefold helical structure. The new three-fold helical silk polymorph was first observed, in our initial experiments,[102] 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/meter.[107,108] 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/meter, 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/meter 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 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 6 and Figure 7 show the type of semicrystalline film morphologies which we observe. Typical electron diffraction patterns for the LB and cast films respectively are shown in Figure 8a and Figure 8c. Following the nomenclature of previously discovered polymorphs we will call this structure silk III. The name silk III has been used previously[109] for another reported

structure of silk fibroin also displaying pseudo-hexagonal symmetry, which never-the-less appears to be different from the structure presented here.

Bombyx 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.[11] The crystallizable sequence of silk fibroin can essentially be considered as a repeating hexapeptide: (Gly-Ala-Gly-Ala-Gly-Ser)$_x$.[1,2,20-22] Many silk-like polypeptides possess a crystalline polymorph that incorporates a threefold 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 polytripeptides such as poly(Gly-Gly-Ala).[79-83] 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 threefold helical crystalline chain conformations. Nevertheless, the strict alternation of glycine with the larger residues, serine and alanine, results in favorable sterics if the silk takes on a threefold helical conformation. This is illustrated in Figure 9 by diagramming the placement of the different residues as seen looking down the chain axis. In the threefold (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 twofold (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.
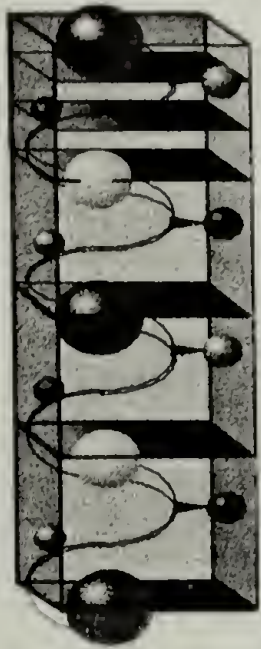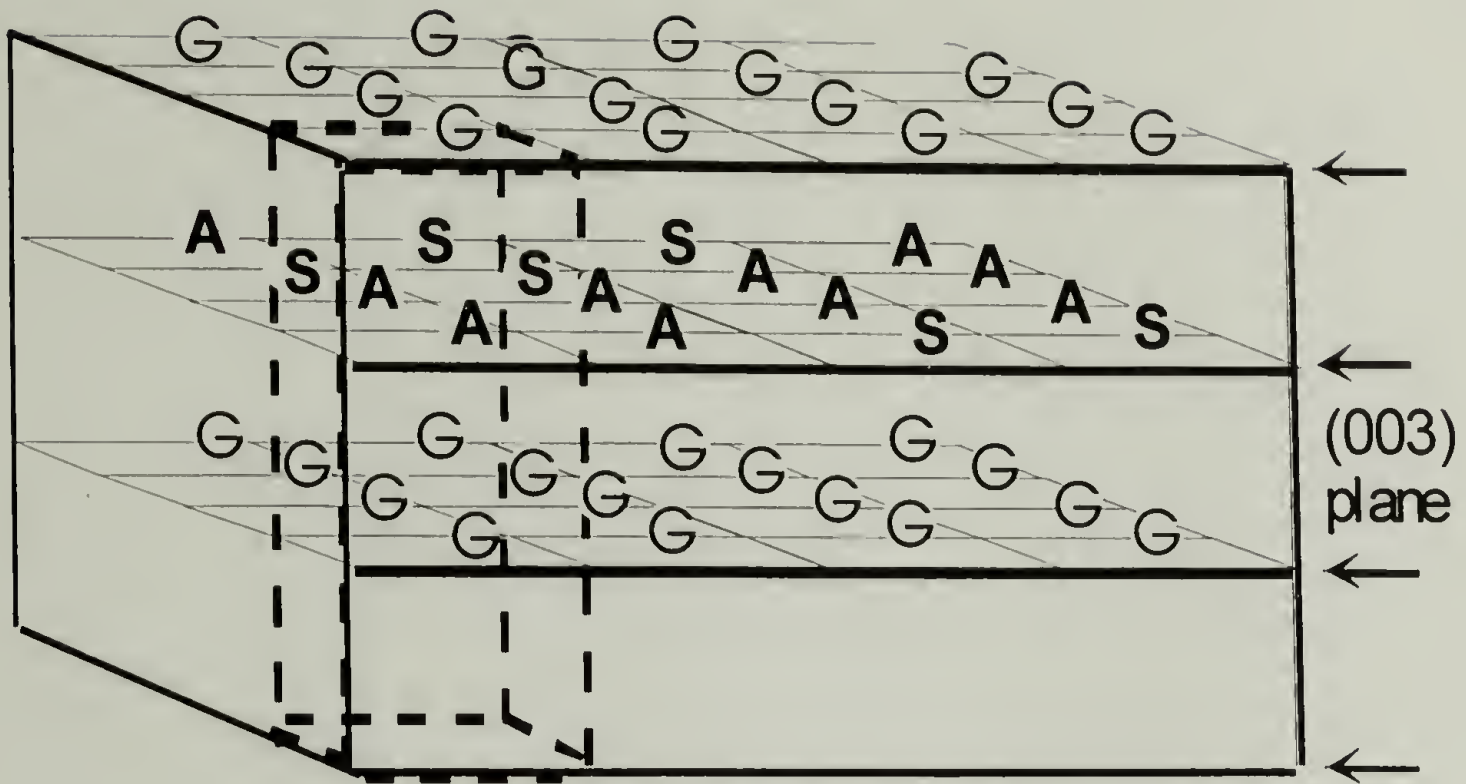
Figure 10: There are different possible lateral arrangements of residues in the 003 planes in a silk III crystal. An ordered arrangement incorporating alternation of planes of glycine and of the larger residues serine and alanine is shown. Such an ordered arrangement would result in a larger unit cell than the primitive three residue cell observed.

2.3 Experimental

2.3.1 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 one week using a 12,000 to 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. Ultra-thin silk films were deposited onto TEM grids using the Langmuir-Blodgett 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 hour or 24 hours at 25°C and a surface pressure of 16.7 mN/meter before being deposited onto TEM grids.
2. Unsupported cast films were prepared from a 3 wt% aqueous silk solution.

2.3.2 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 was 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 Coulombs/cm^2 at 200 kV for the diffraction spot pattern to completely disappear at cryogenic (-160° C) temperatures.

In order 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. In order 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.

2.3.3 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.[110,111] 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.

2.3.4 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.[112,113] 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. In order 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 by 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 6 and Figure 7, 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 effect 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 measured for polyglycine II.[114] 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 was 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[6,9], but differ significantly in their aspartic acid and glutamic acid contents. Fibroin contains approximately 1.39 mole% aspartic acid and 1.06 mole% 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 were easily distinguished from 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 mole % aspartic acid and 1.0 mole% glutamic acid.[6] 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[6,9] 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.
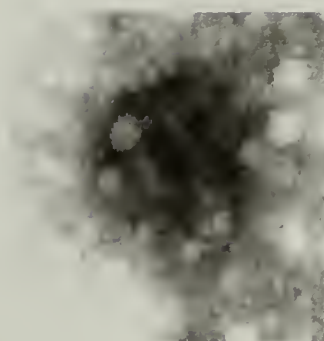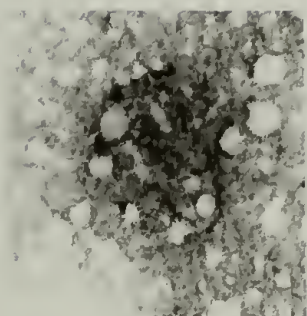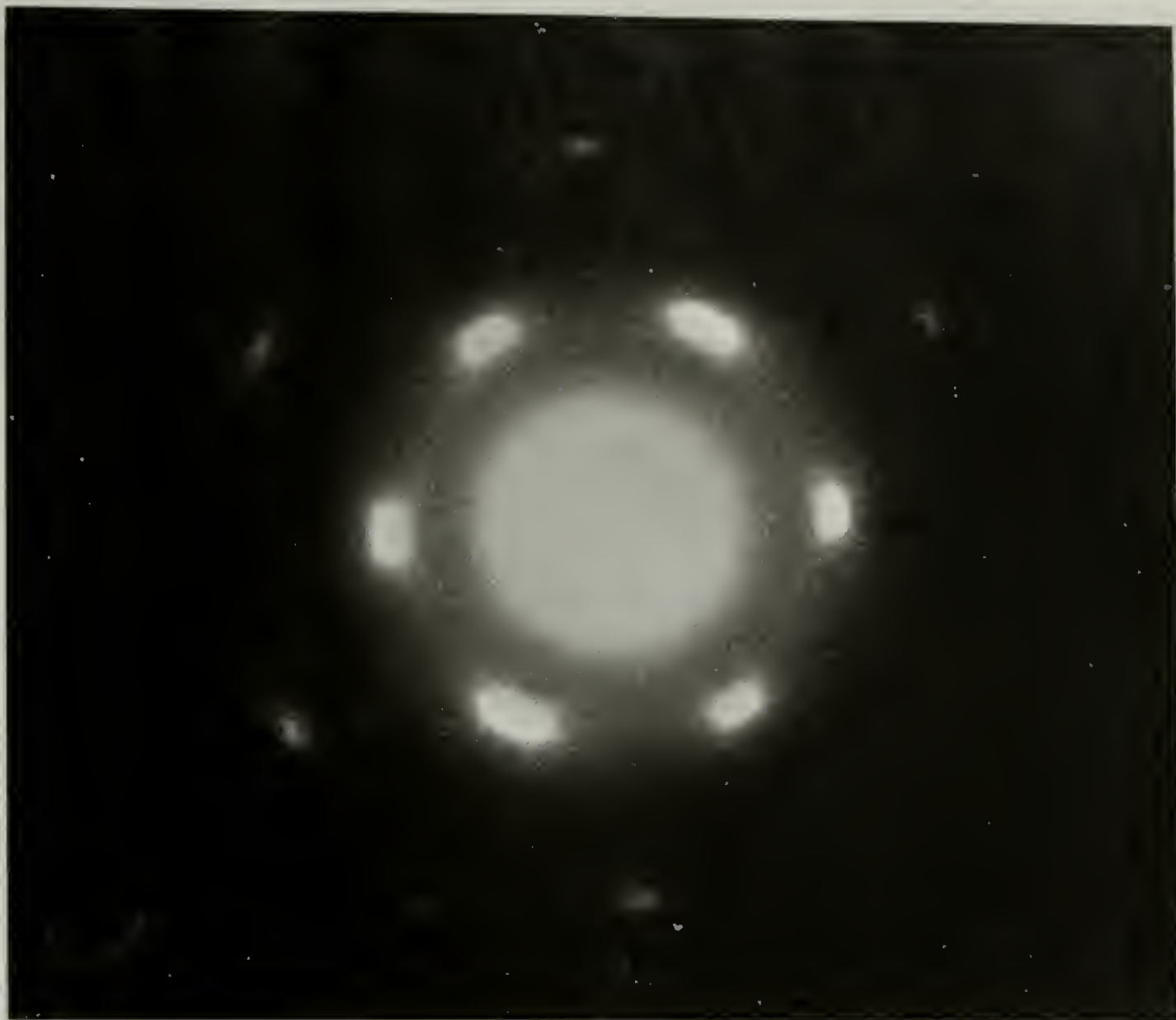
Figure 11: Silk III lamellar crystallites (a) Typical hexagonal single crystal diffraction pattern from a silk III lamella. (b) Contrast reversed image of a typical hexagonal lamella from a darkfield STEM image. (c) Low pass Fourier filtered image of the crystallite in *(b)*. The outlines are easier to see when the graininess from the scan is filtered out.

2.3.5 Results and Discussion

Figure 6 shows a bright field TEM micrograph of a silk LB film (surface pressure 16.7 mN/meter) from sample 1, indicating that the films consisted of very small needle-like crystallites embedded in a featureless amorphous matrix. The darkfield STEM image in Figure 7 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/meter 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 8a and c for LB films and cast films respectively. Additionally, faint images of very thin triangular lamellar crystals were sometimes visible in the LB films. Figure 6 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 11 shows a spot patterns taken from a single lamellar crystal at room temperature and a morphology image for a typical lamella. Single crystal patterns such as

this are frequently superimposed on rings, as seen in the Figure, because the lamellar crystal invariably rests on a mat of needles within the region from which diffraction data is 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 was used for quantitative comparison of observed diffracted intensity to the intensities predicted by structural models. Table I lists the d-spacings and integrated intensities for the ten 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.[114] 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 $00l$ 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 threefold helical conformation[114] resulting in the following unit cell dimensions: $a = b = 4.8$

Å, c = 9.3 Å, α = β = 90°, γ = 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, α = β = 90°, γ =

120°. The Miller index assignments of experimentally observed silk III diffraction maxima

based on this unit cell are listed in Table I. The spacings and symmetry of the six fold

pattern of spots in the single crystal silk III diffraction patterns (Figure 11) indicate a similar

hexagonally packed structure, analogous to that found for polyglycine II.

2.3.5.1 Unit Cell Modeling

The silk III crystal structure was modeled using the *Cerius* software package

from Molecular Simulations Incorporated running on a Silicon Graphics Indigo

workstation. Based on the similarity of silk III diffraction patterns to polyglycine II,

crystal structures with a three-fold helical chain conformation were investigated. In

polyglycine II the glycine residues comprising the helix are non-chiral, 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[115,116] 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).[116-118] 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 (threefold) helix is more stable than the right-handed helical

conformation. The left-handed (threefold) 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 3-fold left-handed helices, while collagen forms a left-

handed helix that is also close to a three-fold helical conformation.[32,79,118-126]

A model unit cell for helical silk III was constructed using polyglycine II as a general structural motif. Ramachandran[121,122] 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.

There are several distinguishable ways of packing threefold 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 10. For example, the packing arrangement in Figure 10 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 threefold helix, or three average residues in the c direction, a 3_2 screw axis, and dimensions very similar to 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 non-random. This longer periodicity would lead to a unit cell which 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 threefold 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.[114] 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 in the silk III structure was simulated by constructing a 2x2x2 super cell, as shown in Figure 12. This super cell contains two repeats of the primitive cell in each of the three crystallographic directions. This results in a group of eight sub-cells 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 super cell can then be superimposed in space to produce a primitive unit cell with an electron density distribution which 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 in order to simulate more complicated crystal structures.

A series of models using 2x2x2 super cells (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 forcefield with H-bonds and Van 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 approximately 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 to 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,[127,128] were calculated to determine which models best matched the data. R is calculated using:

$$
R = \frac{\sum\limits_{hkl} \left| F_{theor} - sF_{exp} \right|}{\sum\limits_{hkl} F_{ave}} , s = \frac{\sum\limits_{hkl} F_{theor}}{\sum\limits_{hkl} F_{exp}} \tag{1}
$$

In Equation (1) the F's are the structure factors and $F = I^{1/2}$ where I is the integrated intensity. F_{theor} is the value of the structure factor predicted by the model, and F_{exp} is the experimentally measured structure factor, which in this case is the average over the unoriented cast film diffraction patterns.

As indicated in Table I, the highest angle peak is compound, consisting of the 104, 113, and 202 reflections. In our calculations of R the experimentally determined and modeled intensity for these three overlapping peaks are lumped together. Once the best models (R less than 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 super cell shown in Figure 12 was the best model. Artificial periodicities in the super cell 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.
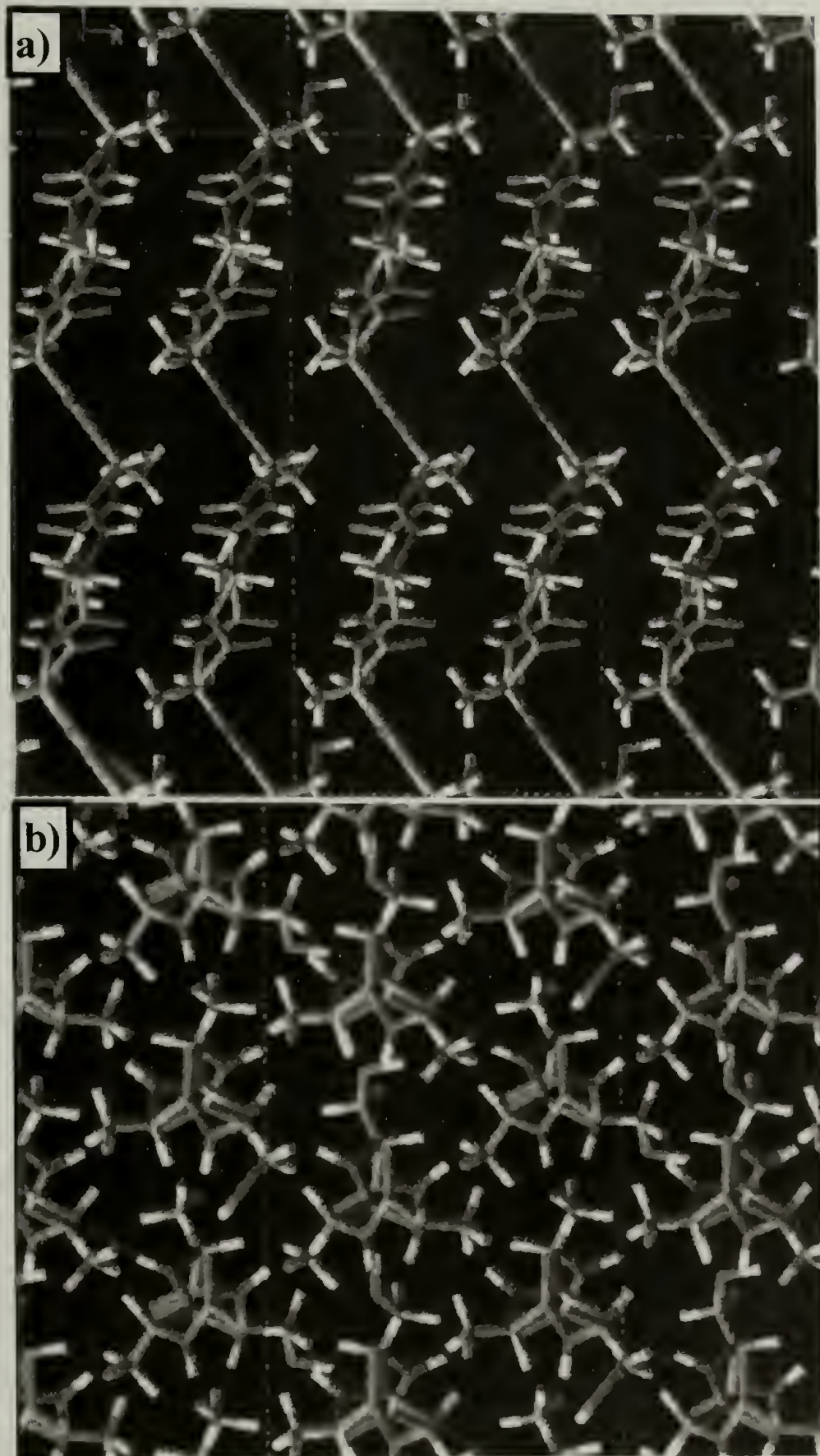
Figure 12: Superlattice structure, or "super cell", for the silk III model, consisting of 2 primitive unit cell repeats in each crystallographic direction. There are a total of eight primitive cells in the super cell used to construct the average primitive unit cell. Oxygen atoms are red, carbon atoms are gray, nitrogen atoms are blue and hydrogen atoms are white. (a) view down the *c* axis   (b) view down the *a* axis
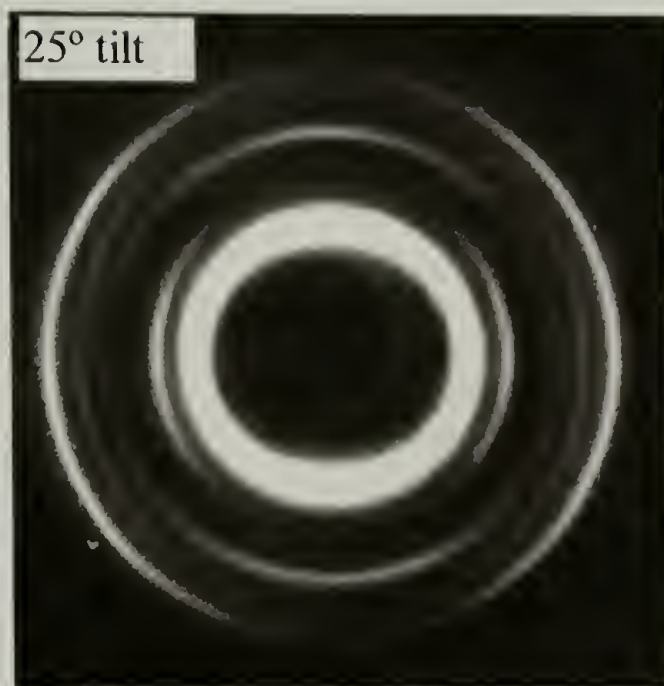
Figure 13: Simulated diffraction from the trigonal model

Packing energies and total energies were calculated for the silk III model which provided the closest match to the electron diffraction data. Energies were also calculated for silk I, and silk 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 forcefield 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 threefold 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 coordinates and Marsh's coordinates respectively.[8,72] 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/mole·residue for silk I, 51.0 kcal/mole·residue for silk II, and 25.5 kcal/mole·residue for silk III. When all atoms were allowed to move in an unconstrained minimization, an energy of around -5.5 kcal/mole·residue was obtained for silk II, and -2.3 kcal/mole·residue was obtained for silk III. In the unconstrained minimization the imposed unit cell periodicity provided the only conformational constraint, and the twofold silk II β-sheet and threefold 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.[129,130] 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.[130,131] 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°. 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 threefold 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 effect electron diffraction

results in proteins and polymers, while it is not a major concern in an X-ray diffraction study. In order 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:

$$D = \frac{\sum\limits_{hkl} \left| F_{ave} - sF_{exp} \right|}{\sum\limits_{hkl} F_{ave}}, s = \frac{\sum\limits_{hkl} F_{ave}}{\sum\limits_{hkl} F_{exp}} \tag{2}$$

In Equation (2), F_{ave} 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.
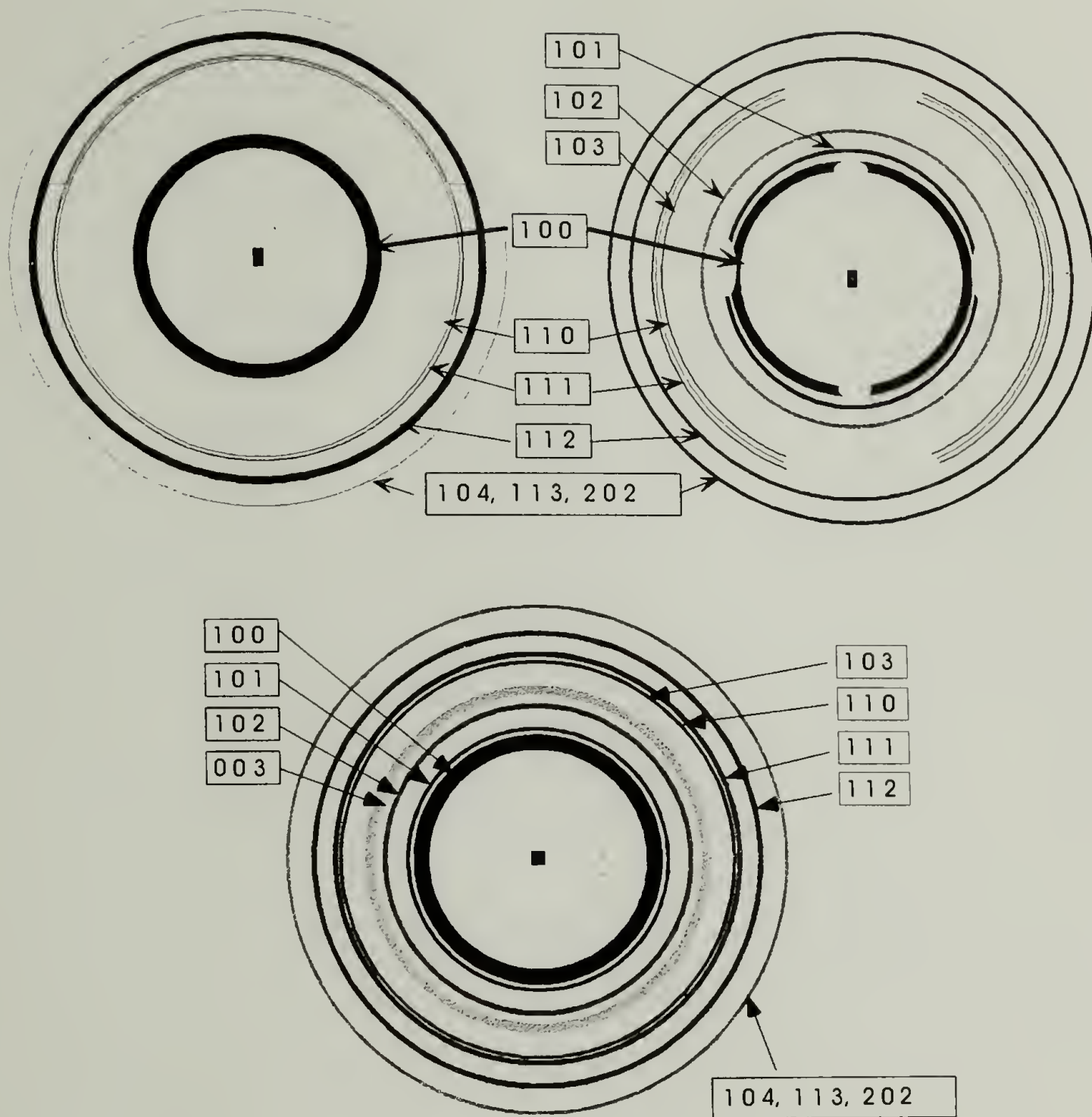
Figure 14: Schematic diagrams of the diffraction rings that result from uniaxially oriented silk III LB films and unoriented silk III cast films in the electron diffraction geometry. (a) uniaxially oriented LB film, 0° sample tilt (b) LB film, 25° sample tilt (c) unoriented cast film

Table 1: Comparison of Diffraction data to Simulations from the trigonal silk III model

Experimental and model *d* spacings and relative intensities

R = 0.11

Index	Experimental Diffraction					Predicted Model Diffraction	
	d	std dev d		I (rel)	std dev I	d	I (rel)
(1 0 0)	4.56	0.025		100	0.0	4.42	100
(1 0 1)	3.93	0.023		47	8.2	4.00	45
(1 0 2)	3.29	0.060		28	9.3	3.22	43
(0 0 3)	2.89	0.011		4	0.5	3.14	2
(1 0 3)	2.67	0.058		11	1.9	2.56	11
(1 1 0)	2.48	0.007		3	1.2	2.55	3
(1 1 1)	2.36	0.006		5	2.5	2.46	2
(1 1 2)	2.23	0.080		4	0.7	2.24	2
(2 0 1)	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
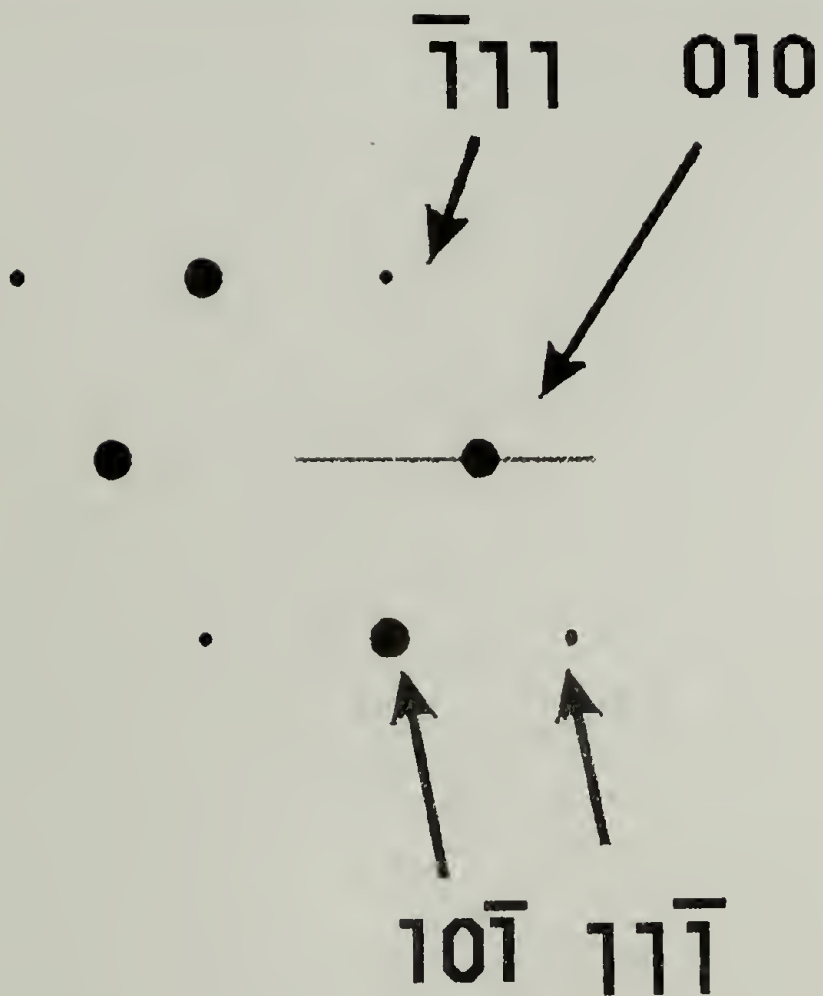
Figure 15: Tilted single crystal diffraction. (a) Tilted, room temperature, single crystal diffraction at approximately 20° sample tilt. (b) Simulated model single crystal diffraction pattern at 16° tilt.
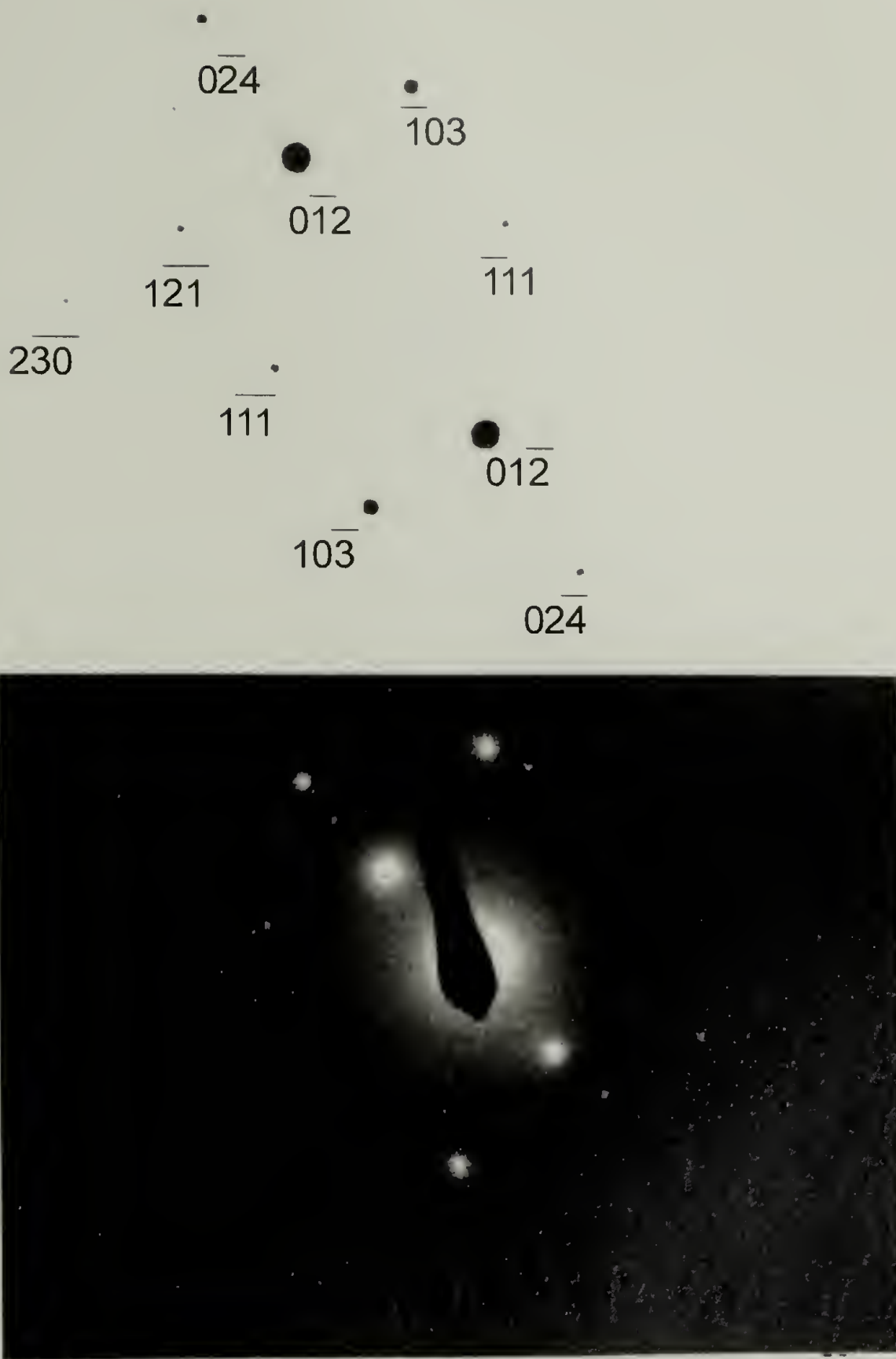
Figure 16: Different zone axes are sometimes observed for the silk III lamellae. (a) Cryogenic electron diffraction pattern obtained for a crystal of silk III near the [3 2 1] zone axis. (b) Simulated model diffraction for a single silk III crystal near the [3 2 1] zone axis.

2.3.5.2 Diffraction from Different Zone Axes of the Silk III Structure

The hexagonal diffraction pattern shown in Figure 11 was 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 15a shows a pattern which 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° to 25°. A simulated pattern at 16° to the [001] axis, Figure 15b, 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 obtain, and identical d -spacings were observed 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 based on the silk III structure. An simulated diffraction pattern, seen in Figure 16a, from the silk III structural model (shown in Figure 12) from near the [321] zone axis compares well with the corresponding electron diffraction pattern shown in Figure 16b. 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.

2.3.5.3 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 8a (LB film) and Figure 8c (unoriented 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 $hk0$ reflections should be observed. Tilting the sample will result in hkl reflections, with non-zero l, appearing as arcs in the diffraction patters.

Experimentally observed diffraction patterns from uniaxially oriented LB silk films at tilts of 0° and 25° are shown in Figure 8a and b. 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 14a and Figure 14b show 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 8a and Figure 8b, and the labeled reflections correspond to experimentally observed diffraction 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 14c, showing all of the experimentally observed reflections. If the uniaxial orientation in the LB films is perfect then Figure 8a and Figure 14a should contain only *hk0* 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 8c, Figure 14c) and in the tilted LB pattern (Figure 8b and Figure 14b).

The 101 and 102 reflections are present in both the powder patterns and in 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.

2.4 Conclusions

A new crystal structure for silk fibroin protein, with threefold 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

"fiber-like" electron diffraction ring patterns taken at several different tilts, and by unoriented

"powder-like" patterns. The dimensions measured for the unit cell of silk III, when

compared to 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 based on diffraction evidence.

CHAPTER 3

THE INTERFACIAL STRUCTURE OF SILK FIBROIN :
REFINEMENTS TO THE MODEL

3.1 Summary

A new crystalline polymorph of *Bombyx mori* silk, which forms at the air-water interface, has been characterized. A previous study found this structure to be trigonal, and to be distinctly different than the two previously observed silk crystal structures, silk I and silk II. This new structure was named silk III. Identification of this new silk polymorph was based on evidence from transmission electron microscopy and electron diffraction, coupled with molecular modeling. In the current paper, additional data enables us to refine our model of the silk III structure. Some single crystal electron diffraction patterns indicate a deviation in symmetry away from a perfect trigonal unit cell to monoclinic unit cell. The detailed shape of the powder diffraction peaks also supports a monoclinic cell. The monoclinic crystal structure has an non-primitive unit cell incorporating a slightly distorted hexagonal packing of silk molecular helices. The chains each assume a three-fold helical conformation, resulting in a crystal structure similar to that observed for polyglycine II, but with some additional sheet-like packing features common to the threefold helical crystalline forms of many glycine-rich polypeptides.

3.2 Introduction

Many of the biological functions that proteins play involve an interface. The behavior of a protein at an interface; its conformation, segregation of functional groups by the interface, and its ability to form larger structures such as crystals or mesophases, is thus an important and very rich aspect of a protein's physical chemistry. Studies have been done on proteins at interfaces[86,98,132,133] and the air-water interface is very accessible for controlled experimental studies. In many of these studies, the interface is found to alter the protein conformation, often resulting in denaturation.[84,85,87,92,93,96,97,106] Most of the interfacial protein studies have involved fairly complicated globular proteins, many of which normally fulfill a specific function at an interface. On the other hand, fibrous proteins, such as silk fibroin, consist predominantly of only a few amino acids in highly repetitive sequences. These fibrous proteins often crystallize in regular or slightly irregular helical conformations, and commonly exhibit crystalline polymorphism. A crystal of a fibrous protein can be entirely composed of some secondary structural motif, such as an α-helix or β-sheet, and thus can be considered a simplified model of a specific facet of conformational, or secondary structural, development in more complicated proteins.
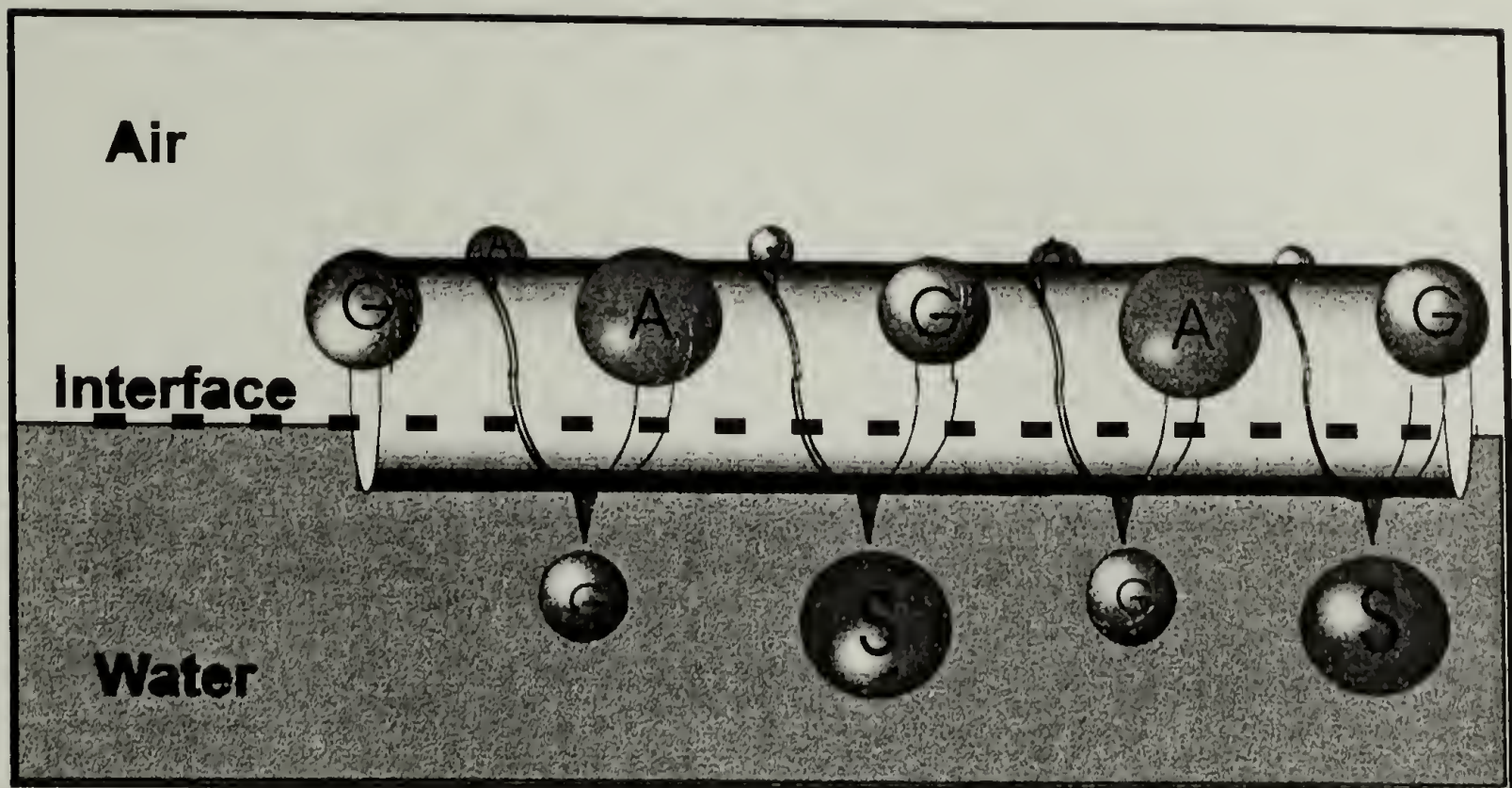
Figure 17: Hexapeptide fibroin sequence in a threefold conformation. The hexapeptide crystallizable repeat sequence in fibroin, (Gly-Ala-Gly-Ala-Gly-Ser), can act as a surfactant in the threefold conformation because the hydrophilic serine residues and hydrophobic alanine residues are separated to opposite sides of the helix.

Crystalline dimorphism in *Bombyx mori* silk fibroin has been an active topic of study for several decades[36,48,49,56]. There are two other known crystalline forms of *B. mori* silk fibroin, silk I and silk II. The conditions needed to form these two polymorphs have been well documented in the literature.[10,41,48,49,55,61,64] Silk II has an monoclinic unit cell with the protein chains in a pleated β–sheet (two fold "zigzag") conformation, and its structure is well understood.[36,48,49,56] The second polymorph, silk I, is not well characterized, because it has not been possible to obtain oriented samples.[36,48,49,56] Silk I appears to have an orthorhombic unit cell. The exact structure of silk I and its poly(*Ala-Gly*) analog may involve a "crankshaft" chain conformation.[1,72,74,134]

In an earlier paper, we reported the discovery of a third crystal structure of *Bombyx mori* silk, silk III, which has been observed in ultrathin films formed at the air-water interface.[103] Initial studies of the new structure indicated a trigonal symmetry, suggesting a threefold helical polyglycine II type of conformation for the silk chains. New data, subjected to a more refined analysis, from a larger population of samples suggest that the apparent trigonal symmetry is often slightly distorted, resulting in a monoclinic cell, but a three-fold helical chain conformation remains a key feature of the structure. Because the silk fibroin protein is composed of chiral L-form residues, the silk threefold helix is expected to be a left-handed helix. Relative intensities from unoriented diffraction data also support a 3/2 helical conformation. The approximate hexapeptide sequence of the crystallizable portions of *B. mori* silk fibroin,[1,2,20-22] (*Gly-Ala-Gly-Ala-Gly-Ser*)$_X$, results in favorable sterics for a threefold conformation. The combination of this hexapeptide sequence with the silk III, threefold helix places all of the serine residues on the same side of the helix, a situation which leads to a distortion from hexagonal packing of chains and imparts some sheet-like features to the more refined structure proposed here.[78,79,103,126,135] To date, a silk fibroin polymorph based on a 3/2 helical chain conformation has been observed exclusively in ultrathin free-standing films of silk obtained from the surfaces of dilute (<5%) aqueous silk solutions. For comparison, the concentration of silk in the *B. mori* silk gland is believed to vary from 15 - 30% fibroin as the silk progresses towards the spinneret. [10,11,39]

3.3 Experimental

3.3.1 Sample Preparation

Regenerated silk fibroin solutions were obtained as described in our previous paper.[103] Dialyzed regenerated fibroin 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 weight percent. A surface excess layer of silk then formed at the air-water interface. Ultra-thin silk films were deposited onto TEM grids using the Langmuir-Blodgett dipping technique, and also by placing droplets of dilute aqueous silk solution onto TEM grids as described previously.[103] In the latter case the films are referred to as "cast films". The solution dried to form an unsupported silk film spanning the holes in the TEM grid. Two sets of TEM samples were prepared:

1. LB silk films which remained on the trough for either 1 hour or 24 hours at 25°C and a surface pressure of 16.7 mN/meter before being deposited onto TEM grids.

2. Unsupported cast films were prepared from a 3 weight percent aqueous silk solution.

3.3.2 Characterization

The samples of LB silk and cast silk 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 fibroin crystal structure within the films. Low dosage techniques were

used to minimize beam damage. In order to further decrease the effect of beam damage, electron microscopy and diffraction was performed using a cryogenic sample holder at a temperature of approximately -160° C. The presence of a preferred crystalline orientation in the LB films, and the powder (random) orientation of the crystallites in the cast films were determined using pairs of diffraction patterns taken at two different tilt orientations between the sample and the incident beam. The experimental details of this tilting technique have been described previously.[103] Ten such *diffraction tilt pairs* were taken from different areas of sample 1 (LB films) at tilts of 0° and 25°. Five tilt pairs were recorded for each of the unsupported cast films in sample 2 (cast films) in order to demonstrate a consistent powder-like orientation.

Electron diffraction patterns were recorded on X-ray film (Kodak DEF-5) and analyzed using the trigonal, primitive unit cell structural models described in our previous paper[103] as a guide. The changes proposed for the monoclinic crystal structure represent only small modifications to the trigonal structure, and the general features of both diffraction patterns are the same, and thus recalculation of R using the monoclinic model is not justified. The models for the monoclinic structure were evaluated based on their ability to account for the subtler features of the unoriented, uniaxially oriented, and single crystal diffraction data in a consistent manner. Molecular mechanics minimization and energy calculation along with quenched molecular dynamics simulations were used in tandem with diffraction simulations to compare possible models for the features observed in the silk III crystallites. All these calculations were performed using the *Cerius* molecular modeling software from Molecular Simulations Inc.

The silk samples used in our studies were screened for purity using a combination of elemental analysis, electron diffraction, spatially resolved elemental mapping (energy dispersive X-ray spectroscopy – EDS – and parallel electron energy loss spectroscopy – PEELS), and radiation damage studies. This work is described in detail in an earlier paper.[103] Amino acid analysis of the purified regenerated silk solutions and of silk films collected off of the interface surface indicate that the solutions and films are composed of fibroin. Nickel, gold, and nylon TEM sample grids did not introduce impurities into the sample, and all of the quantitative diffraction data were obtained using nickel or nylon grids. There was some concern that the hexagonal single crystal diffraction patterns could be due to hexagonal ice deposited on the sample at cryogenic temperatures. However these diffraction patterns are present at room temperature as well as at cryogenic temperature, and the innermost set of six fold diffraction spots has a larger d spacing than the inner reflections observed for hexagonal ice. The hexagonal ice, occasionally observed as an impurity in the cryogenic diffraction studies, is accounted for in our analyses of the diffraction patterns.

Table 2: Angles between six-fold diffraction spots

Measurement of angles in six-fold single crystal spot patterns

RT Spot pair		1	2	3	Cryogenic	1	2	3
		60.2	60	59.7				
		61	60	59		58.5	60.5	61
		61	59.5	59.5		56.4	61.4	62.3
		60.7	60.3	59.1		57.9	60.7	61.4
		62.3	59.5	58.2		58.2	60.5	61.3
	average	61.04	59.86	59.1	**average**	57.75	60.78	61.5
	stddev	0.777	0.351	0.579	**stddev**	0.933	0.427	0.56
for best 4	avg	60.73	59.95	59.33	avg	58.2	60.57	61.23
	stddev	0.377	0.332	0.33	stddev	0.3	0.115	0.208

Table 3 Trigonal versus Monoclinic Unit Cell

Changes in Crystallographic spacings due to the Monoclinic Unit Cell

Monoclinic Unit Cell Indices	Monoclinic Unit Cell d spacings		Trigonal Unit Cell Indices	Trigonal Unit Cell d Spacings
101	8.21		---	---
$\overline{1}11$	7.88		---	---
102	6.55		---	---
110	5.37		---	---
103	5.17		---	---
111	5.16		---	---
112	4.66		---	---
200, 020	4.56		100, 010 $\overline{110}$	4.56
021	4.44		---	---
$\overline{2}20$	4.34		110	4.56
202	4.11		101	3.93
$\overline{2}22$	3.94		$1\overline{1}1$	3.93
203	3.69		---	---

3.4 Results and Discussion

Bright field TEM micrographs of silk LB films from sample 1, indicate that the films consisted of very small needle-like crystallites embedded in a featureless amorphous matrix. These mats of needle shaped crystallites result in rings of diffracted intensity, as described in our previous paper.[103] Additionally faint images of very thin lamellar crystals were sometimes visible in the LB films. The crystals associated with the silk III single crystal diffraction (spot patterns) are very beam-sensitive, requiring only 0.05 - 0.10 Coulombs/cm^2 at 200 kV for the diffraction spot pattern to completely disappear at cryogenic (-150° C) temperatures. Figure 18 shows six fold diffraction pattern produced by several of the silk III lamellar crystals, which were described in our earlier paper. In Figure 18 note the pair of faint reflections *inside* the innermost set of six-fold reflections indicated by arrows. A non-primitive crystalline unit cell is needed to account for these reflections. Cast films produced similar needle like crystalline morphologies to the LB films and similar diffraction patterns except for features due to the different orientations of the two types of films.[103]

Figure 18: Six fold diffraction patterns from LB films. Note the weak inner pair of reflections, indexed as the 1 $\overline{1}$ 1 and $\overline{1}$ 1 $\overline{1}$ using the proposed larger unit cell.
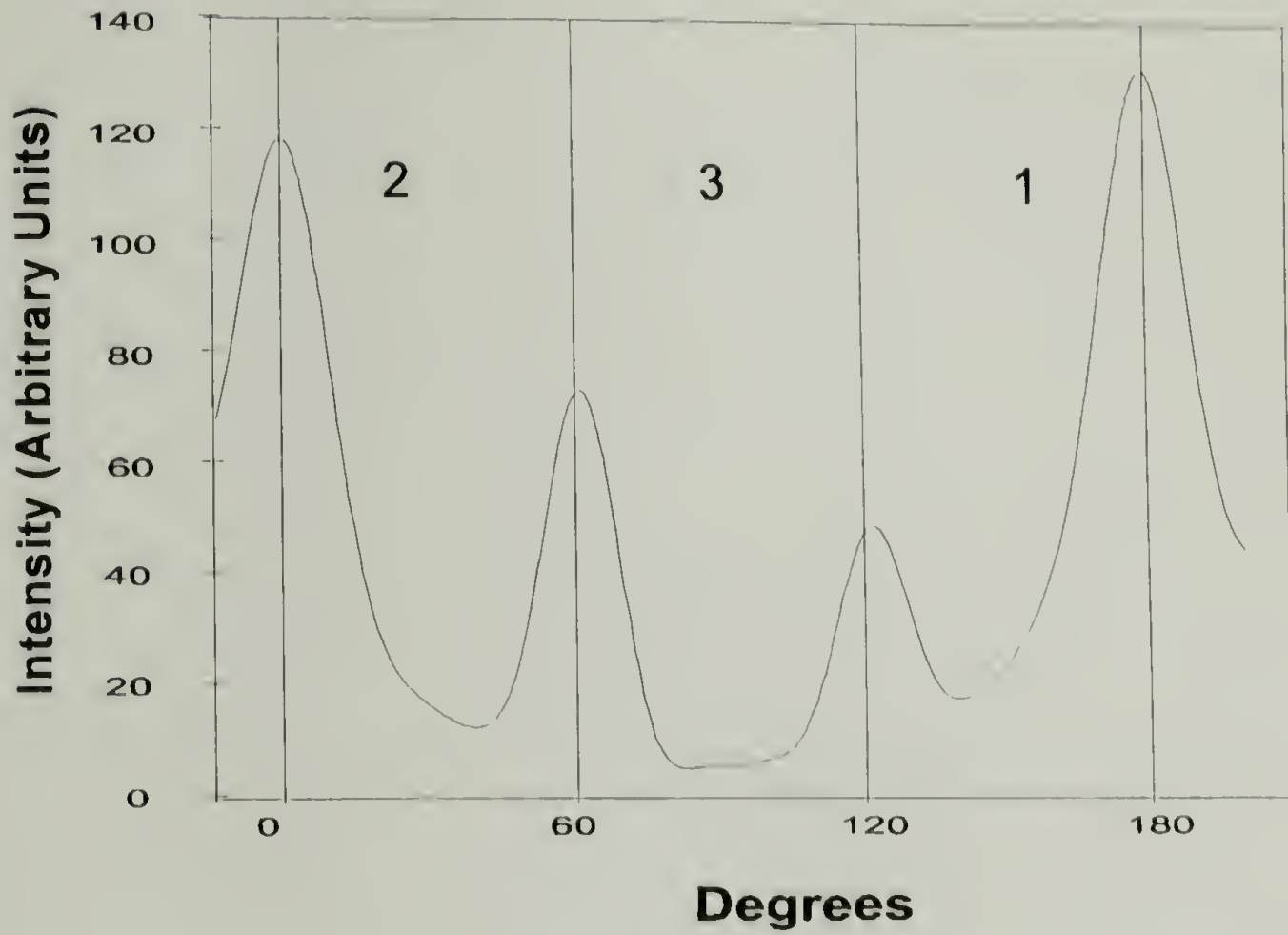
Figure 19: Single crystal densitometer scan. The angles made by the diffraction spots in a hexagonal single crystal pattern are not uniformly 60°.

A hexagonal unit cell for the lamellar single crystals is supported by the occurrence of single crystal diffraction patterns with hexagonal symmetry and by the fact that the six fold spot patterns are often superimposed on rings of diffracted intensity at the same d-spacing, from smaller crystallites of LB silk. The six fold diffracted spots always fall on top of LB silk film reflections from families of planes that would possess hexagonal symmetry in the trigonal unit cell. The small deviations from a perfectly hexagonal unit cell which are observed are not large enough to create an obvious distortion away from hexagonal symmetry in these spot diffraction patterns. Computer analysis of digitized diffraction patterns determined the azimuthal angles between the spots to an accuracy of 0.1°. Table 1 is a summary of the results of measurement of 9 room temperature and cryogenic single crystal, apparently hexagonal, spot patterns. This data indicates that the spots are approximately 2° off of a 60° separation at cryogenic temperatures, as is demonstrated in Figure 19, a scan from a typical cryogenic single crystal pattern. In the Figure, the vertical bars are spaced 60° apart and would intersect the centers of the Gaussian peaks obtained for the six-fold diffraction spots if the diffraction pattern were perfectly hexagonal. Interestingly, the observed distortion is not reproducibly apparent in six-fold single crystal diffraction patterns obtained at room temperature. Sheet-like ordering of serine residues can cause a deviation of γ from 120°, and thus slight differences in the azimuthal angles observed in single crystal patterns.

Tilted, uniaxially oriented diffraction data, and measurements of cryogenic single crystal diffraction patterns indicate that the six fold diffracted spots actually lie on two different, very closely spaced, diffraction rings. The small differences in six fold diffracted spot spacings that occur for one of the pairs of spots are consistent with 100

peak splitting, to be discussed below for electron diffraction powder patterns and uniaxially oriented LB film diffraction patterns. The presence of low angle spots, as in Figure 18, not accounted for in the primitive trigonal unit cell, indicates that the unit cell is also larger than initially supposed. These reflections can be accounted for using a larger unit cell that preserves the essential chain packing geometry. If the unit cell angle γ deviates slightly from 120° the 100 family of reflections is split into non equivalent components with slightly different spacing. When the trigonal unit cell is distorted by changing the angle γ, the (100) and (010) planes are still equivalent, but the spacing of the (1$\bar{1}$0) planes is affected by the distortion in the unit cell parameter. An aspect of the chain packing such as regular chain folds or preferred serine location and orientation could possibly account for the distortion away from true hexagonal packing, by creating a non-equivalent nearest neighbor packing direction. Thus the deviation of γ from 120° can explain the slightly different spacings observed for different 100 family diffraction spots.

3.4.1 Polycrystalline data

In Figure 20 an image of the innermost ring of an LB film diffraction pattern, recorded with a slow scan CCD camera, shows that this ring is actually composed of two closely spaced reflections, denoted with arrows. The slow scan image was obtained at 120kV and -167° C using a camera length of 1,200 cm and a Zeiss omega filter to remove the background inelastic scattering. The two peaks in Figure 20 result from splitting of the 100 reflection of the originally proposed trigonal cell when the angle γ is altered away from 120°. In the Miller indices of the non-primitive unit cell to

be proposed below, the components of this strong inner reflection result from the splitting of the 200 and the $2\bar{2}0$. Henceforth, we will refer to reflections by their non-primitive, monoclinic indices.

Tilted diffraction patterns, taken at a 25° tilt to the orientation axis, also reveal that the strong inner diffraction ring splits into at least two rings with slightly different radii, as seen in Figure 25. Since the ring in the untilted pattern, as shown in Figure 25, is broad and encompasses the spacings of the component rings observed when the sample is tilted, it is reasonable to assume that the innermost ring is in fact compound.

The shape of the strong inner peak in the fibroin thin film diffraction data, requires a number of Gaussian components to be fit well. Figure 21, is a plot of intensity vs. inverse d-spacing, produced by integrating a representative digitized ring pattern around the azimuthal angle. Component reflections are present on the outer edge of the strong inner peak; they are well-resolved on the negatives but appear as shoulders on the densitometer scan due to a broad amorphous scattering peak in that region. The most apparent of these reflections are indexed as the 203 and the $2\bar{2}2$ using the monoclinic unit cell. Component reflections indicated by the arrows in the Figure are always present and reproducible across a large number of diffraction patterns (7 unoriented patterns and 16 uniaxial LB patterns). Single crystal data clarify the contributions of the component peaks in split 200 reflections from uniaxially oriented LB films. Single crystal spot patterns obtained at cryogenic temperatures consist of reflections with close d spacings specifically the 200, 020, and $2\bar{2}0$. The measured d spacings for the single crystal spot patterns match the spacings obtained for the set of overlapped maxima in

azimuthally integrated uniaxially oriented, polycrystalline diffraction patterns. The angular deviation in the six-fold single crystal diffraction spot geometry also agrees with the distortion in the unit cell parameter γ necessary to obtain the closely spaced reflections observed in Figure 21.

Further evidence for the existence of a distorted modification of the silk III structure is provided by preliminary data on the structure of fibroin films formed at the aqueous -chloroform, liquid-liquid interface. The diffraction pattern obtained from one of the fibroin films formed at this interface is shown in Figure 22. In these samples the crystallites are large enough so that the inner component rings can be clearly distinguished. These rings are indicated by arrows in the Figure and their Miller indices with respect to the monoclinic non-primitive unit cell are given. The d spacings measured for the rings in the aqueous-chloroform interface diffraction pattern closely match those found at the air-water interface for silk III. The fibroin films formed at the water-chloroform interface were observed to sink to the bottom of the more dense chloroform phase, implying a density greater than 1.54. This density is higher than expected for either silk I or silk II, but is commensurate with the densities of approximately 1.6 calculated for both the trigonal silk III model and the non-primitive monoclinic model.
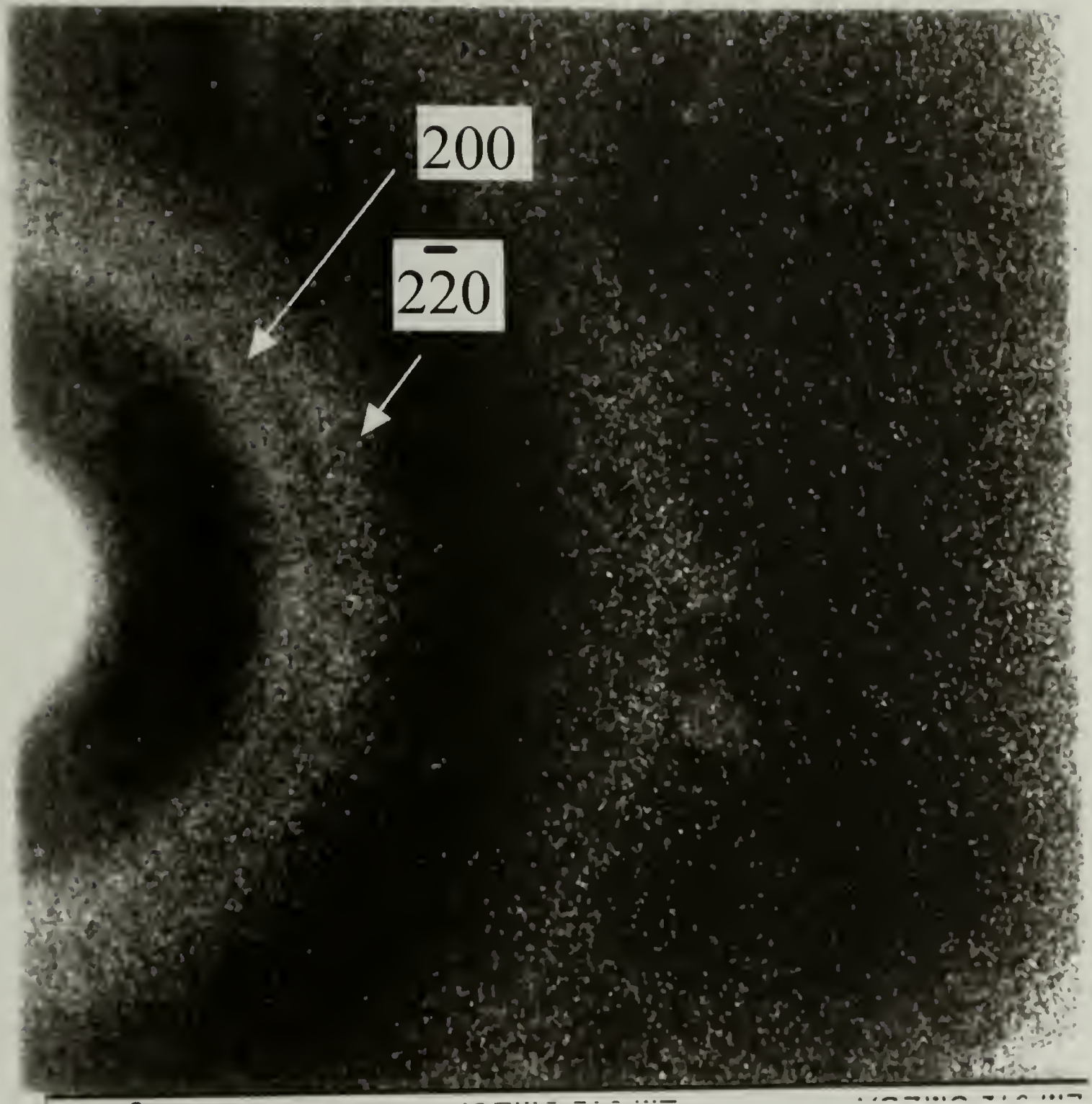
Figure 20: Slow scan image of a silk LB film diffraction pattern showing the closely spaced component rings in the inner (4.5 Å) reflection. These reflections are indexed using the monoclinic unit cell as the 200 and $2\bar{2}0$. The camera length used in the slow scan diffraction pattern is 1200 cm.
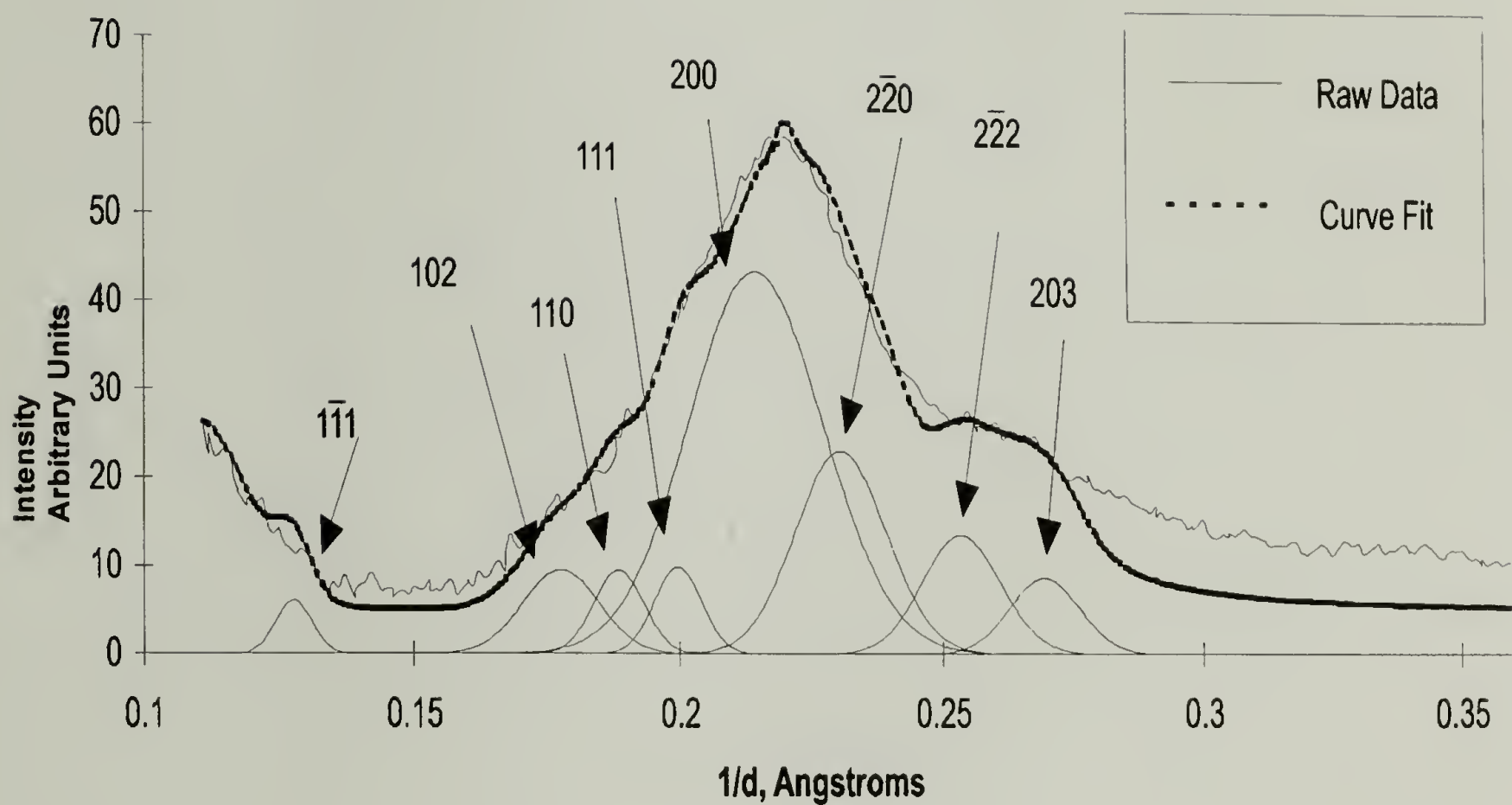
Figure 21: Densitometer scan of a typical diffraction pattern from a uniaxially oriented LB film. Many reflections such as the 111, 203, and 102 cannot be indexed using the trigonal model.
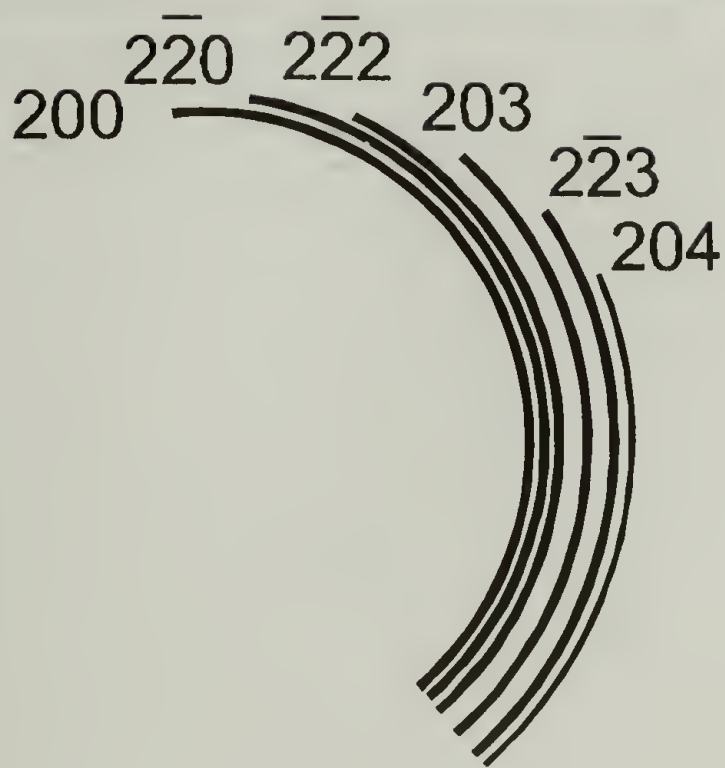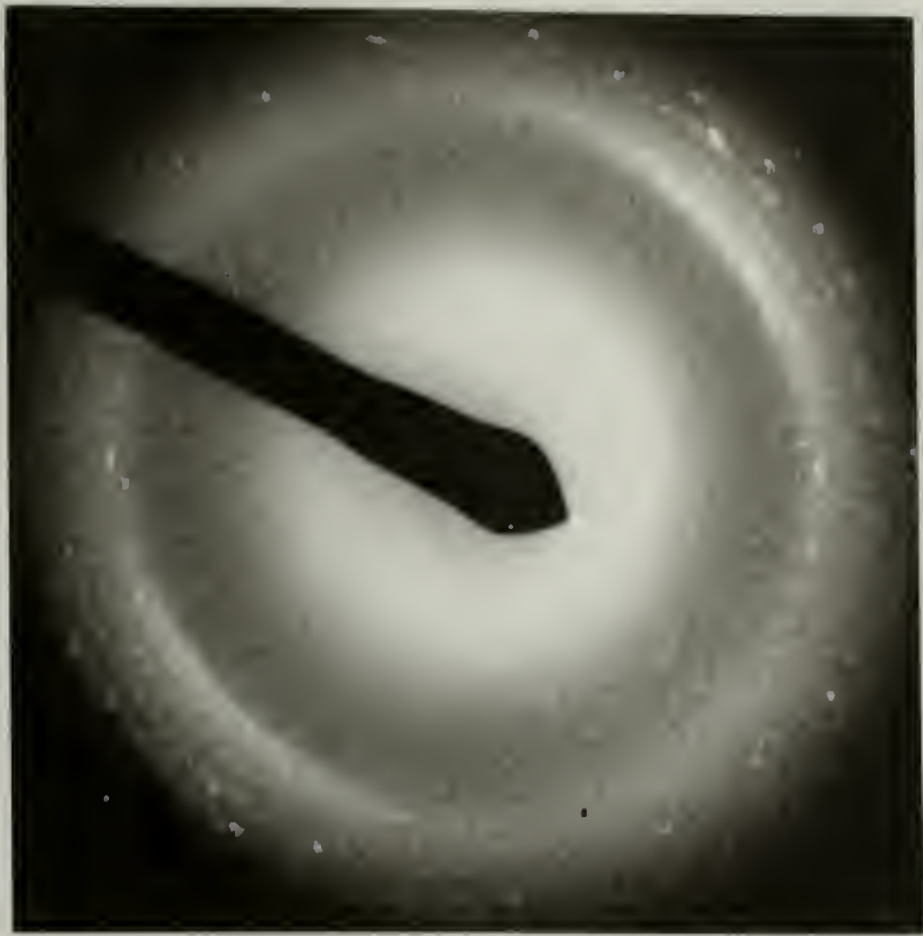
Figure 22: Polycrystalline diffraction rings from the chloroform interface. These rings can be indexed using the monoclinic modification of the silk III unit cell, but cannot be indexed using the original trigonal model.

3.4.2 Refinements to the Basic Hexagonal Silk Crystal Structure

The lattice spacings, symmetry, and peak intensities in diffraction patterns of fibroin crystallized at the air-water interface are similar to those found previously by Keith and Padden[136] for polyglycine II. Thus polyglycine II was used as an initial structural motif when constructing the threefold helical chain conformation. Initial models described previously[103], with a primitive unit cell, had dimensions of $a = b = 5.1$ Å, $c = 9.42$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$ for a unit cell containing a single turn of a single threefold helical chain. These dimensions are similar to the dimensions of the polyglycine II unit cell allowing for a small lateral expansion in the a and b directions to accommodate the alanine and serine residues in silk. Simulated diffraction from these primitive, trigonal models was compared with the curve fitting results from the experimental data, and a very reasonable reliability index, R, of 0.11 was obtained. The trigonal, primitive unit cell model describes the general form (major reflections) of the data very well and probably is a fair representation of the crystal structure for a significant portion of the crystallites. It is likely that crystals with the trigonal crystal structure described in our earlier paper coexist with crystals having a slightly more ordered molecular packing scheme and monoclinic unit cell symmetry. The ordered packing scheme proposed for this monoclinic silk III structure incorporates slightly different nearest neighbor packing distances to account for the difference in the sterics of the glycine-serine side of the helix as compared to the alanine-glycine sides.
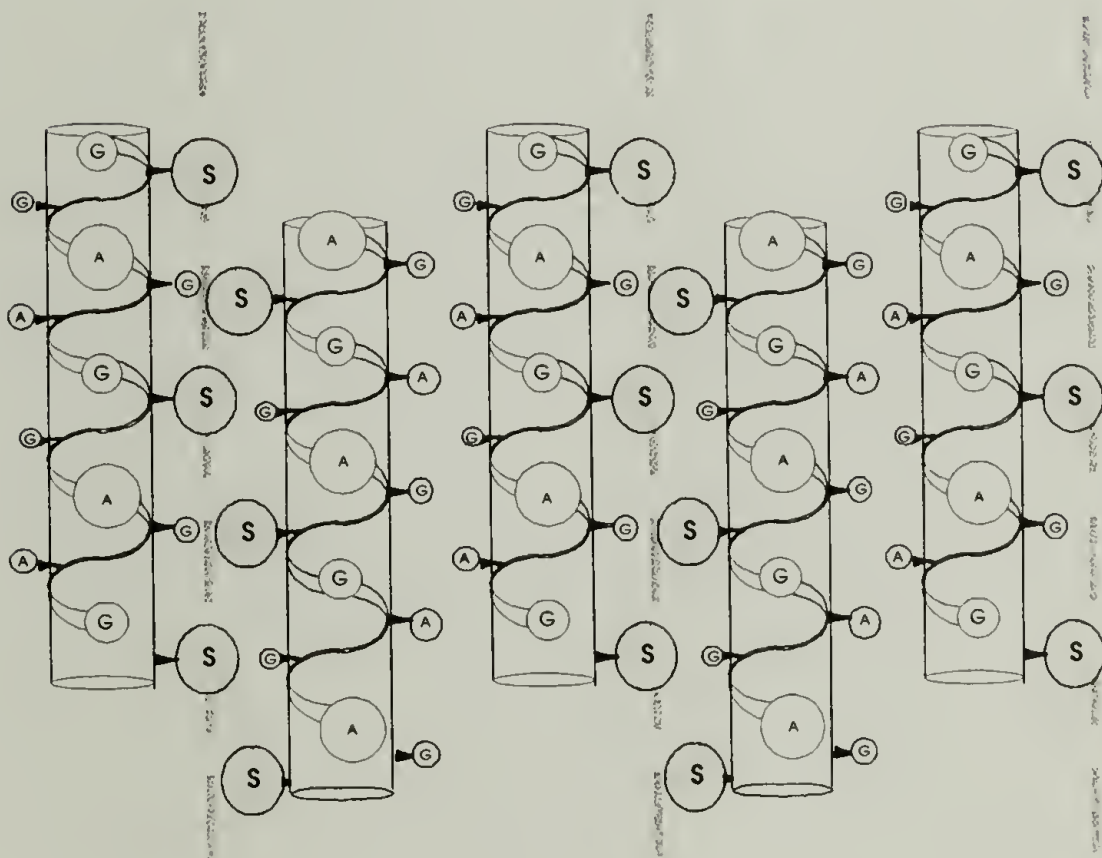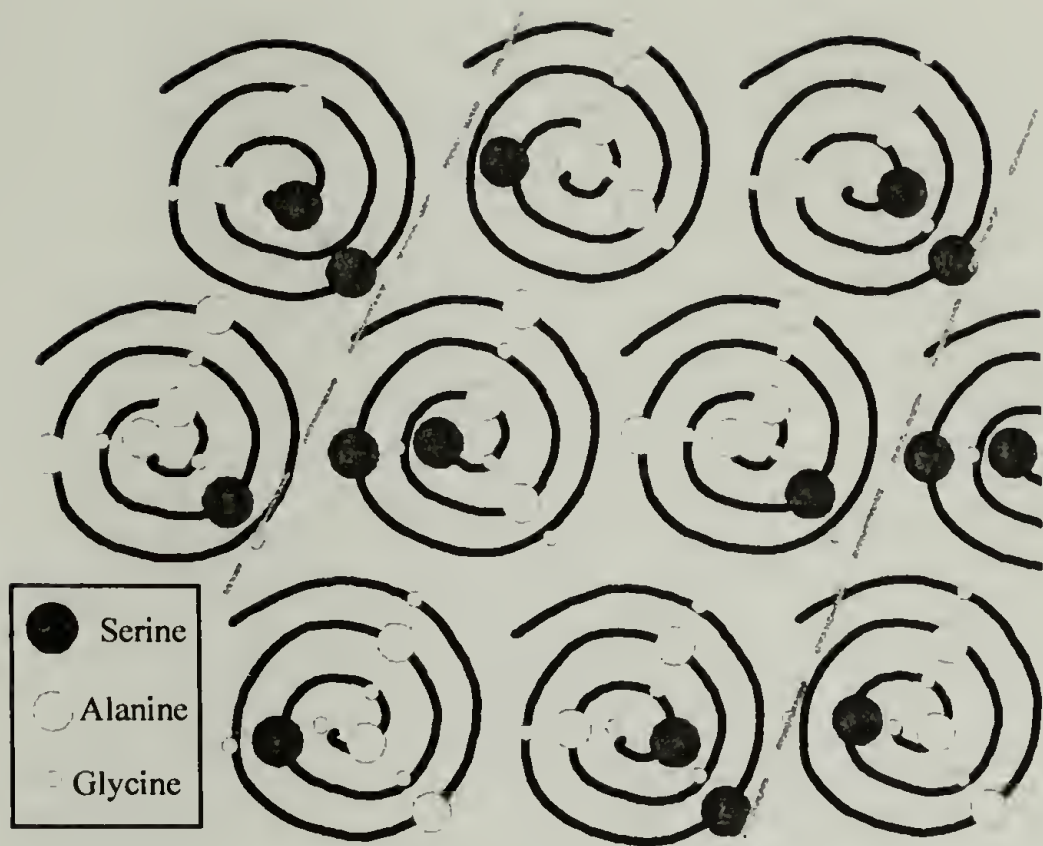
Figure 23: The combination of the six residue repeating sequence poly(*Gly-Ala-Gly-Ala-Gly-Ser*) (representing the crystallizable component of silk fibroin) with a threefold helical conformation, results in a column of alternating glycine and serine residues parallel to the helical axis. The alternation of glycine and serine allows closer packing of helices because the glycine residues act as spacers. (a) Top view (b) side view; In both views the dashed gray line shows the location of the clustered serine residues.

Simulated diffraction patterns from model structures containing sheets of serine and a slight distortion from perfect hexagonal packing, leading to a monoclinic (but almost trigonal) unit cell, provide a very good match to the observed experimental diffraction. These models account for the previously unexplained features of the diffraction patterns. The increased size of the resulting non-primitive unit cell (over a primitive cell), due to the regular arrangement of residues, results in the additional lower angle diffracted reflections observed, but not predicted for the primitive trigonal unit cell. Two low angle spots can be seen in Figure 18. They occur at 7.8 Å, and can be indexed using the distorted non-primitive cell as the $1\bar{1}1$ and $\bar{1}1\bar{1}$ reflections, an assignment which is consistent with their d spacing, their intensity, and the geometry of the diffraction pattern. There is no indexing assignment possible using the smaller trigonal primitive unit cell.

The combination of the six residue repeating sequence poly(*Gly-Ala-Gly-Ala-Gly-Ser*) and the 3/2 helix leads to placement of the serine residues on one side of the helix, as can be seen in Figure 23a and Figure 23b. In the Figure it is clear that one third of the cylindrical surface described by the helix is rich in serine residues. The other two thirds of the surface contain columns of vertically alternating glycine and alanine residues and are equivalent. Thus the distortion of one third of the nearest-neighbor distances between molecular helices may indicate a preferred orientation for serine within the crystalline unit cell. An arrangement that clusters the bulkier more hydrophilic serine residues should also be favorable in terms of energetic and packing considerations. Regular chain folding would also distort one of the three nearest-neighbor distances, but the distortion due to chain folds is expected to be much smaller

than that observed for LB silk. Attempts to model the slight distortion from hexagonal symmetry focused on the arrangement of serine. Such models result in a sheet-like arrangement of threefold helical molecules that cluster the serine residues in the $(1\bar{1}0)$ planes.

The polypeptides that are most comparable to silk incorporate a regular alternation of glycine and larger residues, often as a repeated *di* or *tri* peptide sequence. Among the silk-like polypeptides, the occurrence of a crystalline polymorph that incorporates a 3/2 helical, conformation is fairly common.[79-83] This polymorph occurs for copolypeptides with more than 50% glycine and appears to be more common in poly*tri*peptides such as poly(*Gly-Gly-Ala*).[79-83] Although *B. mori* silk contains almost exactly 50% glycine in its crystallizable segments, making it poorer in glycine than these polytripeptides, the strict alternation of glycine with the larger residues results in favorable sterics if the silk takes on a 3/2 helical conformation.[103] In the polytripeptides that were found to crystallize in a threefold helical structure, the repeated residue sequence alters the packing of the polypeptide helices away from the perfect hexagonal packing reported for polyglycine II. In the polytripeptides, the threefold helical conformation results in the different residue types lining up "vertically", i.e. in the direction of the chain axis, with residues of the same type. For example, a polytripeptide with a repeating sequence of [*Ala-Gly-Pro*]$_x$, results in a threefold helical conformation with a column each of alanine, glycine, and proline running down the helix. The resulting polypeptide helix will have bulkier side groups in a column parallel to the helical axis.[78,79,125,126,135,137] In the poly*tri*peptide case there is no alternation of large and small residues within the column parallel to the chain axis, and no spacer

residues to decrease the packing distance. In response to the different sterical requirements for the different sides of the helix, the threefold polytripeptide helices form hydrogen-bonded sheet-like structures, with the bulky residues accommodated between the sheets.

A sheetlike arrangement in crystals of the silk III threefold helical conformation results in a non-primitive unit cell. While a silk III sheetlike structure should be similar to the structures found in the polytripeptides there are some distinct differences. In a threefold conformation of silk fibroin, while the serine residues also line up along the same edge of the helix, they alternate with small glycine residues which act as spacers. The regular alternation of bulkier residues (alanine and serine) with glycine along the helical axis and as nearest neighbors in the chemical repeat sequence eases packing difficulties in the 3/2 helix. Hydrophilicity and hydrophobicity may play a role in structure determination in addition to sterics, because the threefold conformation separates the hydrophilic serine residues from the hydrophobic alanine residues. In order to maximize interactions involving serine, the fibroin threefold helices pack so that each serine-glycine column faces a serine-glycine column from a neighboring helix. A preferred packing, or serine-rich crystallographic direction, is needed to capitalize on the energetics of the threefold conformation in a crystalline array. The resulting sheetlike structure has the serine residues lying between the sheets, as depicted schematically in Figure 23. The intersheet distance is smaller than in the polytripeptide threefold helical crystals because of the glycine spacers and the relatively small size and high flexibility of the serine residue. A crystal with this sheetlike structure does not

significantly disrupt the three-dimensional hydrogen bonding patterns that exist for a perfectly hexagonal polyglycine II - like packing.

Besides the additional reflections due to the distortion of the unit cell and incorporation of more than one chain, certain of the component reflections which make up the fine structure of the first large maximum can only be accounted for if the unit cell is long enough to encompass six residues in the c direction rather than the three residues originally supposed. The reflection referred to as the $1\bar{1}1$ (Figure 18) can only be indexed using a unit cell with a six residue periodicity in the c direction. This reflection would have non-integer indices in a primitive unit cell. In order to have a six residue periodicity there must be a non-random arrangement of alanine, glycine and serine residues in the (006) planes. In other words, the hydrogen bonding between neighboring chains must show preferences for bonds between particular types of residues as in the proposed sheet structure.
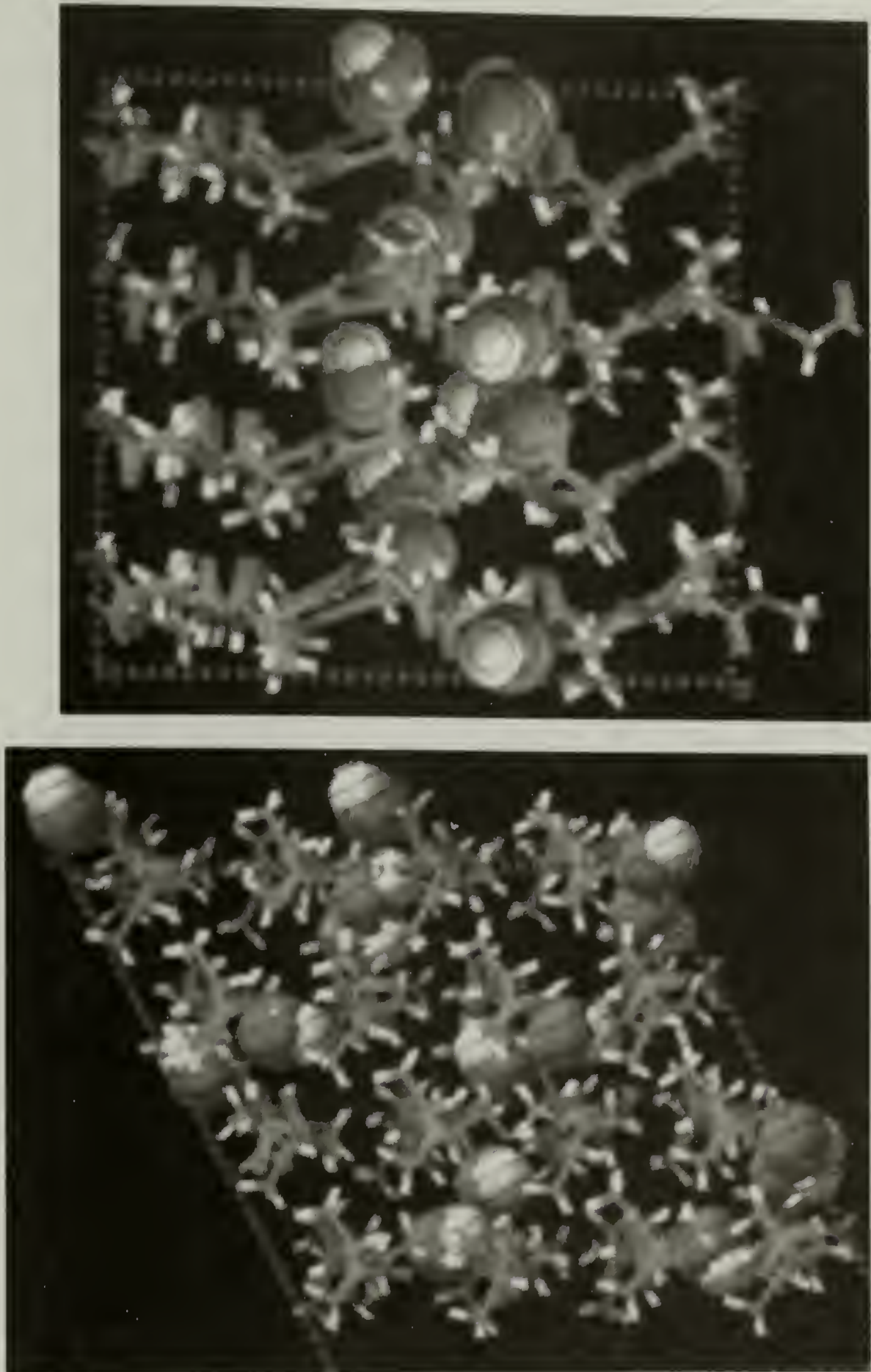
Figure 24: Molecular model showing sheet-like arrangement of serine residues in the distorted silk III modification. (a) View looking down the c (chain) axis (b) View looking down the b axis.

Models for the nonprimitive, monoclinic silk III structure were constructed, and energy minimization was carried out using the Dreiding II forcefield with H-bonds and Van der Waals forces considered in the calculation. Different chain packings and residue placements were investigated, but the sequence *Gly-Ala-Gly-Ala-Gly-Ser* was always preserved for each chain. The unit cell parameters used in the simulation were $a = 10.2$ Å, $b = 10.2$ Å, $c = 19$ Å, $\alpha = \beta = 90°$, $\gamma = 116.5°$. These result from measurements of experimental diffraction data and reflect the fact that a four chain unit cell was used: The Miller index assignments of experimentally observed silk III diffraction maxima based on this non-primitive, monoclinic unit cell are listed in Table 3, along with the initial assignments from the primitive trigonal unit cell model. Two chains each in the a and b directions and two full turns of the helix in the c-direction are included in the nonprimitive, monoclinic unit cell. This four chain unit cell has alternating up and down chain orientations along the b axis to simulate the possibility of chain folding. Serine residues are clustered along the $(1\bar{1}0)$ planes, as shown in Figure 24a and Figure 24b. The serine oxygens have been emphasized in the Figure for clarity. The planar arrangement of serine residues is most apparent from the view looking down the chain axis, Figure 24a. The three-dimensional clustering of serine residues is evident in Figure 24b, a view down the b-axis.

An energy of -6.6 Kcal/mole · residue was calculated for the threefold helical structure using the unit cell described above. For comparison, the energy calculated for the well-established silk II structure is -5.5 Kcal/mole · residue using the Dreiding forcefield. The more favorable energy for silk III reflects a small difference that could well be attributed to an artifact of the force-field parameterization, but the similarity of the values gives us confidence in our structure. A series of dynamics simulations each

followed by energy minimization was used to relax the structure. The resulting helices were slightly irregular, with the glycine residues before serine in the residue sequence (running from amino terminal to carboxyl terminal) adopting higher ψ torsions and lower ϕ torsions than the average ϕ and ψ of -91° and 156° respectively, to compensate for the lower angles adopted by the bulkier serine residue. In Table 3 the average torsional angles are listed for each residue in the six residue sequence for fibroin, and it can be seen that the angles of the glycine residues adjacent to serine adjust to accommodate the steric bulk of the larger residue and any additional distortion due to the extra hydrogen bonds made by serine. The average backbone hydrogen bonded distance was 2.07 Å between the carbonyl oxygen and amide hydrogen, with COH and OHN angles of 155° and 161° respectively. Out of the 96 possible interchain H-bonds per unit cell between carbonyl and amide groups on the main chain 24 were not formed. However each of the 16 serine groups in the model formed *at least* 1 interchain H-bond, so little or no overall stabilization due to H-bonding was lost.

3.4.3 Tilting Experiments with Uniaxially Oriented Samples

The crystallites in the fibroin LB films have a uniaxial orientation, like a sedimented mat, which results in clear changes in the diffraction pattern when the samples are tilted. Two orientations were compared. Simulated diffraction patterns were obtained at 0° and 25° tilt of the incident electron beam direction to the axis of uniaxial orientation (normal to film) in order to make a comparison with the *diffraction tilt pair* data from the LB silk films. The experimental and simulated diffraction patterns are compared in Figure 25a and Figure 25b for the monoclinic model unit cell. In the experimental diffraction pattern at 25° tilt and the simulated diffraction, also at 25° tilt, the 200, $2\bar{2}0$,

202, and $2\bar{2}2$ reflections (indexed with respect to the monoclinic cell) are indicated by arrows. The simulated pattern from the trigonal model looks similar to the real pattern at 0° tilt (not shown), but at 25° tilt, as shown in Figure 25c, the 202 reflection shows a marked difference in the distribution of intensity with respect to that observed for the experimental data and the simulation using the monoclinic cell. In the experimental diffraction pattern and the monoclinic simulation the arced 202 reflection covers a considerably larger azimuthal angle (looks almost ring-like) with an intensity maximum rotated by 90° to the arcs observed in the trigonal simulation. The difference in the appearance of the 202 arc between the experimental (monoclinic simulation) and the triclinic simulation is a result of the splitting of the 202 and 200 reflections in the monoclinic structure. The 202 is split into non-equivalent 202 and $2\bar{2}2$ reflections, due to the deviation of γ from 120°. Splitting of the 200 into non-equivalent 200 and $2\bar{2}0$ reflections results in overlap of the 202 and $2\bar{2}0$ reflections and contributes intensity to the 202 ring.
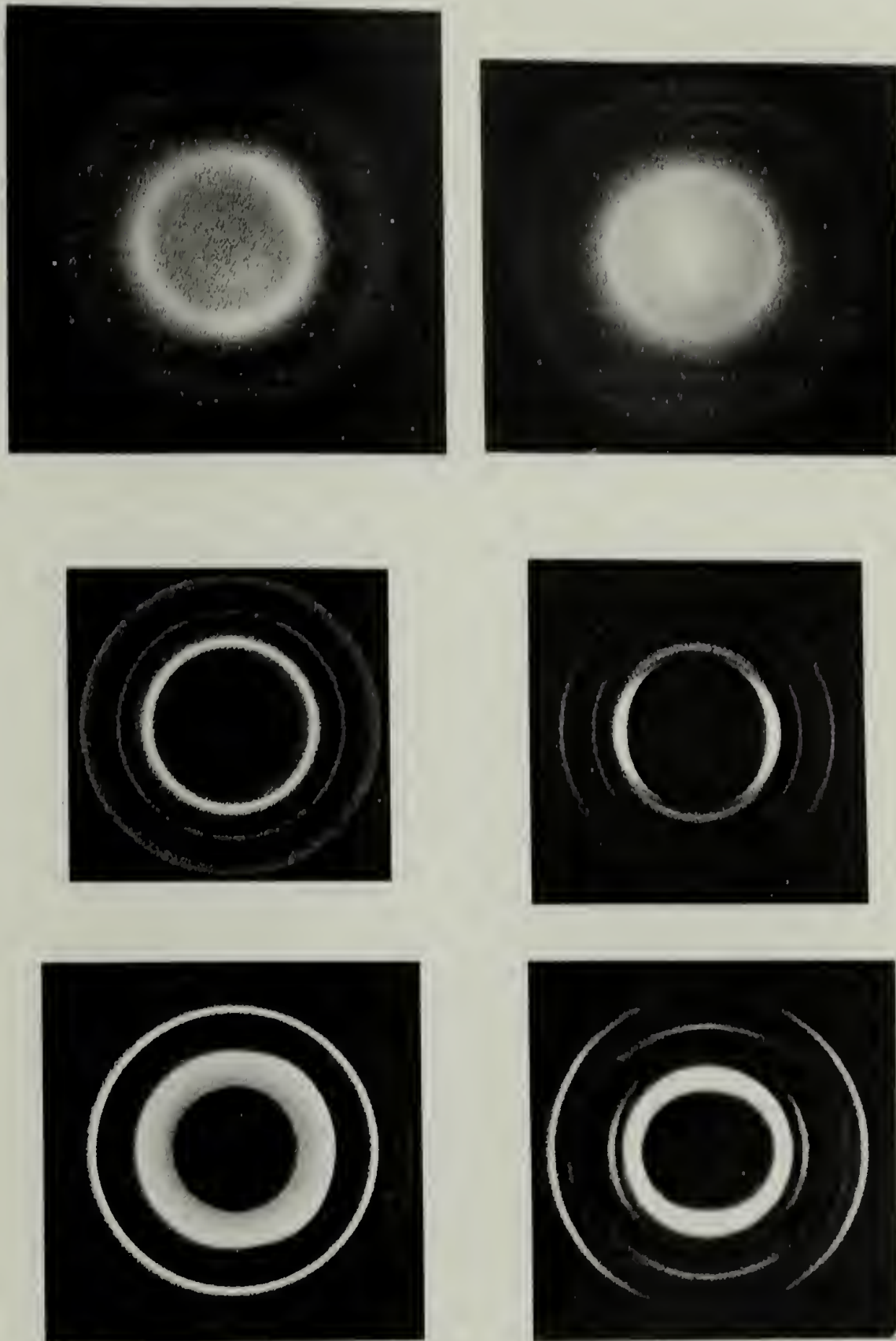
Figure 25: Close up of the inner rings in a tilted uniaxially oriented diffraction pattern, demonstrating the "splitting" of the 100 reflection when the unit cell is enlarged and distorted. In the untilted patterns these reflections appear as a single ring centered at 4.5 Å, whereas in the component arcs seen in the tilted pattern have *d* spacings of 4.3 and 4.7 Å. (a) Experimentally observed diffraction from a uniaxially oriented sample. The untilted pattern is on the left. The pattern on the right has been tilted 25°. (b) Simulation of tilted diffraction from monoclinic model showing splitting of rings (c) Simulated tilted diffraction for the primitive trigonal model fails to reproduce the splitting of the 4.5 Å ring.

93

Table 4: Variation in Torsional Angles

Measurement of average torsional angles
for the silk fibroin crystallizable hexapapetide
sequence (GAGAGS). The sequence runs from
the amino to the carboxyl terminal.

Residue	Psi	Phi
Glycine	86.3	-15.5
Alanine	86.2	-23.4
Glycine	97.7	-17.0
Alanine	72.7	-19.7
Glycine	109.4	-42.7
Serine	81.1	-23.1

3.5 Conclusions

The crystalline diffraction data from the interfacial (air-water interface) films can be fit very well by assuming a trigonal unit cell containing a single helical turn. This simple model requires average or randomized residue identities along the chain. Some of the weaker features of the diffraction data cannot be fit using the trigonal model and require a slightly modified unit cell that retains the essential features of the packing, a hexagonal or nearly hexagonal array of chains is still used. However the modified cell is non-primitive, making it larger, and also incorporates a slight distortion in the unit cell angle γ, lowering the unit cell symmetry from trigonal to monoclinic. A hexagonal crystal habit is observed for the silk III lamellae.

CHAPTER 4

ORIENTATION EFFECTS IN SILK FIBROIN LANGMUIR-BLODGETT FILMS : EVIDENCE OF SURFACTANCY

4.1 Summary

A threefold helical crystal structure of Bombyx mori silk fibroin has been observed in films prepared from aqueous silk fibroin solutions using the Langmuir-Blodgett (LB) technique. The films were studied using a combination of transmission electron microscopy and electron diffraction techniques. Films prepared at a surface pressure of 16.7 mN/meter have a uniaxially oriented crystalline texture, with the helical axis oriented perpendicular to the plane of the LB film. Films obtained from the air-water interface without compression have a different orientation, with the helical axes lying roughly in the plane of the film. In both cases the *d*-spacings observed in electron diffraction are the same and match a threefold helical model crystal structure, silk III, described in the previous chapters. Changes in the relative intensities of the observed reflections are consistent with the different oriented crystalline textures proposed. In addition to diffraction evidence, there are some differences in the morphology images of the films, most notably at the edges of tears in the films. In the films that have an orientation that places the helical axis in the plane of the film, there are fibrous objects spanning the tears in the film. In the films with a helical axis orientation perpendicular to the plane of the film, the edges of rips in the film are clean. These observations are

consistent with the helical orientations proposed, based on the electron diffraction data obtained.

4.2 Introduction

A new crystal structure of Bombyx mori silk fibroin was observed in films taken from the air-water interface of aqueous fibroin solutions.[102-104,107,108] The crystal structure, silk III, involves an approximately hexagonal packing of silk molecules in the left handed threefold helical chain conformation.[103,104] Because this conformation separates the serine and alanine residues, creating a hydrophilic column of residues parallel to the helical axis, surfactant behavior of fibroin at the air-water interface is believed to play a role in selecting the conformation and subsequent crystal structure at the interface. However, much of the data used to characterize the structure was for LB films compressed to 16.7 mN/meter. These films possess a uniaxial crystallite orientation similar to a sedimented mat, with all of the crystallite c (or chain) axes oriented perpendicular to the plane of the film. If the fibroin is behaving as a surfactant, assuming a threefold helical conformation that separates its hydrophobic and hydrophilic residues, the chain axis of the fibroin molecules and of any resulting crystallites should lie in the plane of the film. While it is possible that compression on the Langmuir trough could reorient whatever structures are present at the interface, data on the orientation of uncompressed films are needed for comparison.
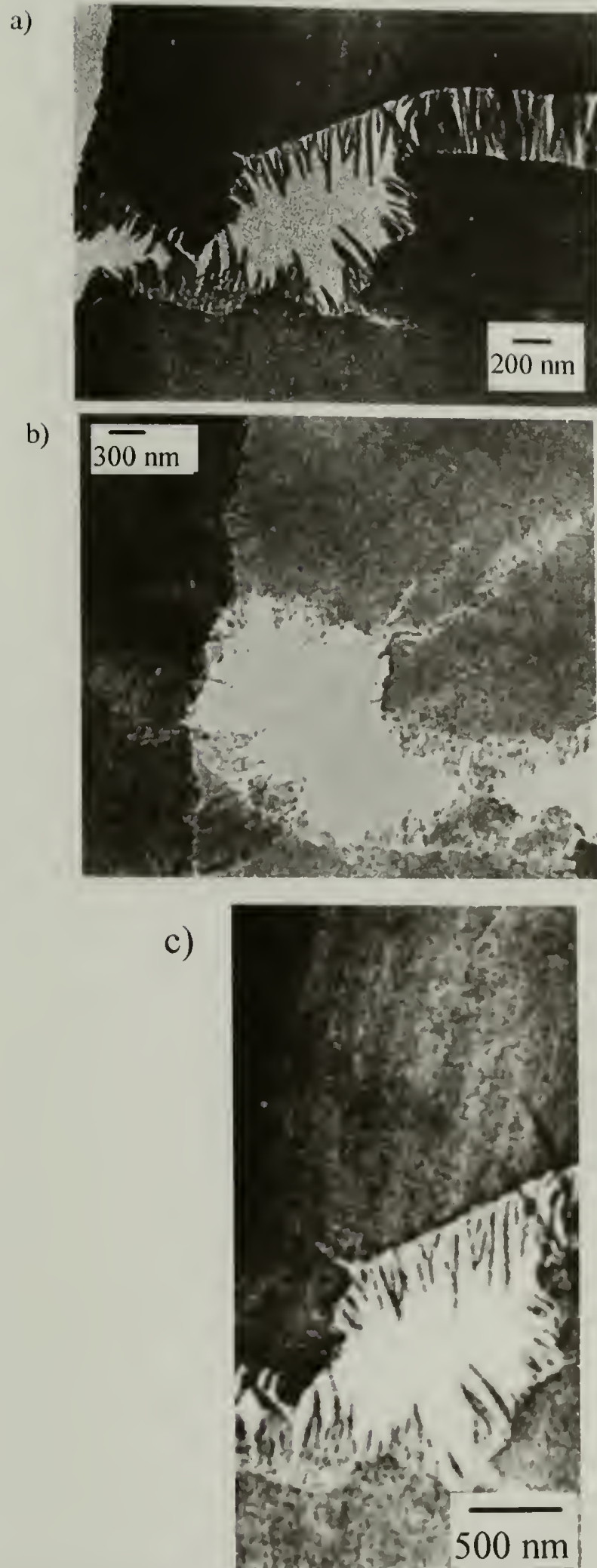
Figure 26: Uncompressed film morphology (a) Propagating crack spanned by fibrillar objects (b) Edges of continuous film regions have fibrils protruding from them (c) Close up of a region in *(a)* showing dark bands leading up to the protruding fibrils

Solid film formation has been observed previously on uncompressed surfaces of aqueous fibroin solution, but only a cursory examination of the morphology with no supporting diffraction data was reported.[138] Electron diffraction and TEM morphology data have been obtained for uncompressed surface films, and TEM morphology images of uncompressed films are shown in Figure 26. These data clearly indicate a silk III crystal structure with the helical axes lying in the plane of the film. The relative intensities in the diffraction patterns are altered in a systematic fashion, consistent with this orientation. In addition there are consistent morphological differences between the LB films compressed to 16.7 mN/meter and the uncompressed films. The presence of fibrils in the uncompressed films, spanning cracks and protruding from the edges of the film is evidence for an in plane orientation for the silk III helical axes. Edges and cracks in the LB films compressed to 16.7 mN/meter are clean. The oriented crystalline textures observed are very uniform, and can be controlled by varying the surface compression treatments used. A variety of well-controlled oriented film textures are possible, making these films potentially useful as templates to grow other protein crystals.

4.3 Experimental

Bombyx mori cocoons were degummed using boiling water and $CaCO_3$ to remove the sericin, leaving pure fibroin. Amino acid analysis has been used to assess the protein composition of fibroin prepared in this manner and no sericin was detected. The degummed fibroin was rinsed thoroughly with distilled water and dissolved in a 9.1 M solution of LiSCN in water. In order to remove the salt, the fibroin and LiSCN solution was then dialyzed against frequent changes of distilled water for several days. The dialyzed fibroin solutions were filtered to remove dust and any protein precipitate. Five different film treatments were used at the air water interface to create different crystallite orientations within the films.

1) Aqueous solutions of silk fibroin were left standing undisturbed (uncompressed) until solid skins formed at the surface. These films were collected onto TEM grids.

2) Aqueous silk solutions were placed into an Langmuir trough. After a surface excess layer had formed and stabilized, the surface was compressed to 16.7 mN/meter and held at that pressure for times ranging from 1 hour to 1 day.

3) An LB experiment was performed as in (2) except the surface was compressed to 34 mN/meter.

4) LB experiments were performed where the surface film was allowed to age uncompressed for 1 day prior to being compressed to 16.7 mN/meter and held at that pressure overnight.

5) Dilute solution (~1%) was dropped onto uncoated grids. The grids were dried on filter paper. These are referred to as cast films.

The fibroin thin film characterization and diffraction data analysis techniques used have been described the previous two chapters, and have been published previously.[103,104]

4.4 Results and Discussion

The silk films on the Langmuir trough and the uncompressed silk films are systems dominated by surface phenomena. The regenerated silk solutions used to prepare our thin films exhibit a marked tendency to foam, indicating surfactant behavior, and the residue sequence in the crystallizable regions of *B. mori* fibroin suggests a mechanism for the surface activity of this molecule. The six residue repeating sequence of silk, *Gly-Ala-Gly-Ala-Gly-Ser*, in a threefold helical conformation creates an "edge" of the helix that is somewhat hydrophilic, while the rest of the helix is hydrophobic. This allows the chains to behave as polysurfactants: The hydrophilic, hydroxy groups on one side of the helix interact with the water subphase and hydrophobic side chains on the other side of the helix protrude into the air. Thus we conjecture that the selection of a threefold helical structure at the interface is driven by hydrophilic and hydrophobic interactions at the air-water interface. In such an interface-assisted model of silk III crystallization, once a monolayer of 3/2 helical silk chains organizes at the interface, crystallites could nucleate and grow from the water subphase based on the symmetry patterns initially established.

Table 5: Relative intensity changes with different orientations

Index	Angle to [100]	Angle to [001]	d spacing	Cast	Uncompressed	16.7 mN/meter
111	54	43	5.16	19	49	7
112	69	25	4.66	9	42	5
200,020	——	90	4.56	100	100	100
021	68	47	4.44	4	49	n/obs
2$\bar{2}$0	——	90	4.34	10	14	15
202	62	28	4.11	15	32	8
2$\bar{2}$2	76	28	3.94	15	23	6
006	90	——	3.10	25	55	n/obs
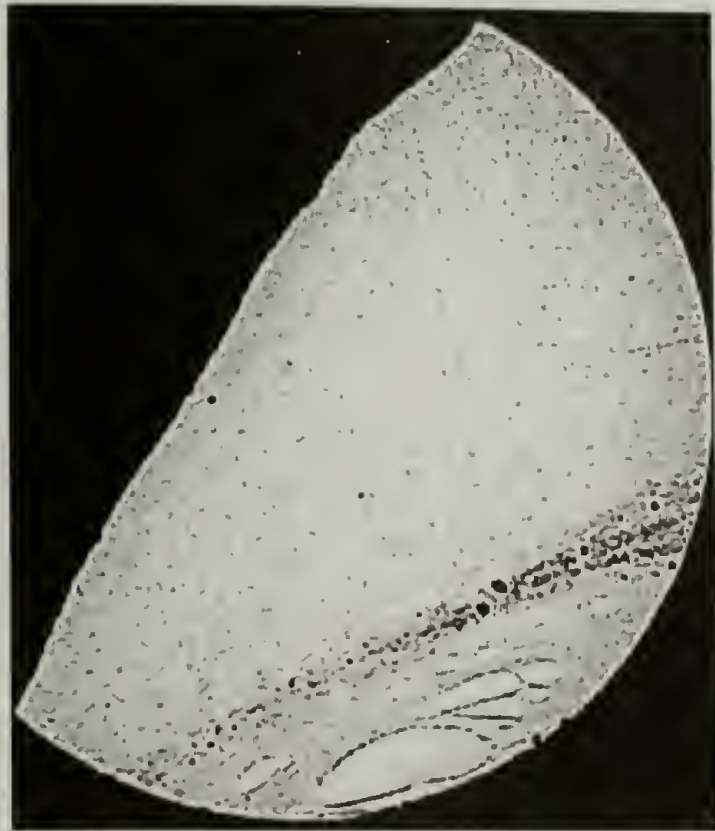
Figure 27: LB film morphology (16.7 mN/meter)
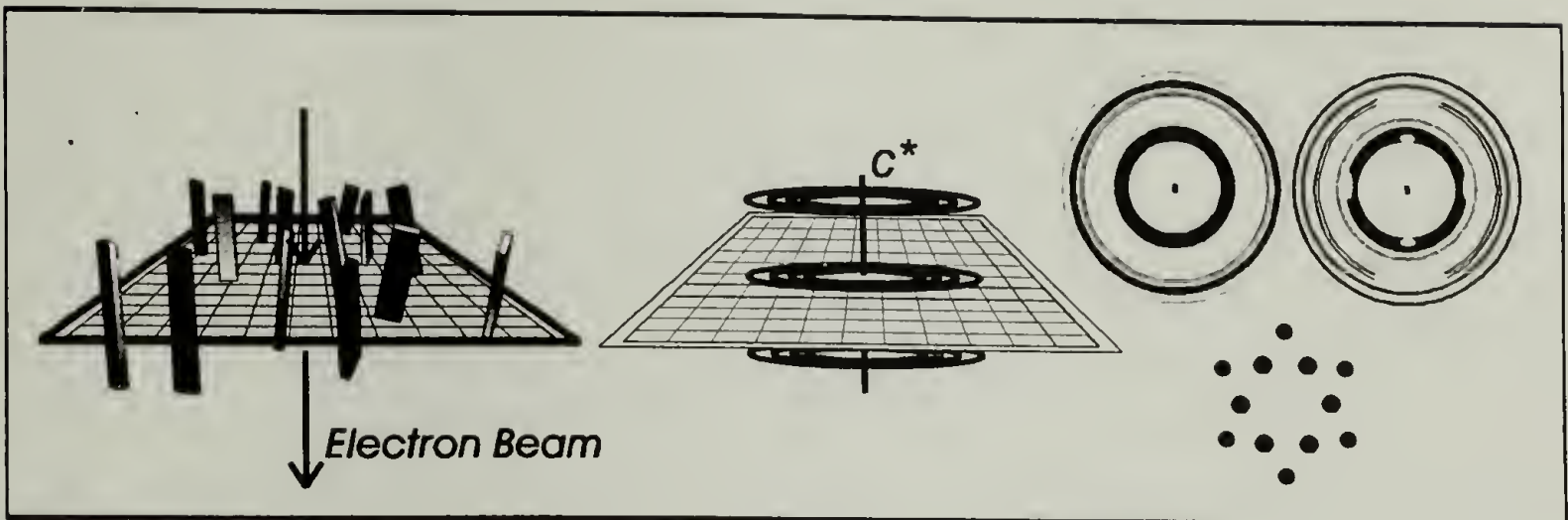
Figure 28: Schematic of LB film orientation at 16.7mN/meter

Electron Beam
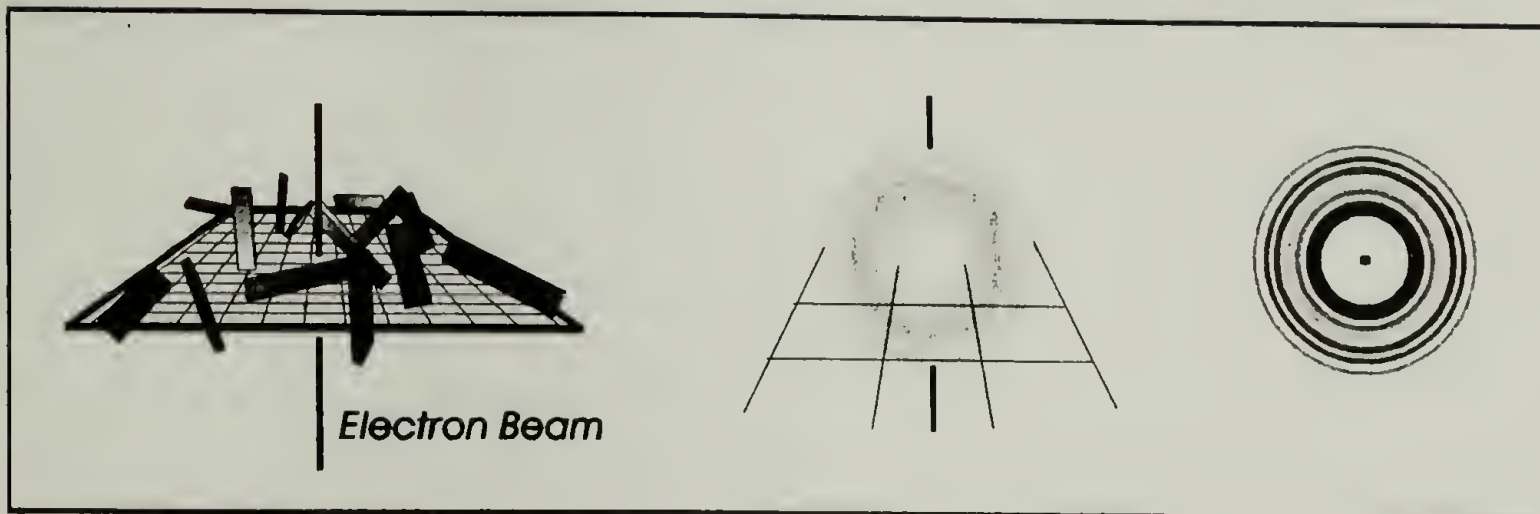
Figure 29: Schematic of unoriented cast film

Because the formation and behavior of the silk films at the air-water interface is similar in many respects to the behavior of other, more complex, proteins that change their structure (denature) at the interface and experience conformational changes as they adsorb, a similar mechanism may be present in silk adsorption to the air-water interface[84,87,96,106]. Nucleation of specific crystalline polymorphs and oriented textures by a Langmuir film has been observed previously, most notably for glycine.[132,133,139] In the studies on glycine crystallization, a Langmuir monolayer of another substance (not glycine) was used to assist the crystallization of glycine at the interface and the packing and electrostatic interactions of the monolayer resulted in a particular crystalline orientation in the glycine crystallites that formed. For LB Silk III, the layer nucleating the silk crystallites is a surface excess layer of silk; the same protein as the resulting crystals.

The surface excess layer in the LB trough has been shown to stabilize after approximately an hour, by which time surface tension measurements plateau.[140] Given sufficient time, a solid skin will form at the air-water interface of aqueous solutions left undisturbed at room temperature. The formation of a solid skin has also been observed by Happey and coworkers,[138] who found that the film appeared fibrous in the TEM. They did not report any electron diffraction data for these films. We deposited some of the same type of uncompressed films onto TEM grids (Sample 1), and observed a fibrous film in TEM, similar to our LB and cast films (Samples 2, 3 and 5) and to the structure reported by Happey. [103,104,138] Electron diffraction data taken from these films looks very much like the patterns recorded for the LB silk films as well There is a major reflection corresponding to a d-spacing of 4.5 Å from the 200 and 020 planes, and the

minor reflections also occur in the correct locations for the silk III structure. All of the reflections expected for silk III are present, as in the unoriented cast films (Sample 5), but the diffracted rings are sharper than in the unoriented patterns. In addition, the 200 and 400 (indexing using the four chain monoclinic unit cell described in an earlier paper) are attenuated, whereas the 003 and 102 reflections appear stronger than in the unoriented cast (Sample 5) films. This effect is quite the opposite of that observed for the LB films. The differences in relative intensity between the three film preparations occur in a systematic fashion, and can be attributed to different oriented crystalline textures within the films.

In the uncompressed films, the crystals are oriented such that their c axes are in the plane of the film. This leaves them free to rotate about two axes as shown in Figure 31. Rotation about the first axis results in a disc of concentric rings for each layer, as shown at the top of the Figure. The $00l$ reflections remain points on the c* axis after rotating about the first axis. Rotation about the second axis causes the $00l$ reflections to spread out into a disc of concentric rings perpendicular to the two axes and parallel to (coincident with) the Ewald plane. The $hk0$ reflections are each spread out over a spherical shell as I an unoriented crystalline texture, but in this case the intensity is not evenly distributed and is instead concentrated at the poles of the sphere, far from the Ewald plane. The result is an attenuation of these reflections. The hkl reflections spread out over a shell that can be described as a truncated sphere (Figure 30b). Intensity is again concentrated at the edges of the surface, away from the Ewald plane, but for hkl reflections that are close to the $00l$ in reciprocal space the resulting attenuation of intensity in the Ewald plane is less severe. Since these reflections make a smaller radius with the first rotation axis, the first rotation results in a ring with a small radius. When this ring is rotated about the second rotation axis, the result is a narrow

band centered on the Ewald plane. Even though the diffracted intensity is stronger at the edges of this band, it is nevertheless distributed over a smaller region of reciprocal space than it would be in the unoriented case. The result is that hkl reflections will be less attenuated than $\bar{h}kl$ reflections, and can even have enhanced intensity in the biaxial orientation if they are sufficiently close to the $00l$ reflections on the reciprocal lattice. Since the 200 reflection remains the strongest reflection in all of the orientations observed, scaling of the hkl reflection intensities against the 200 to attain relative intensities results in a higher relative intensity for all non-$hk0$ reflections, because they are all scaled against a reflection that experiences the maximum possible attenuation due to orientation.

In the uncompressed films, if the crystallites formed from silk helices with their axes in the plane of the film, rings would still be expected observed in electron diffraction patterns. Within the plane of the film, the direction of the helical axis (c axis) is unconstrained. Diffraction rings are in fact observed for the uncompressed film samples. The change in orientation in the uncompressed films (Sample 1) would account for the presence of reflections that are weak or absent in the LB films. Some of these reflections should also be enhanced when compared to the powder patterns, as is evident by comparing the relative intensities of the silk III reflections for the unoriented drop cast films, the LB films and the uncompressed films in Table 5.

As can be seen from the comparison of relative intensities in Table 5, the 200 and $2\bar{2}0$ reflections retain the same intensities relative to each other in all three film preparations compared, consistent with the orientations proposed. Intensification of hkl reflections is more pronounced for hkl planes whose normal, the $[hkl]$ axis, makes a higher angle to the $[100]$ axis. The corresponding reciprocal lattice points are near a

001 reciprocal node, and the [*hkl*] axes of strongly intensified reflections are also at a lower angle to the [001] axis. This trend can be seen by comparing the data in Table 5 for the 111 and 112 reflections. The [111] axis is at an angle of 26° to the [100], whereas the [112] is at an angle of 68° to the [100]. The 111 reflection in the uncompressed film has a relative intensity roughly 2.5 times its value in the unoriented cast films, whereas the relative intensity of the 112 reflection increases by a factor of approximately 4. In diffraction from the LB films compressed to 16.7 mN/meter, the attenuation in the relative intensity of *hkl* reflections can be related to the angle that the [*hkl*] axis, normal to the (*hkl*) plane, makes with the [001] axis. Recall that the compressed LB films are uniaxially oriented about the [001] axis. The [202] and [2$\bar{2}$2] axes are both at an angle of 28° to the [001] axis, and the 202 and 2$\bar{2}$2 reflections are both attenuated by a factor of approximately 4 in diffraction from the 16.7 mN/meter LB films. The trend in relative intensities is not always clear due to measurement difficulties in separating the contributions of component overlapping reflections during peakfitting.
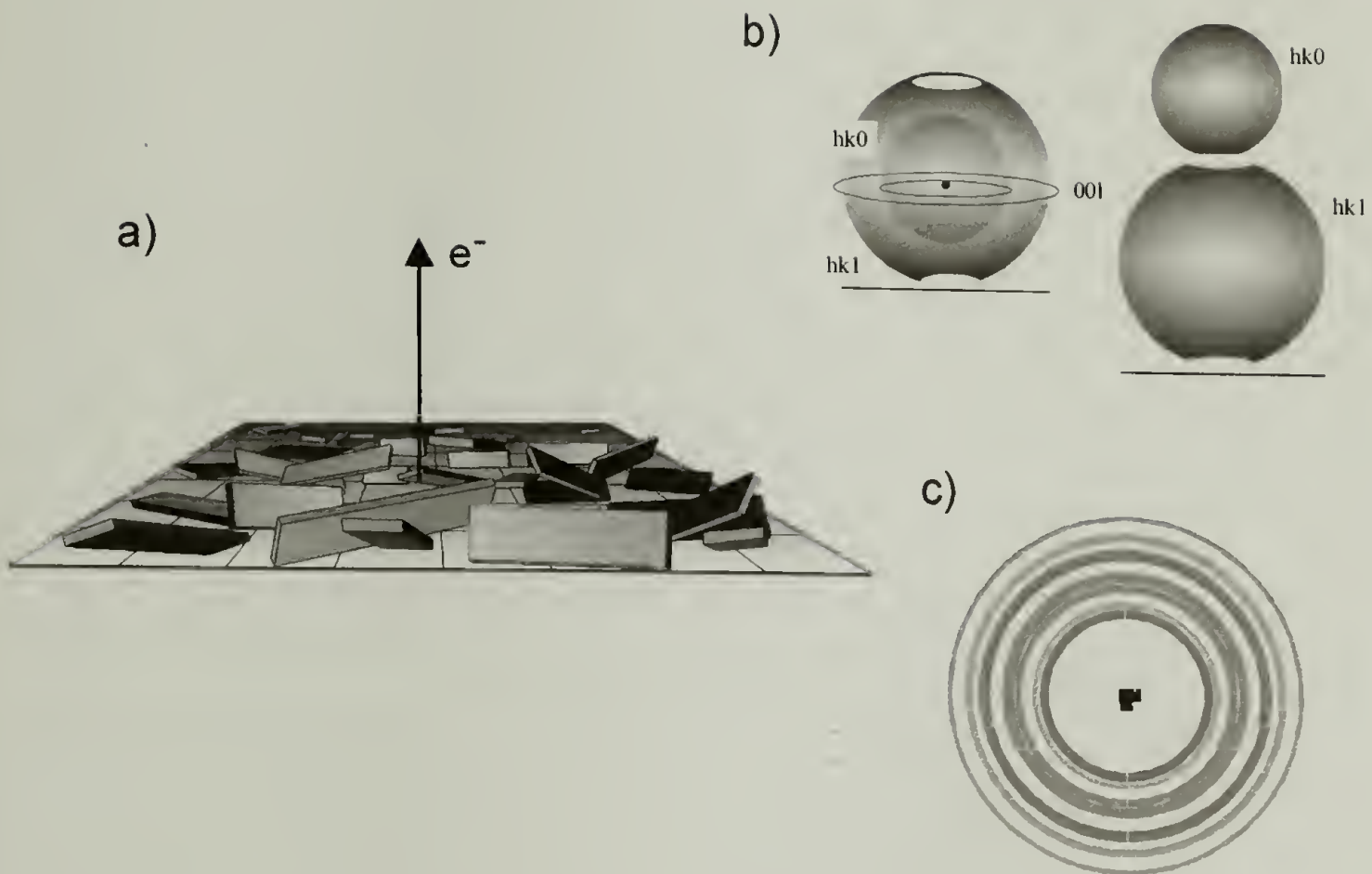
Figure 30: Crystallite orientation at the air-water interface, uncompressed surface (Sample 1). (a) The c axis is always in the plane of the film resulting in more intense *00l* reflections, because nearly all of the (*00l*) planes are at the Bragg condition. In contrast only a much smaller fraction of the *hk0* planes are at the Bragg condition (b) Reciprocal space schematic of the orientation in *(a)* (c) The *hk0* reflections are attenuated in the resulting polycrystalline diffraction pattern.

In addition to the diffraction data, the morphology of the films supports an in-plane orientation for the helical axes. Figure 26a is a TEM image of a crack in one of the films. As can be seen from the , there are fibrils spanning the crack, a behavior one would expect if the helical silk molecules had their axes in the plane of the film. Similar morphologies have been observed in chain folded lamellae, where the chains are perpendicular to the sample plane and unfold to span cracks in specific crystallographic directions. The fibrils observed in morphology images are too large to represent single chains and result from aggregation of locally aligned chains. A similar mechanism for the spanning fibrils cannot be invoked here because the diffraction evidence indicates a polycrystalline sample with very small (roughly 100 - 200 Å) crystallites and not a lamellar single crystal. Furthermore, as can be seen from Figure 26, the fibrils protruding from the edges of the film do not prefer any particular direction and occur on all of the film edges. The most reasonable explanation is that the fibrils result from localized clustering of silk chains oriented with their helical axes in the plane of the film. In fact, close examination of the morphology image of a film edge in Figure 26c reveals fibrous bands that end at the protruding fibrils, implying fibrous clustering of silk chains within the film. Similar morphological features, including dark bands with the correct diameter to be fibrils (approximately 20 nm), are observed throughout the films. Contrast this with the morphology image from a uniaxially oriented LB film where the chain axes are perpendicular to the plane of the film, shown in Figure 27. There are numerous small holes and tears in the film depicted in the Figure, but in this case the edges of the film are clean.
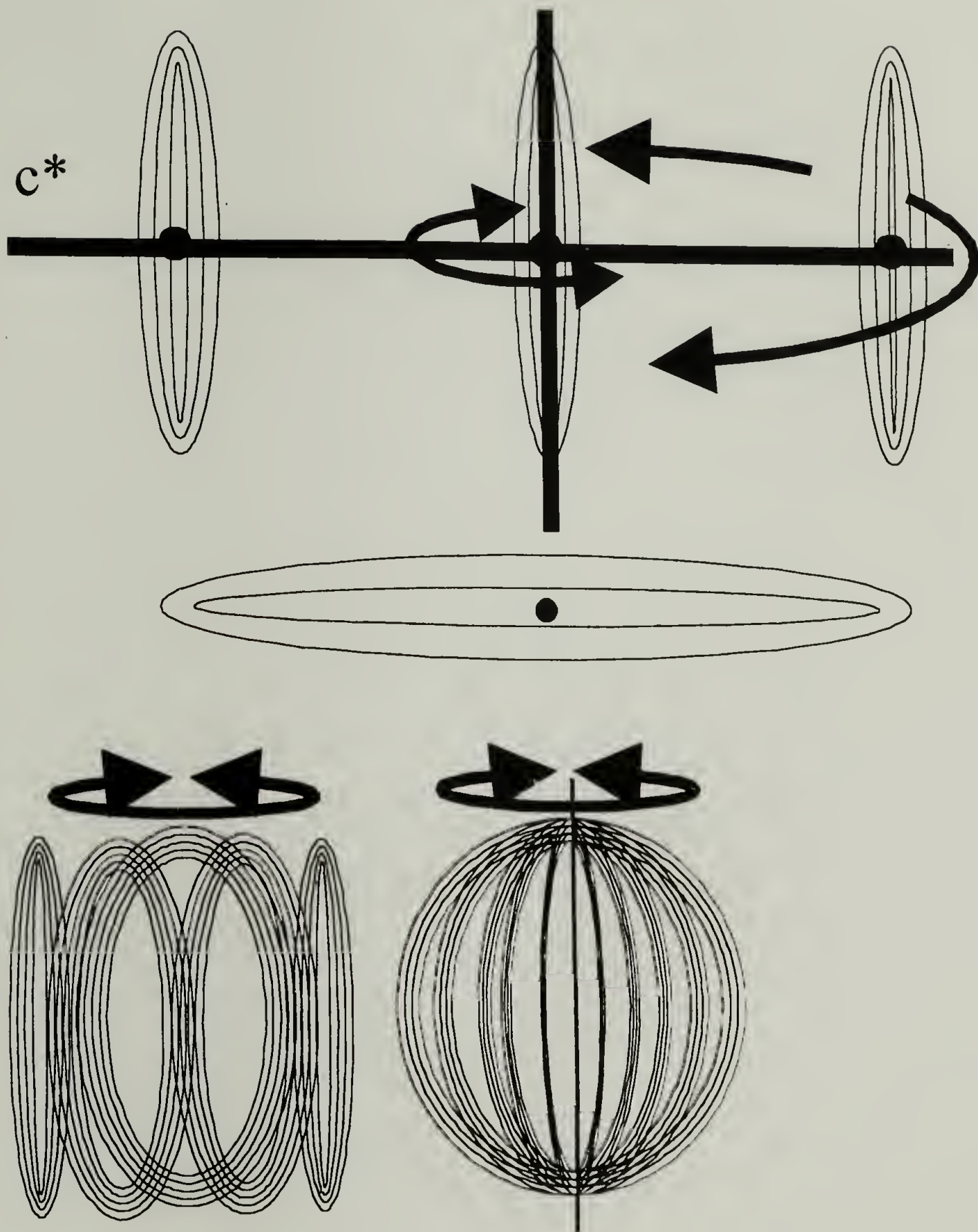
Figure 31: Two rotation axes are present in the uncompressed films. (*Top*) Uniaxial orientation of crystallites about the c axis results in rings symmetric about the c* axis. The *00l* reflections are points on the c* axis. (*Middle*) When the second axis is added, the *00l* reflections become a series of rings in the plane perpendicular to the two axes. (*Bottom*) The *hkl* reflections are distributed over a series of truncated spherical shells while the *hk0* reflections are distributed over a series of spherical shells. In both cases the intensity is concentrated away from the Ewald plane.

The LB films at 16.7 mN/meter have a consistent crystalline orientation that is different from what one would intuitively expect.[103,104] If the air-water interface causes the silk molecules to act as amphiphiles resulting in a threefold helical conformation, the silk surface layer would consist of silk helices lying in the surface, with their chain axes parallel to the surface. The crystallites observed in the LB films are oriented with their chain (or *c*) axes *perpendicular* to the surface, as depicted in Figure 28. This result appears inconsistent with the polysurfactant model until the data from the uncompressed films and from films aged at multiple compression states are considered. As can be seen from the schematic of the crystallite orientation and the reciprocal space diagram in Figure 28, the orientation of the helices with their chain axes perpendicular to the surface (Sample 2) results in attenuation of the *hkl* reflections and absences of *00l* reflections, as compared to the unoriented cast films (Sample 5) diagrammed in Figure 29. In the uniaxially oriented 16.7 mN/meter LB films the relative intensities of the *hk0* reflections are strengthened. This effect is apparent in the relative intensities listed in Table 5. The 200 reflection is stronger in the uniaxial LB patterns (Sample 2) than in the unoriented patterns (Sample 5), while the 202 is weaker. The 006 is absent in the uniaxially oriented LB film samples. The fact that the diffracted intensities observed in the LB films are an orientation effect is demonstrated by using the goniometer to tilt the films. Diffraction rings are observed at $0°$ tilt whereas arcs are observed when the film is tilted.[103] Analysis of the 16.7 mN/meter LB films has been treated in greater detail in the previous chapters.
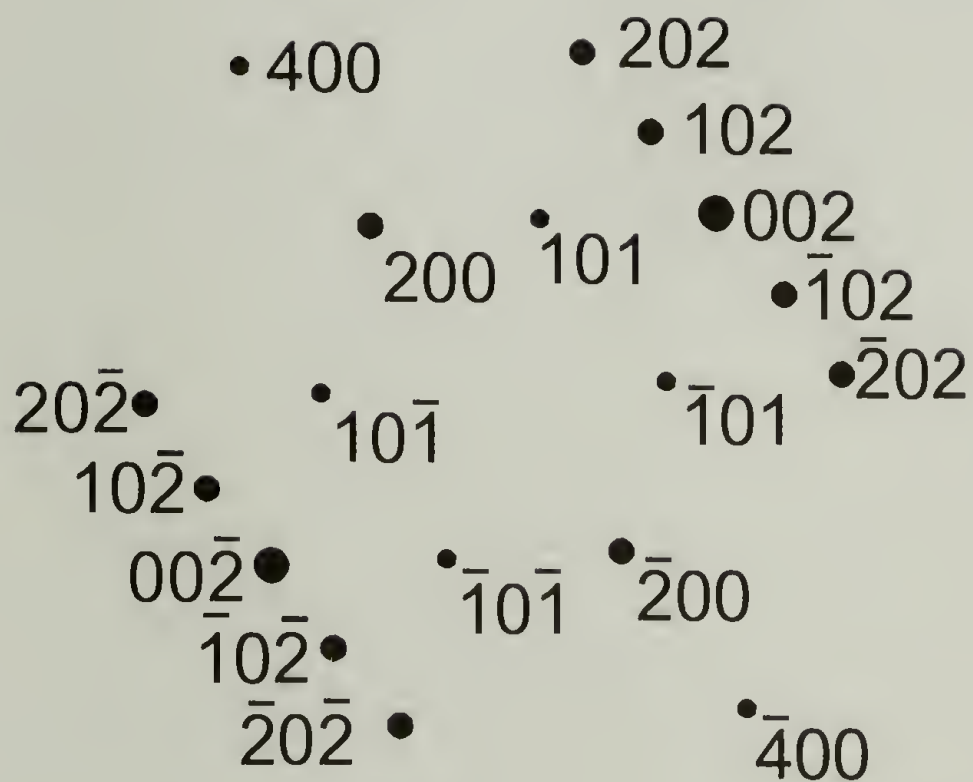
Figure 32: Silk II single crystal diffraction pattern from a sample film taken from a surface compressed to 34 mN/meter. Diffraction is from the [010] zone, looking down the chain axis.
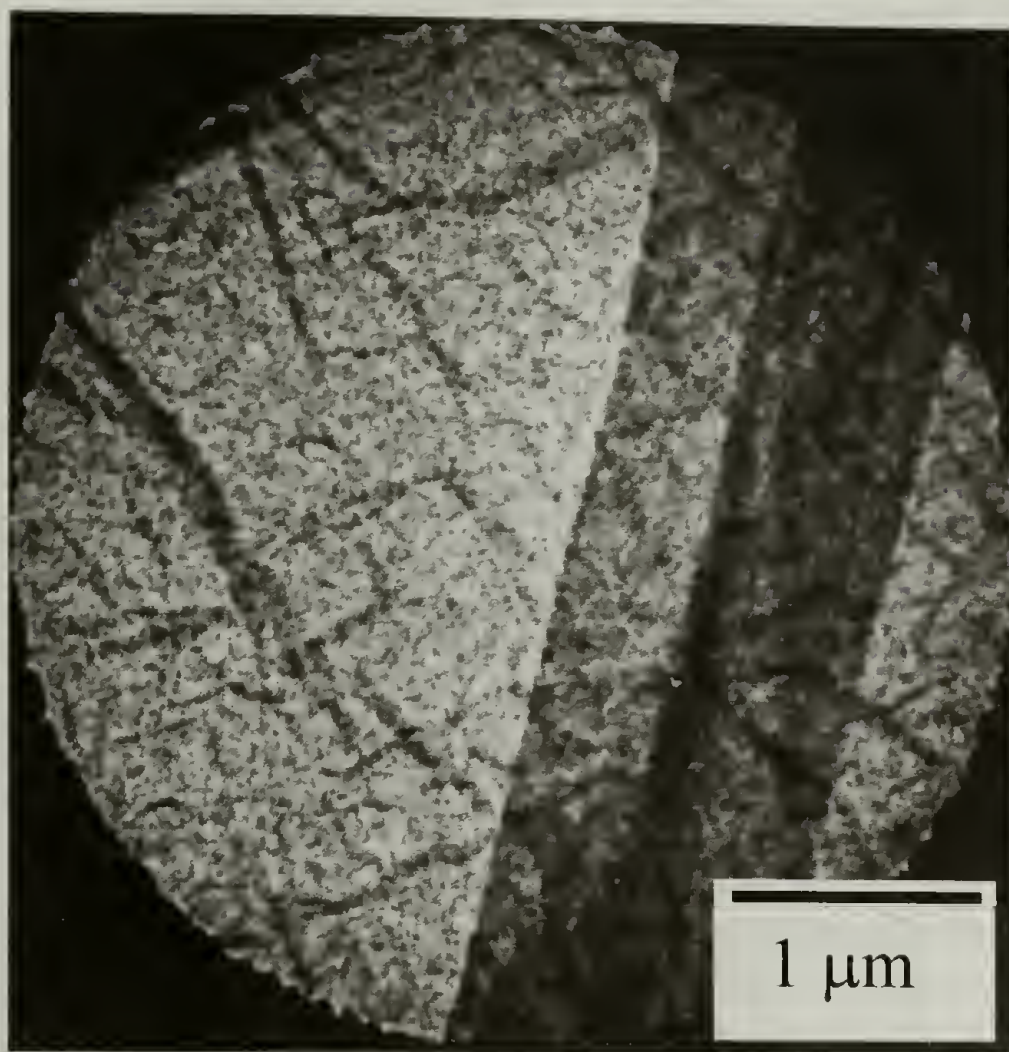
Figure 33: Morphology from an LB film that was aged prior to compression.
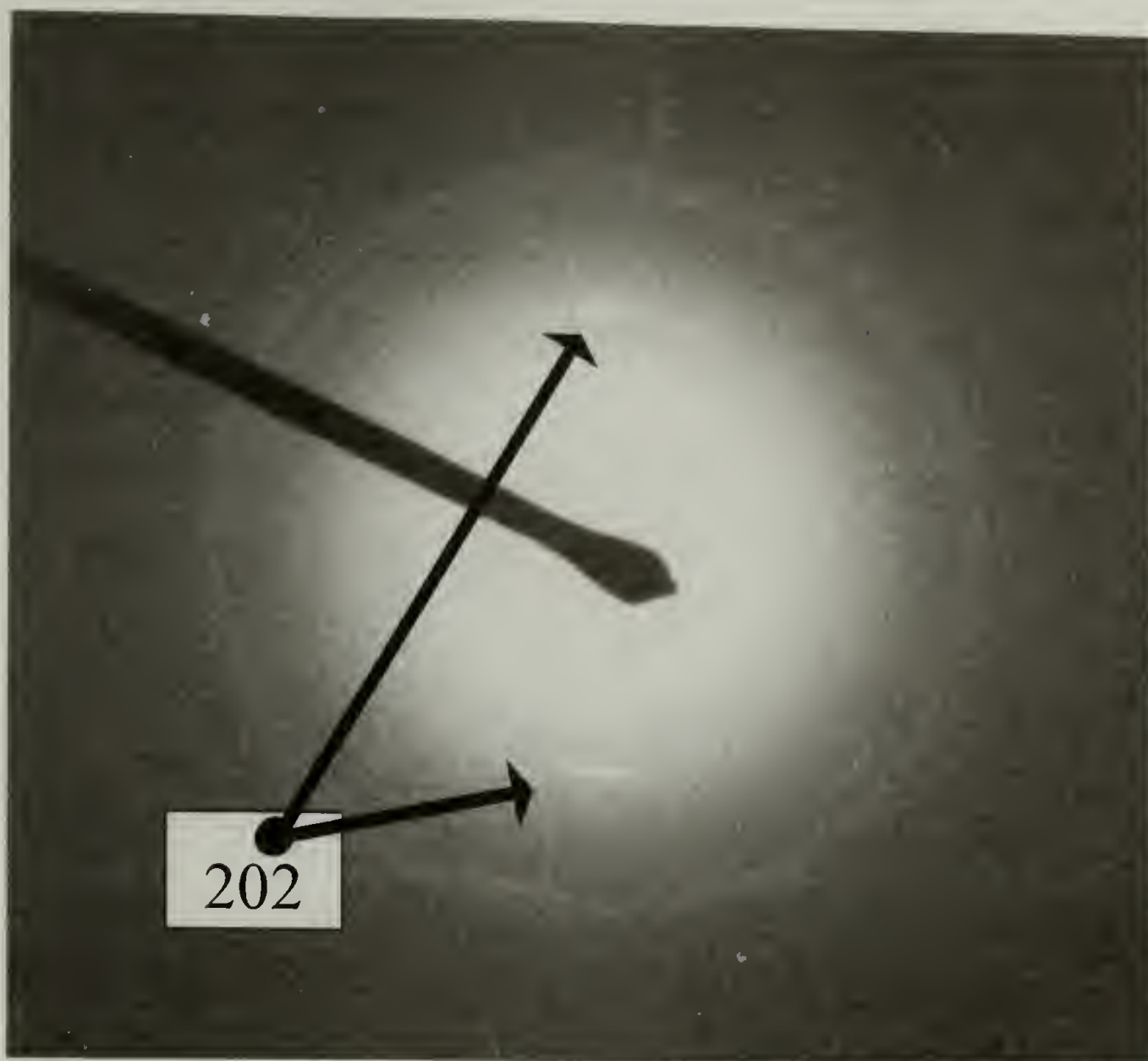
Figure 34: Diffraction from an LB film aged prior to compression

When the LB films are compressed to a higher pressure, 34 mN/meter instead of 16.7 mN/meter, the lamellar single crystal patterns observed possess the silk II structure, but again the orientation places the chain axis perpendicular to the plane of the interfacial film.[1,40,50,51] Figure 32 is a single crystal diffraction pattern obtained from a 34mN/meter LB film (Sample 3). The spacings and symmetry match the [010] zone of silk II and are indexed accordingly. Recall that in the published crystal structure for silk II the chain or helical axis is given as the b axis.

A combination of orientations can be produced by combining the compression treatments used to prepare the films. In sample preparation method 4, the surface layer was aged prior to compression in order to form a solid surface film of adsorbed fibroin. The surface was then compressed to 16.7 mN/meter and again allowed to age. TEM morphology images reveal fibrous films with a crumpled appearance such as the one in Figure **33**. In films aged on the trough without pressure and then aged at 16.7 mN/meter (Sample 5), two different orientations are observed. The uniaxial orientation described for the 16.7 mN/meter (c axis perpendicular to the plane of the sample film) coexists with oriented regions that have the 202 axis in or near the plane of the film. There are two orientation effects observed in these films using diffraction evidence. Many areas of the film yield diffraction patterns with a very sharp arc on the 202 reflection ring, such as the pattern shown in Figure 34. The diffraction pattern in the Figure suggests an orientation that places the [101] axis near the plane of the sample

film. In addition, six fold diffracted spot patterns and diffracted rings that both match the silk III structure are often observed. The six fold spots and diffraction rings indicate regions with in a uniaxial orientation about the c axis, which in this case is perpendicular to the plane of the sample. As discussed earlier, the spot patterns are obtained from large flat lamellar crystals of silk III. The lamellae have their c axes perpendicular to the film surface.

These films appear to posses two different dominant orientations, with the orientation being different in different regions of the film. The presence of six-fold spot patterns from large lamellar crystals of silk III in these films and the ordinary LB films at 16.7 mN/meter suggests that the applied surface pressure is required to grow large lamellar crystals of silk III. These large lamellae are not observed in either the cast films (Sample 5) or uncompressed surface films (Sample 1). The mixed orientation observed in the films from sample 4, which experienced a combination of aging under no applied surface pressure (similar to Sample 1) and 16.7mN/meter surface pressure (similar to Sample 2), is evidence for an orientation transition resulting from applied surface pressure. In these films with combined compression treatments there are two orientations that coexist. The first is identical to the final uniaxial orientation favored by the 16.7 mN/meter surface pressure. The second is an intermediate orientation which is closer to that observed for uncompressed films, where the chain axis is in the plane of the film.

118
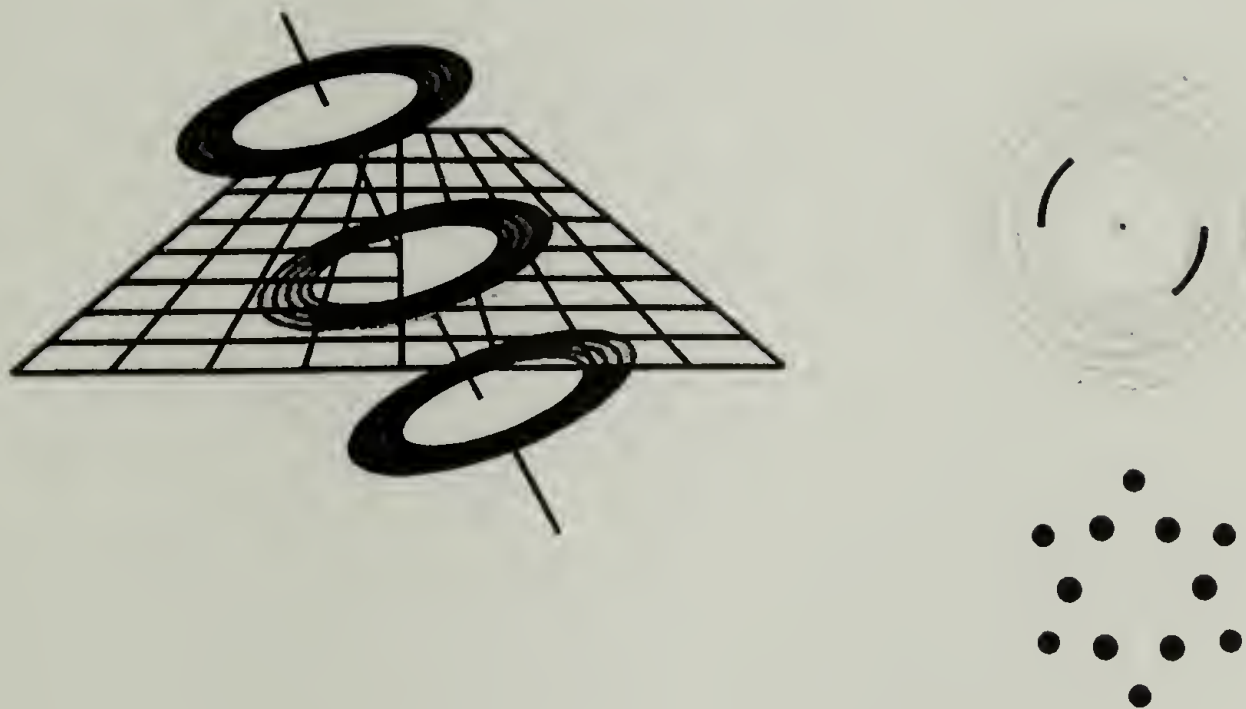
Figure 35: A combination of surface pressure treatments can be used to produce a combination of orientations.

In light of the above discussion, an explanation of the orientation behavior in the silk LB films is possible which is consistent with our observations of ultrathin silk films. The silk excess layer that forms at the air water interface results in aggregates, or micelles, of the threefold helical structure with an asymmetric shape. These aggregates possess the orientation one would intuitively expect, with the chain axis lying in the plane of the interface. When the trough is compressed these aggregates realign so that they each occupy a minimum of space at the interface. Crystallization of the silk III structure could then continue at the new orientation.

The presence of silk II at higher pressures is consistent with our explanation, since fibroin molecule helical (crystallizable) segments, once oriented with their axes perpendicular to the plane of the film, would take up less space in the silk II β-sheet conformation than in the silk III threefold helical conformation. The silk II crystallites observed at higher pressures have an orientation that is consistent and is commensurate with this mechanism, oriented with their chain axes perpendicular to the plane of the film.

4.5 Conclusions

It appears that the 3/2 helical conformation is induced by the air-water interface during film formation, probably due to the silk protein acting as a macromolecular surfactant. The orientation of the helices in the interfacial film can be altered by applying

a relatively gently surface pressure of 16.7 mN/meter after the protein film has adsorbed to the interface. A higher applied surface pressure of 34 mN/meter causes the molecules to orient and transform to silk II. Combining surface pressure treatments by aging the interfacial films on the trough at different pressure conditions in sequence freezes in the chain conformation under the initial treatment condition, probably through crystallization. The larger, more solid structures formed at the interface by aging the films are not reoriented by changes in surface pressure to the same degree as the freshly adsorbed films. The result is a film with two orientations, each of which predominates in different small regions of the film.

CHAPTER 5

A HYDRATED STRUCTURE OF SILK III OCCURS AT AQUEOUS-ORGANIC INTERFACES

5.1 Summary

Truncated hexagonal lamellar crystallites of *Bombyx mori* silk fibroin have been observed with a structure that is a hydrate of the silk III structure reported previously.[103,104] These crystallites form at the interface between hexane and aqueous silk fibroin solution. At the water - hexane interface, and similar water - organic interfaces such as water - chloroform, a surface excess layer of fibroin forms resulting in films which were studied by transmission electron microscopy. Diffraction data obtained for the crystallites match hydrated models based on the layered monoclinic modification of the silk III crystal structure.[104] A banded morphology has also been observed for *B. mori* silk fibroin films obtained from an aqueous-hexane interface; the period of the banding is approximately 1 μm. The diffraction pattern from the banded structure observed at this interface is similar to the pattern expected for a cholesteric liquid crystal. In some places the material in the banded region has crystallized locally and the films yield a diffraction pattern containing additional off-meridion reflections that would not be expected for a cholesteric, but the director orientation appears to be preserved. These patterns thus appear similar to an X-ray fiber pattern and can be indexed using the same unit cell as the hexagonal lamellar crystallites. Surfactancy of fibroin and subsequent aggregation and

mesophase formation may help to explain the liquid crystallinity reported for silk, and long suspected to play a role in the biological silk spinning process. [10,11,39,47,104]

5.2 Introduction

As part of our ongoing studies of the structures formed by *B. mori* silk at interfaces, we have investigated the structures formed at the interface between an aqueous silk fibroin solution and another liquid, in this case hexane. Previous studies of silk at the air-water interface resulted in the discovery of a crystal structure involving approximately hexagonal packing of threefold fibroin helices. [103,104] The choice of a threefold conformation at the interface is believed to be driven by surfactancy of the fibroin crystallizable repeat sequence. This sequence is approximately (*Gly-Ala-Gly-Ala-Gly-Ser*)$_x$, and a left-handed threefold helical conformation, which is sterically reasonable, [103,104] separates hydrophobic alanine and hydrophilic serine residues to opposite sides of the interface. The air-water interface is very thin, on the order of 2-3 Å, with a molecularly sharp concentration profile making it essentially two dimensional with respect to the width of a fibroin helix. [96] The presence of amphiphilic fibroin molecules at this interface does not appreciably diminish its sharpness[84,85,87,90,92,93,96,106] In studies of fibroin at the air-water interface, the conformation of the molecules at the interface is transmitted to additional molecular layers which crystallize below the interface by a structural templating mechanism. [103,104,132,133]

cluster

400 nm
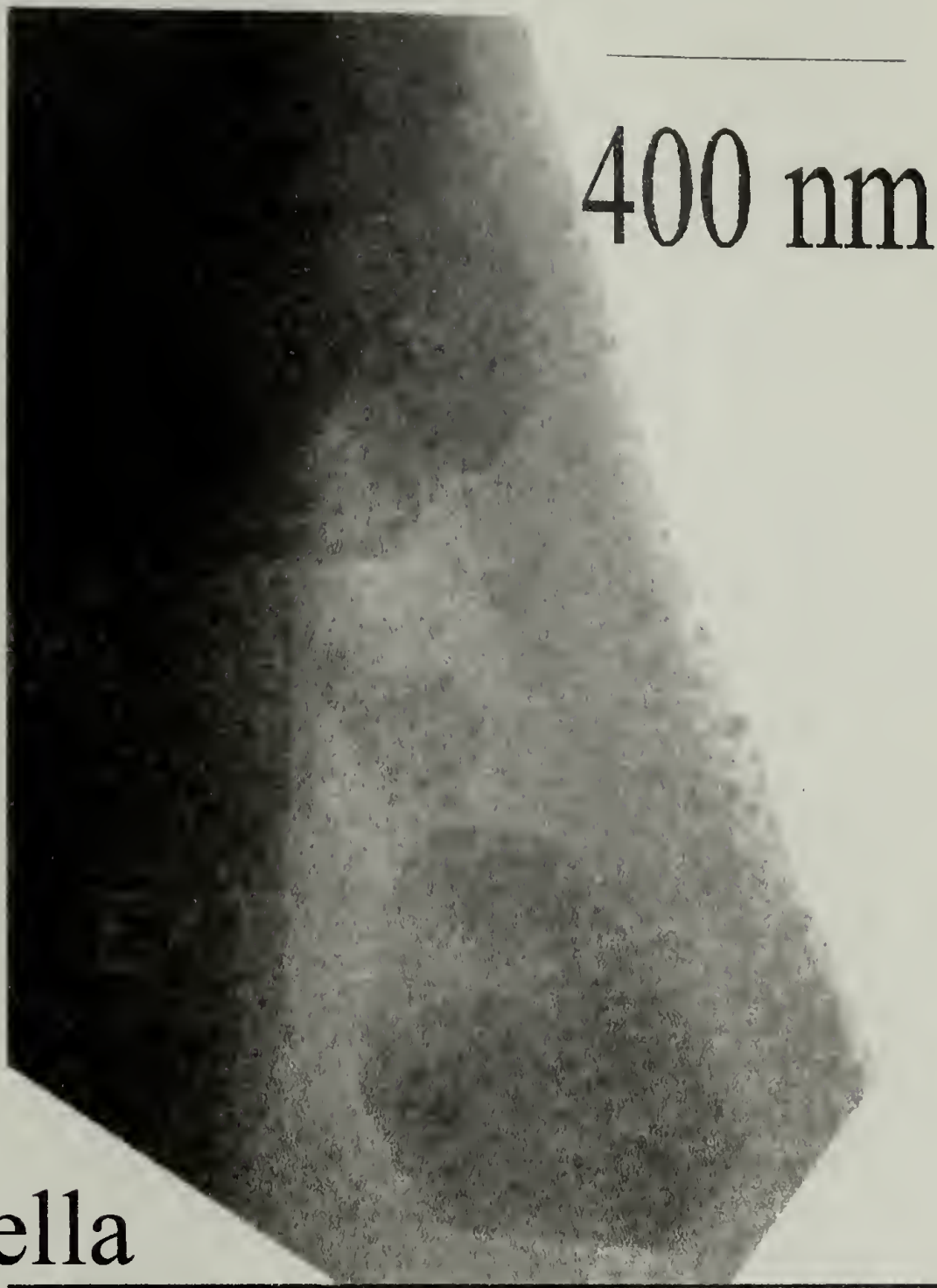
lamella

Figure 36: Crystallites from the water -hexane interface. (a) A TEM image of crystallites from the water -hexane interface interfacial film shows that they possess a truncated hexagonal habit. (b) Diffraction pattern from the crystallites in the image. The strong primary pattern is due to the dark mass of tilted crystallites labeled *cluster* in the morphology image. The faint overlaid spots (circled) are due to the single lamella in the morphology image. (c) Diffraction pattern schematic. The gray "x" and "o" markers indicate the locations of the inner spots from the faint single crystal pattern due to the single lamellar crystallite. A pair of 022 reflections are marked "o", and the 111 family reflections are marked "x".

Continued next page

124

b)

c)

205

135

123 113

X

102

X

121 o

240| 120 | ▪▪100   ▪▪ | |

o

X

020, 110   X
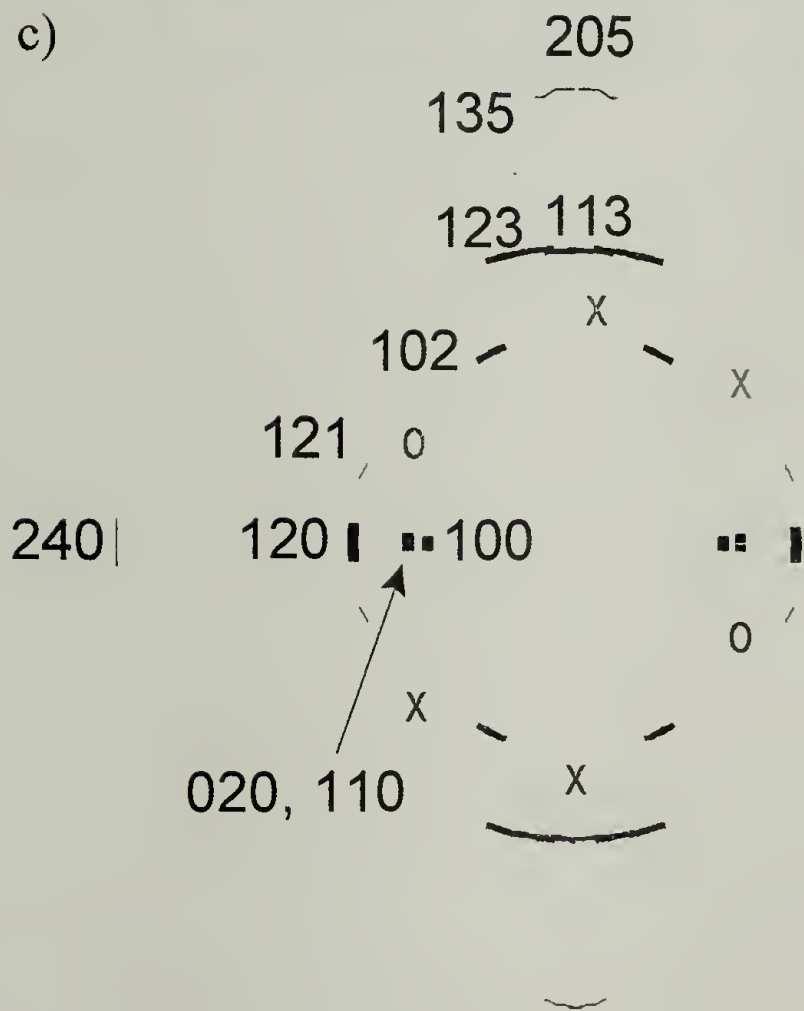
Figure 36: *continued*

Liquid-liquid interfaces can in many cases be much thicker than the air-water interface[89,105,141-143] incorporating a significant chemical gradient on a length scale of a few or even a few tens of Ångstroms, comparable to the width of several protein helices. The presence of silk fibroin acting as a surfactant is expected to result in a thicker interfacial region as well.[105] Thus an oil-water interface could provide a thin three-dimensional interphase region where the conformation of the dissolved fibroin can develop and form three-dimensional structure under the direct influence of the chemical gradient between the two liquids. An aqueous-hexane interface was chosen as an initial probe of fibroin liquid-liquid interface behavior. This interface, in the absence of fibroin, is believed to be about 10 Å thick.[142,143] The silk at the aqueous-hexane interface forms a film as it ages, and this film can be picked up onto sample grids for observation in a transmission electron microscope (TEM). The hexane was expected to be a better solvent for the alanine residues in silk than the water, forcing them to the hexane side of the interface. The aqueous phase should be a better solvent for serine. TEM imaging shows regions characterized by a regular banded pattern as well as truncated hexagonal crystallites in the films. The crystallites are shown in Figure 36. The banded structure is shown in Figure 37. Diffraction evidence from the crystallites and from the banded regions in the films indicates a hydrated crystal structure incorporating a conformation with ϕ and ψ angles that vary along the backbone of each chain, but incorporate two full helical turns in six residues as described in a previous paper.[104] The conformation observed thus represents a localized rearrangement of a threefold helix, which supports the hypothesis that the thermodynamic interactions among the different residues and the

two solvents on opposite sides of the interface segregate the residues and force a threefold helical conformation.

The oily interior region of a typical globular protein conformation is usually modeled by a solvent such as dioxane or ethanol, both of which are more compatible with water than hexane.[99,115,144]In these proteins there is a statistically significant tendency for hydrophilic residues to reside on the outside of the protein, while the hydrophobic residues are more abundant in the interior of the protein.[115,131,145] In the case of silk fibroin at an aqueous-hexane interface, we may be observing a similar segregation effect in a fibrous protein. Both aqueous solutions of regenerated *B.mori* silk fibroin and washed diluted fibroin from freshly dissected silk glands will readily form bubbles and foams. Some portion of the silk fibroin molecule must be capable of behaving as a surfactant. While the non-crystallizable portions of the chain may be at least partially responsible for the surfactant behavior, they comprise only about 20% of the chain.[1,6,24,25] In our interfacial experiments we obtain highly crystalline films incorporating a conformation that allows the crystallizable portion of silk to behave as a surfactant at the interface.

a)

2 μm
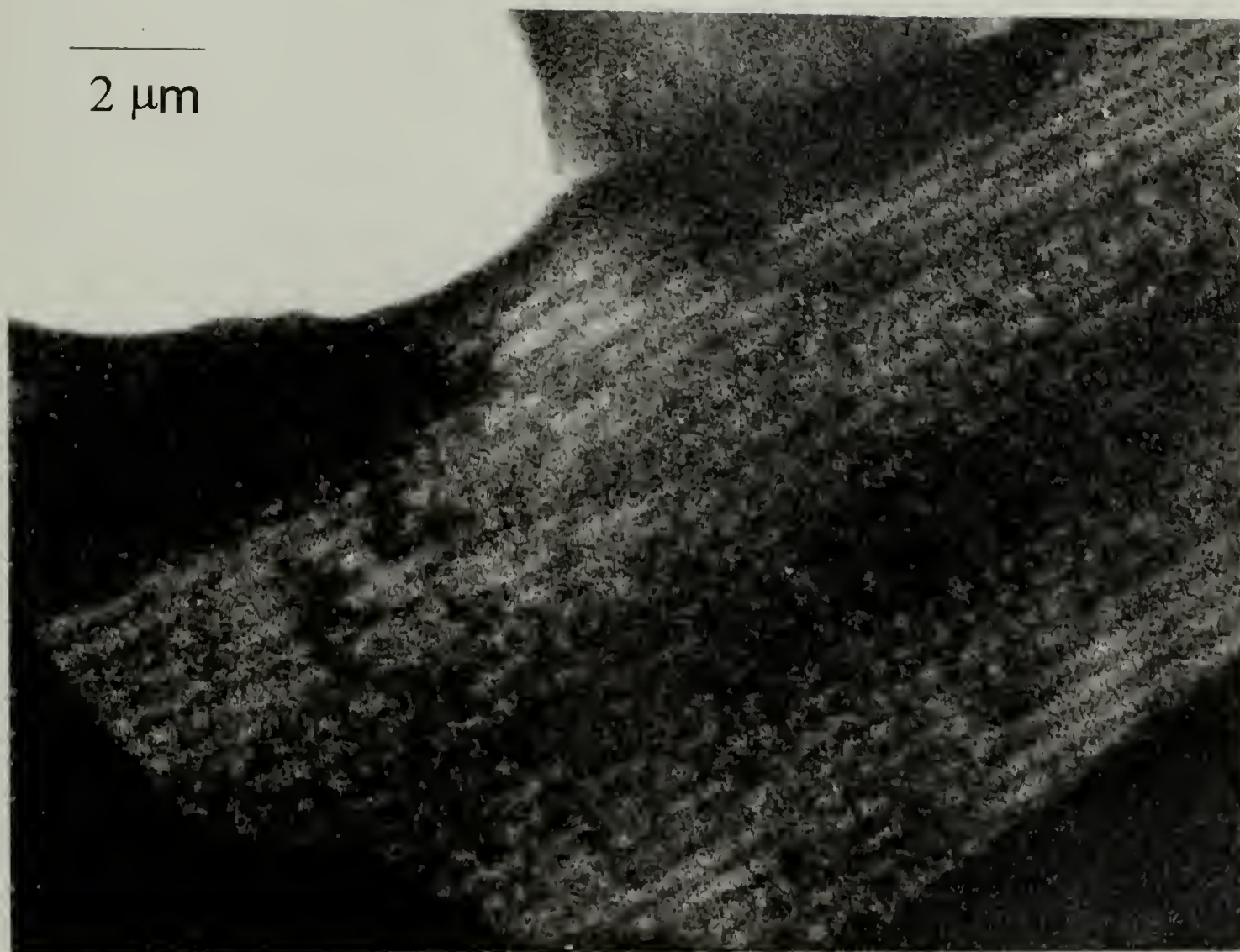
Figure 37: Banded mesophase at the water-hexane interface. (a) TEM image (b) Diffraction pattern from a non-crystalline portion of the banded region. (c) Defocused diffraction image for the diffraction pattern in *b*. (d) Diffraction pattern schematic.

Continued next page

b)

c)

d)

5.9 Å

5.5 Å
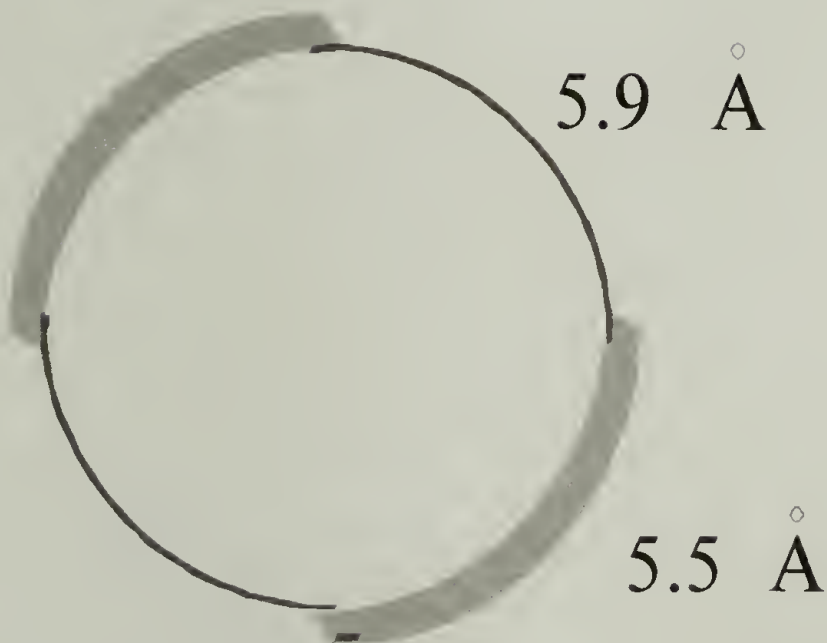
Figure 37: *continued*

5.3 Experimental

Bombyx mori silk cocoons were degummed using boiling water, sodium dodecyl sulfate (SDS) and $CaCO_3$ to remove the sericin, leaving pure fibroin. Other cocoons were degummed without SDS, using only $CaCO_3$ and boiling water. Amino acid analysis has been used to assess the protein composition of fibroin prepared in this manner and no sericin was detected. [103,104] The degummed fibroin was rinsed thoroughly with distilled water and dissolved in a 9.1 M solution of LiSCN in water. In order to remove the salt, the fibroin and LiSCN solution was then dialyzed against frequent changes of distilled water for several days. The dialyzed fibroin solutions were filtered to remove dust and any protein precipitate.

Aqueous fibroin solution was placed in a clean crystallization dish and covered with HPLC grade hexane. The dish was covered to reduce evaporation and the hexane was replenished as needed. Interface films were collected onto TEM grids (with no support substrate) after 3 days (Group A) and after 7 days (Group B). The interfacial material formed a very thin self-supporting film across the grids when dry. The films collected from the hexane interface were examined shortly after their preparation, and some were reexamined three months later in order to determine if further structural evolution took place on the grid. Samples observed after three months of aging comprise Group C.

Interfacial films of silk fibroin were then characterized using a JEOL 2000 FX-II TEM operating at 200 kV accelerating voltage. Samples were maintained at below -150°C during the TEM characterization, utilizing a cryogenic sample holder. Working at cryogenic temperatures was necessary in order to reduce beam damage and to prevent the loss of water from hydrated crystal structures in the high vacuum of the microscope. Electron diffraction, and TEM bright field imaging were used to assess the structures in the films. An internal gold standard was used to determine lattice spacings. Defocused diffraction imaging was used to determine the relative orientations of the diffraction patterns and banding or crystallite facets in the morphology images. Detection of salt contamination and the characteristics of residual salt crystallites have been described in previous papers.[103,104] No salt artifacts were observed in the structures obtained from the water-hexane interface.

The films obtained from the aqueous fibroin - hexane interface were prepared from degummed fibroin where a small amount of SDS, a commercially available surfactant, was used in the initial degumming process. However, the same crystal structure and diffraction pattern was observed in data from aqueous fibroin - chloroform interfacial films, where no SDS, only $CaCO_3$ and water were used in the degumming. The similarities in the diffraction patterns from the crystallites and the banded morphology strongly suggest that the features observed are due to the same molecule in the same conformation. Thus we can rule out a soap impurity as the cause for the unusual morphological behavior in these films.

5.4 Results and Discussion

After 1 week at the aqueous - hexane interface, the films (Group B) became too thick for TEM at 200 kV. Aging the films on the grids (Group C) resulted in film disintegration in all but the thickest regions. In the surviving portions of film, morphologies typical of the fresh films in Group A were observed for Group C. Lamellar crystallites with a truncated hexagonal habit, shown in Figure 36a, and banded structures, shown in Figure 37a, were observed in both the fresh films (Group A) and films aged on the TEM grid (Group C) where the silk film had been picked up after 3 days at the hexane - 5% aqueous fibroin interface. These morphologies, truncated hexagonal lamellar crystals and banded regions, have diffraction patterns shown in Figure 36b and Figure 37b respectively. For the banded morphology in Figure 37, the defocused diffraction pattern in Figure 37c shows the orientation of the banding with respect to the diffraction pattern in Figure 37b.

In Figure 36a, a single lamellar crystallite with a hexagonal habit can be seen in the lower portion of the Figure, and an oriented cluster of crystallites can be seen in the upper portion of the Figure. The major diffraction pattern in Figure 36b is due to this cluster, but a faint pattern of single crystal diffraction spots from the large lamella is superimposed. These spots are circled in Figure 36b. In the diffraction schematic in Figure 36c the spots due to the single lamellar crystallite are denoted by either an "x" or "o", where all of the reflections marked with "x" have the same d spacing as do the reflections marked by "o". Both diffraction patterns, single crystal and oriented cluster,

can be indexed using an orthorhombic unit cell with dimensions $a = 6.1$, $b = 10.4$, $c = 13.6$ Å. The indices for the diffraction pattern due to the cluster are also shown in the diffraction schematic in Figure 36c. Models incorporating two fibroin chains per unit cell, based on the approximately threefold helical conformation described in an earlier paper[104] with one hydration water for every two residues, provide a reasonable fit to the experimental data. In the first publication on the silk III structure, the left-handed threefold helical conformation was mistakenly described as a 3/2 helix.[103] In fact, correct nomenclature for the silk III helical conformation is a 6/2 helix, incorporating six residues to make two full helical turns. This chain conformation does, however, possess a approximate 3_2 screw axis symmetry. In the silk III crystallites observed at the air-water interface, the six residue repeat sequence results in six unique slightly different ϕ and ψ angles for each of the residues in the sequence. The resulting 6/2 helix is irregular, but despite the irregularities in the ϕ and ψ angles, the silk III helical conformation observed at the air-water interface acts as a surfactant, separating hydrophilic serine and hydrophobic alanine residues to opposite sides of the interface. At the water-hexane interface, shortening of the c axis in the silk III hydrate results in a conformation that is more irregular, but still retains the essential silk III 6/2 helical character.

a)

500 nm

Figure 38: (a) Crystallites observed in films taken from the aqueous fibroin-chloroform interface. (b) Diffraction pattern from region in *a*, taken at 0° goniometer tilt (left). Diffraction pattern from region in *a* at 22° goniometer tilt (right). (c) Schematics of the diffraction patterns in *b*.

b)

0° tilt   25° tilt

c)
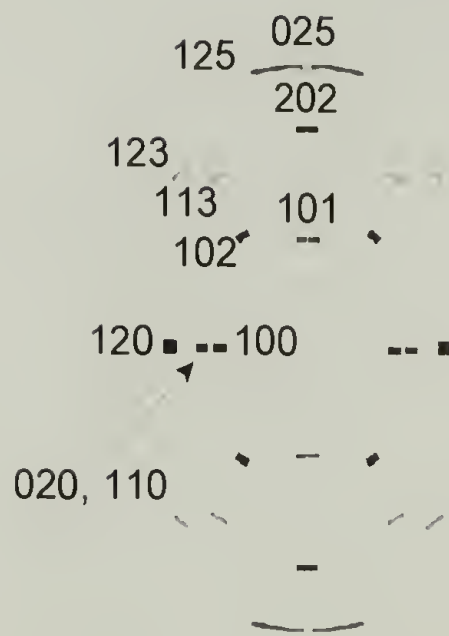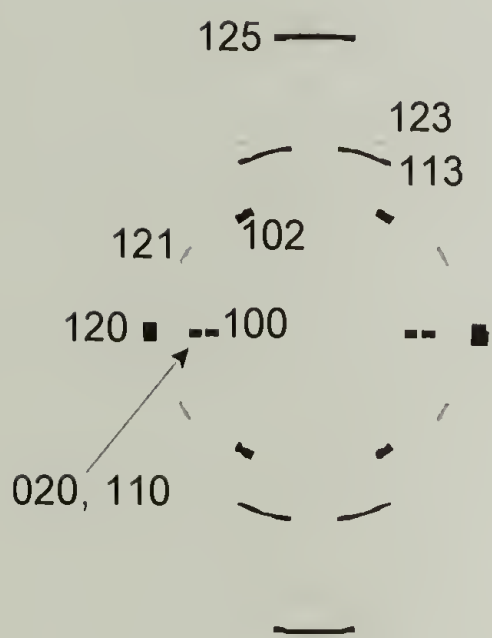
Figure 38: *continued*

In our first publication concerning silk fibroin at the air-water interface a structure was proposed based on a primitive unit cell with a perfect hexagonal packing of threefold-helical chains possessing trigonal symmetry.[103] In our second publication concerning silk fibroin at the air-water interface,[104] this structure was slightly modified to allow for preferential chain packing orientations which result in sheet-like arrangements of serine residues. The packing in this structure was only slightly distorted away from a perfect hexagonal packing of helices, with a unit cell angle γ of 116° rather than the 120° expected for perfect hexagonal packing and trigonal symmetry. The slight deviation from trigonal symmetry results in a monoclinic modification of silk III found at the air-water interface. It results from a packing that places serine residues in the ($1\bar{1}0$) planes, resulting in hydrophilic ($1\bar{1}0$) planes that could provide a location for hydration of the silk III structure. Hydration water, if present, would be expected to further distort the crystalline lattice away from hexagonal packing and trigonal symmetry, resulting in a value for γ less than 116°. The orthorhombic unit cell determined in this publication for silk fibroin at the water-hexane interface is equivalent to a distortion of the previous two silk III structures with a γ angle of 113°.

Hexagonal lamellar crystallites, identical to those found at the aqueous-hexane interface (Figure 36), have been obtained in studies of the aqueous fibroin - chloroform interface and are shown in Figure 38a. Diffraction patterns, shown in Figure 38b and schematically diagramed in Figure 38c, were obtained for a cluster of lamellae from the

aqueous chloroform interface at two different sample tilts, 25° apart. These diffraction patterns were obtained with a sample orientation in which the [001] direction is tilted off the meridian. When the sample is tilted 25°, additional reflections appear along this meridional direction which can be indexed as the 101 and 202, and correspond to d-spacings of 5.3 and 2.6 Å. The model predicts stronger intensity and slightly higher d-spacings for these reflections, but such discrepancies are explained if peak broadening due to small crystallite size is taken into account. If the centers of the 101 and 202 are slightly out of the Ewald plane , but they are strong enough to nevertheless contribute intensity to the diffraction pattern, they will not intersect the Ewald plane at exactly the expected location.

The crystallites found in clusters at the hexane and chloroform interfaces occur at a small range of tilts with the c axis (chain axis) less than 40° from the plane of the film. In the morphology images, crystallites possessing a hexagonal habit also occur, and in many cases appear to be lying flat, or nearly flat in the plane of the film with their c axes (normal to the lamella) perpendicular to the sample plane. As mentioned earlier, one of these crystallites contributes a pattern of faint spots overlaying the main diffraction pattern, in Figure 36. The dark cluster in the thick region of the film in Figure 36a appears to be a group of crystallites oriented with the crystallographic c axis at an angle of approximately 15° to the plane of the film, giving rise to the main diffraction pattern in Figure 36b. The additional faint single crystal spot pattern, overlaying the oriented polycrystalline pattern, consists of diffraction spots aligned with the facets of the

hexagonal crystallite at the bottom of the image. These reflections consist of 022 ("x")

and 111 reflections ("o"), indicating that the diffraction pattern comes from a crystallite

tilted such that the [0$\bar{1}$1] zone axis is parallel or nearly parallel to the incident electron

beam. In the proposed unit cell, the (0$\bar{1}$1) planes make a 38° angle with the $\{hk0\}$

planes, or equivalently the [0$\bar{1}$1] zone axis is 38° from the $<hk0>$ axes. Assuming that

the crystallite has threefold rotational symmetry (a hexagon with alternating long and

short facets), an approximate measure of the crystallite tilt can be obtained.

Measurement of the crystallite facets indicate that it is tilted 35° out of the plane of the

film. Since other crystallites, such as the ones shown in Figure 39, have been observed

with the assumed threefold rotational symmetry, the tilt calculated for the crystal in

Figure 36a is reasonable. The orientation of the diffracted zone, using the proposed unit

cell, is thus consistent with the tilting of the crystallite observed in the morphology

image; both indicate a orientation approximately 35° out of the plane of the sample film.

Even though the orientation for the single lamellar crystallite is significantly different

from that observed in the clusters of crystallites (Figure 36 and Figure 38), the observed

crystalline habit, d-spacings of reflections, and symmetry are consistent with a silk III

hydrate model.
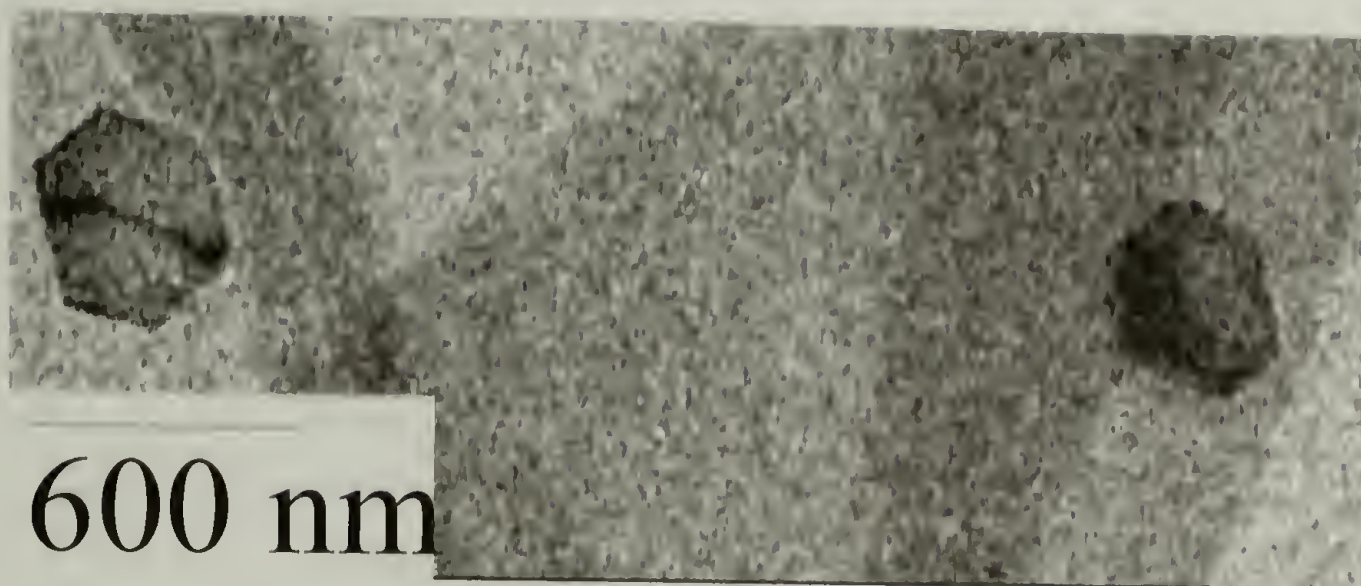
Figure 39: TEM morphology image of hexagonal crystallites of silk fibroin at the water-hexane interface.

Diffraction from the hexagonal lamellae indicates a pronounced deviation from hexagonal packing and symmetry due to hydration of the silk III structure along one of the interhelical nearest neighbor directions, as can be seen from the model in Figure 40. Figure 40a is a view of the model structure looking down the a axis and Figure 40b is a view down the c axis (the chain axis). The structure shown is a "super cell" containing a number of slightly different chain arrangements which are superimposed to generate a single average model unit cell. This is necessary in order to account for the disordered aspects of the silk III hydrate crystal structure. Thus the unit cell boundaries shown in Figure 40 do not necessarily contain units with perfect translational periodicity. The hydrated crystal structure incorporates a substantial distortion away from hexagonal packing of silk helices due to the presence of hydration water located preferentially in the serine-rich (010) planes of the hydrated silk III structure. A comparison of the d-spacings and approximate relative intensities from the model diffraction and the observed experimental diffraction patterns is given in Table 6.

The model structure was relaxed prior to the diffraction simulation using molecular mechanics minimization and microcanonical dynamics, and an energy of -5 Kcal/mol·res was obtained, normalized by weight. Simulations of threefold silk helices using both molecular mechanics minimization techniques and molecular dynamics at 300 K result in reasonably low energies (-2 to -3 Kcal/mol·res) for the threefold helical conformation as it exists in the hydrate structure but without hydration water actually present. The energy of the structure is then reduced as water is added on the (010)

140

plane. The trend persists regardless of whether the energy from the hydrated structure is normalized by the molar weight of the unit cell, by volume or by number of residues. When water is present, the water molecules can act as hydrogen bond bridges between helices along the serine rich interhelical directions; this effect is observed in the molecular dynamics simulations.

In the interfacial films made using 5% aqueous fibroin and hexane (Groups A and C) large regions of a banded morphology coexist with the truncated hexagonal crystallites. As can be seen from Figure 37b, a typical diffraction pattern from the banded region, the banded morphology is not crystalline, rather it has the order of a mesophase. The elliptical ring of intensity in Figure 37b results from two overlapping arcs, as shown schematically in Figure 37d. One of the arcs is sharp and well-defined, corresponding to a spacing of 5.9 Å. The other is broad and diffuse and is due to a distribution of real space periodicities or to several closely-spaced overlapping weak reflections with a center of intensity corresponding to a spacing of 5.5 Å. In the silk III hydrate crystal, the threefold helical chains have a slightly elliptical cross-section, and thus there are three different nearest neighbor packing distances. The non-equivalent nearest neighbor interhelical packing distances result in non-equivalent interplanar d-spacings for the three planes of densely packed helices: (100), (110), and (120). These spacing result in three different strong equatorial reflections in Figure 36. The fact that the chains in the silk III hydrate crystal have non-equivalent near-neighbor packing distances prevents mesophase structures of these chains from achieving close interchain packing distances, such as 5.3 Å (110 reflection) and 4.0 Å (120 reflection), observed for

the crystalline structure. A mesophase interchain packing distance approximately equal to the largest crystalline nearest neighbor interchain distance is consistent with threefold helical chains in a less ordered packing arrangement. The real space periodicities that give rise to the arcs at 5.9 Å correspond to the nearest neighbor distances between the helices in the banded mesophase, and are similar to the largest equatorial spacing in the silk III hydrate crystal structure, nearly matching the 6.0 Å spacing of the 100 reflection (Figure 36b).

Table 6: Silk III Hydrate Model Data

	Silk III Hydrate spacings				
Experimental				**Model Simulation**	
Index	d spacing	relative Intensity		d spacing	relative Intensity
equatorial					
100	6.1	ms		6.1	m
020, 110	5.4	ms		5.3	s
120	4.0	s		4.0	m
240	2.0	w		2.0	w
non-equatorial					
101	5.3	mw		5.6	m
121	3.8	mw		4.0	mw
102	4.5	m		4.5	
113	3.3	w		3.0	mw
123	3.0	vw		3.0	vw
024	2.6	mw		2.8	w
202	2.6	mw		2.8	w
135, 205	2.0	w		2.0	w
125, 025	2.2, 2.4	mw		2.3	mw
	partially crystalline mesophase				
002	6.8	w		6.8	mw
003	4.4	w		4.5	w
004	3.4	w		3.4	-------

There are regions where the fibroin in the banded morphology has partially erystallized, resulting in diffraetion patterns that have many features in eommon with those of the lamellar crystallites. A diffraetion pattern from of one of these regions, along with a defocused diffraetion image showing the orientation of the banding with respect to the diffraction pattern, can be seen in Figure 41a and Figure 41b. With the notable exception of the equatorial refleetions, this diffraetion pattern ean be indexed as shown in the schematic in Figure 41c, using the same orthorhombie strueture as the lamellar diffraction. The differenees in indiees of refleetions present among Figure 36e, Figure 38c, and Figure 41c are due to the different tilt eonditions of the erystallites or crystallized mesophase giving rise to the diffraction patterns. In Table 6, the reflections from the diffraetion pattern in Figure 41a, which represents a region of the banded morphology that lacks fully erystalline order, are ineluded and are listed separately. The absence of the 110, and 120 reflections, corresponding to d-spacings of 5.3 and 4.0 Å, along the equator in Figure 41 are due to incomplete crystallinity in the diffracting region. Occasionally portions of material near the edges of the banded strueture exhibit better crystallinity yielding diffraction patterns with the same sequenee of equatorial reflections as the fully crystalline lamellar diffraction. The faet that the same erystalline structure is observed in the hexagonal lamellae and in the most highly crystalline regions of the banded morphology suggests that the same chain conformation is present in both the banded morphology and the hexagonal lamella.
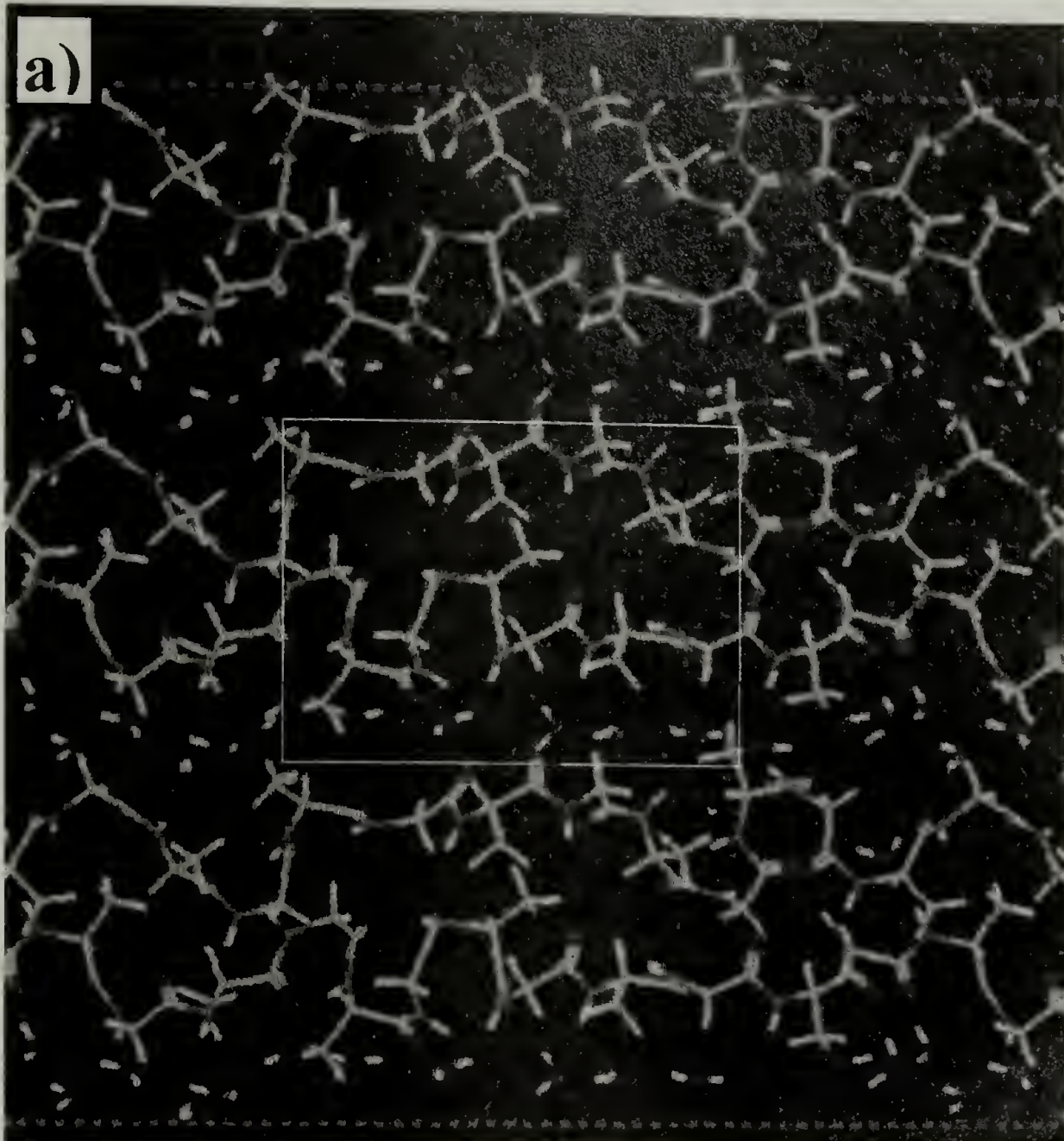
Figure 40: Hydrated silk III model with an orthorhombic two chain unit cell. The images shown are of "supercells" containing multiple unit cells. The box is a single unit cell. (a) View looking down the *a* axis (b) View looking down the *c* axis (chain axis). The light colored chains are the contents of a single two-chain unit cell.

Figure 40: *continued*

b)

a)

c) 3.2 $\overset{\circ}{A}$ 014 ___004__ 3.4 $\overset{\circ}{A}$

4.2 $\overset{\circ}{A}$ 013 ___003___ 4.4 $\overset{\circ}{A}$

002 6.8 $\overset{\circ}{A}$

• • 5.9 $\overset{\circ}{A}$
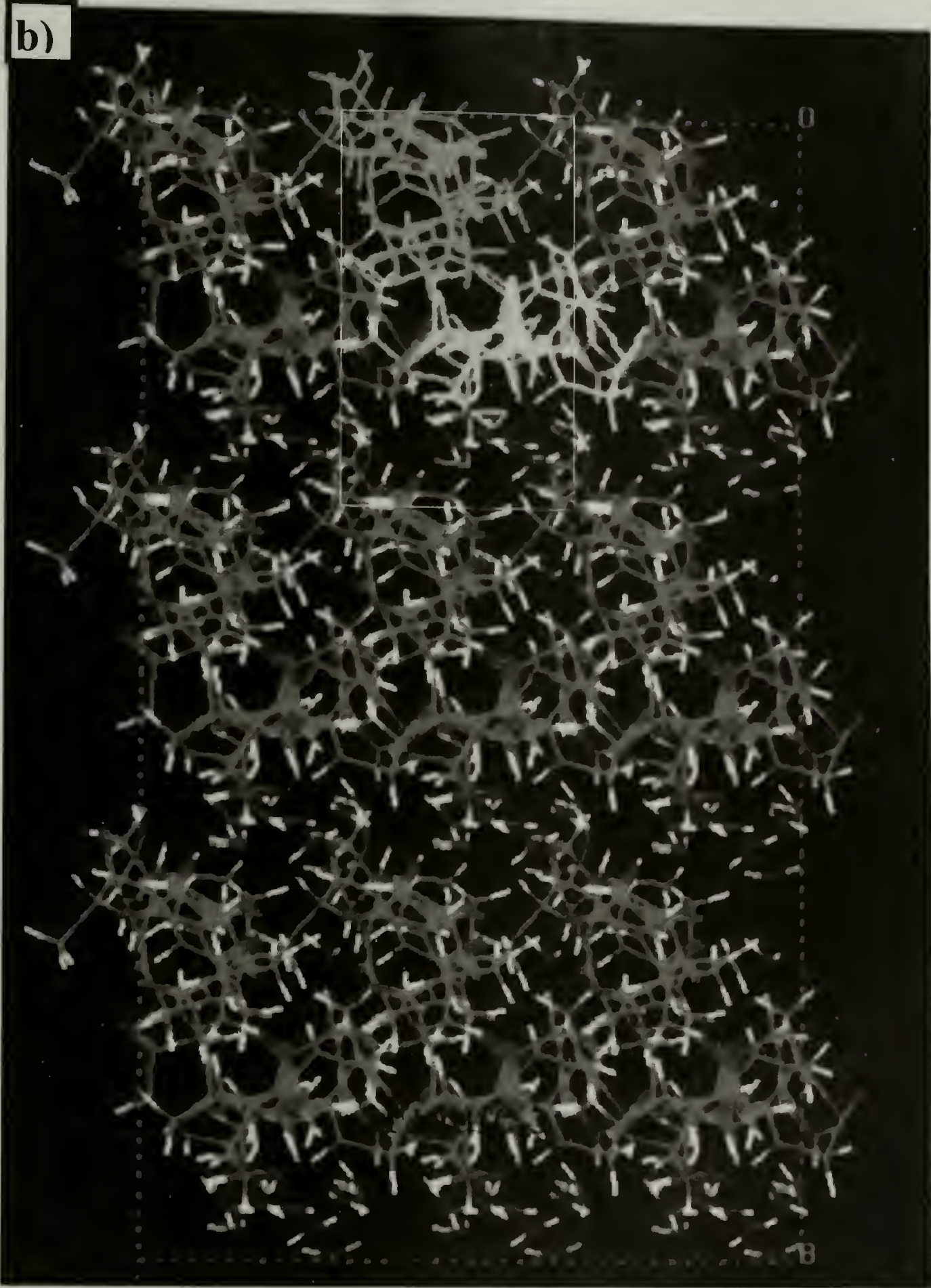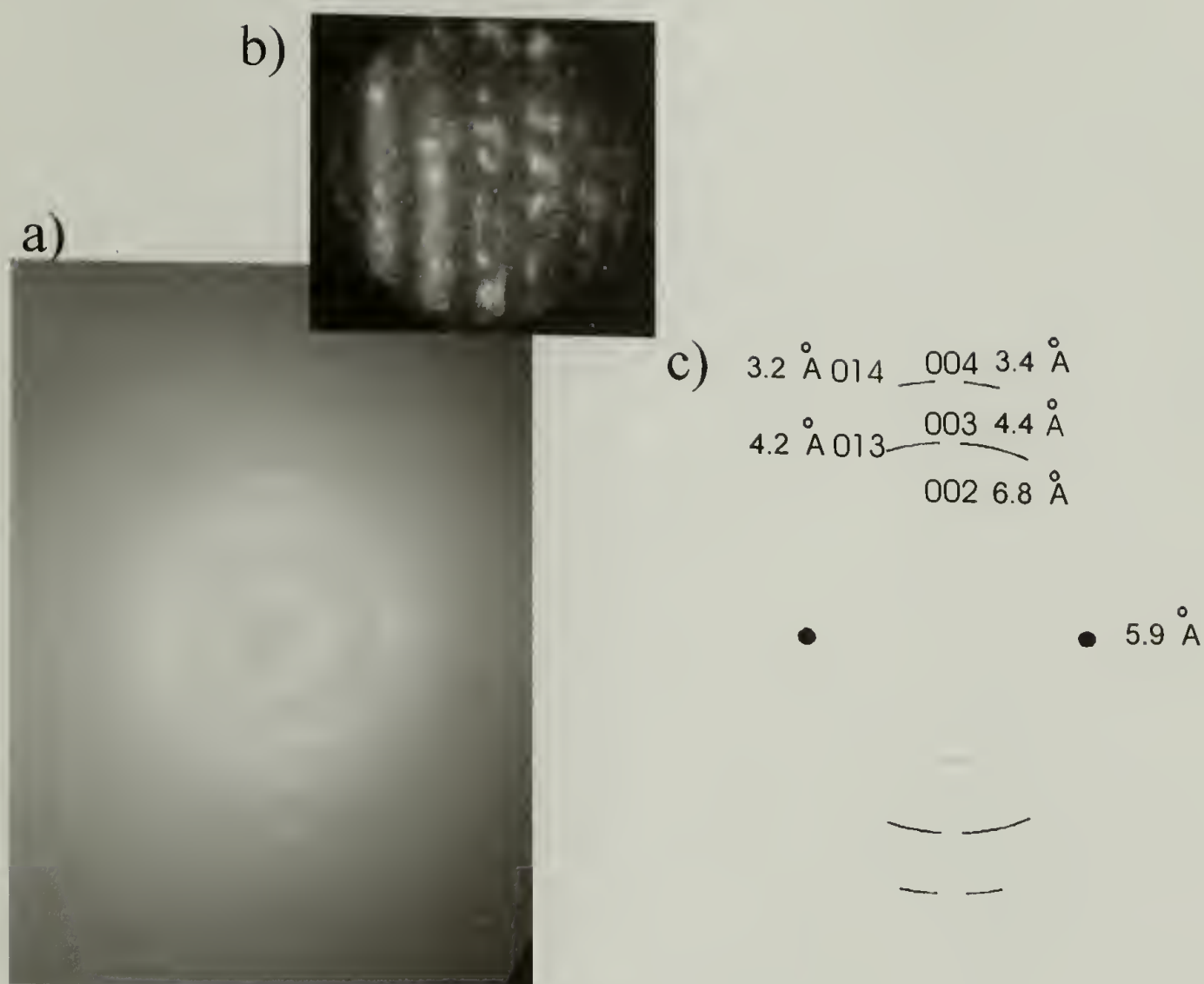
Figure 41: Diffraction pattern and corresponding defocused diffraction image for a region of the banded morphology that has partially crystallized. (a) Diffraction pattern. Although the reflections are largely clustered about the equatorial and meridional axes, off axis reflections indicate limited three-dimensional order. (b) Defocused diffraction image shows the orientation of the bands (c) Diffraction pattern schematic. The indices shown were obtained using the silk III hydrate model.

In the banded regions with less perfect crystallinity, the crystalline lattice is essentially preserved, but disordering causes the loss or merging of some reflections, notably the 110, and 120 in Figure 41. Even in regions that have not crystallized and retain mesophasic order, as in Figure 37, the diffraction pattern indicates d-spacings characteristic of the fibroin molecules in the same threefold helical chain conformation as found in the hexagonal crystals. The differences among the diffraction patterns of the lamellar crystallites, partially crystallized banded morphology, and non-crystalline banded morphology are due to different degrees of order in the lateral packing of chains, and also due to the local sample orientation.

The defocused diffraction images obtained from the partially crystallized and mesophasic regions in the banded morphology suggest a chain orientation that places the helical axis in the plane of the bands as shown schematically in Figure 42. In other words, the helical axis, is always perpendicular to the direction normal to the bands. Different meridional reflections are observed in different regions of the mesophase that have partially crystallized, indicating a change in the orientation of the helices in different parts of the mesophase. The differences in helical orientation in different parts of the mesophase are due to a helical axis orientation that rotates out of the plane of the banded sample film, as shown schematically in Figure 42. In addition, the orientation of the chain axis with respect to the bands in the partially crystalline regions is consistent with the data for the mesophasic region, indicating that the chains retain the same approximate director orientation from the mesophase when they crystallize. This partial crystallization that retains the general orientation of a preexisting liquid crystalline phase

is reminiscent of the lamellar decoration technique used to enhance TEM imaging contrast of defects in nematic liquid crystalline polymers.[146,147] The orientation of the helical axes normal to the bands and the rotation of the helical axes along the banding are behaviors expected for a cholesteric liquid crystal (or its solid phase analog), indicating that a cholesteric free surface may be responsible for the banding.

In the diffraction pattern from the partially crystallized banded structure (Figure 41a) there is a major equatorial reflection at 5.9 Å.   In diffraction from the lamellar crystals this single reflection is split into three closely spaced equatorial reflections, the 100, 110, and 120. The 100, 110 and 120 have d-spacings of 6.0, 5.3, and 4.0 Å and result from planes of nearest neighbor packing in the lamellar crystals. The single equatorial reflection at 5.9 Å seen in diffraction from the banded structure thus occurs at a spacing that approximately matches the largest of the nearest neighbor distances in the lamellar crystals, just as in the non-crystalline cholesteric portion of the banded structure. Despite the apparent merging of equatorial reflections when diffraction from the lamellar crystallites is compared to diffraction from partially crystallized regions of the banded structure (Figure 41), the same unit cell can be used to index meridional and near-meridional reflections that are observed in both types of patterns.
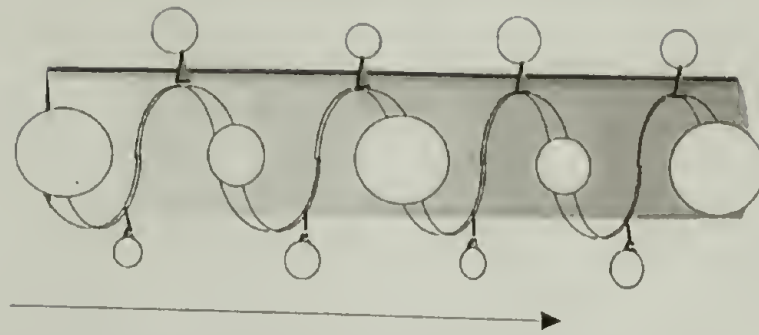
Helical axis direction
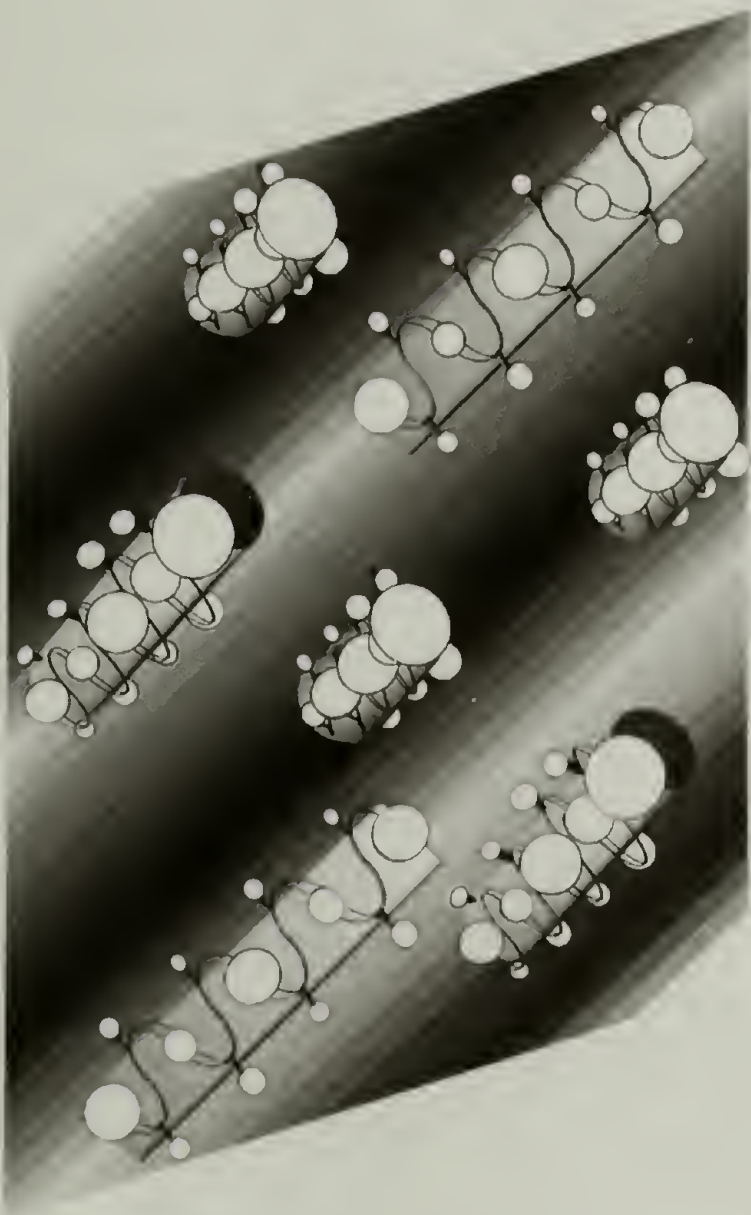
Direction perpendicular to bands

Figure 42: Schematic depiction of the orientation of the silk fibroin helices with respect to the bands in the banded morphology. The helical axis (c axis) is always perpendicular to the direction of banding.

The reflections on the meridion are due to periodicities along the helical axis, while the equatorial reflections arise from the packing arrangement lateral to the chains. Since the differences in the two types of diffraction patterns, lamellar vs. banded morphology, are most strikingly apparent in the equatorial reflections, the differences between the two structures are likely to involve a difference in the packing arrangement lateral to the helical axes. This change will result in a loss of distinction between the nearest neighbor packing directions for the banded structure. The difference is thus attributed to a lower degree of packing order in the banded regions, indicating a mesophase rather than full three-dimensional crystalline order. The spacing of the single, merged equatorial reflection in the partially crystalline banded region is also the spacing of the sharper of the two arcs (5.9 Å) observed in diffraction patterns from the non-crystalline regions of the banded structure shown in Figure 37. This arc has the same orientation with respect to the banded morphology as the equatorial reflections in diffraction from the partially crystalline regions of the banded morphology (Figure 41). In computer simulations, when the chain packing in the silk III hydrate crystal structure is disordered, the calculated equatorial spots begin to merge, but strong meridional reflections remain, just as in the diffraction from the banded morphology.

The banded morphology is very similar to the morphology observed for the freeze-fracture surfaces of cholesteric liquid crystals.[47,148-153] The diffraction data obtained from the banded morphology in Figure 37b indicate a fairly low degree of order, and are generally consistent with a mesophase. Even the partially crystalline

diffraction pattern in Figure 41b, while it can be indexed using the crystalline model, has features that suggest mesophasic rather than fully three dimensional crystalline order. Except for a single equatorial reflection, the pattern in Figure 41b consists entirely of apparently meridional reflections and of reflections lying close to the meridion. A pattern with only reflections on the equator and meridion is generally indicative of a nematic or cholesteric liquid crystalline structure with one dimensional order, but the presence of these off-meridional reflections indicates limited two dimensional order, perhaps involving small crystallites which have formed in the banded texture. The banded repeat distance of about 1 μm, is very large compared to the helical cross-section of a fibroin molecule, but is not inconsistent with the lengthscale of cholesteric pitch in a biological cholesteric mesophase.[151]

5.5 Conclusions

Silk fibroin exhibits surfactancy in conjunction with a novel interfacial crystal structure found for fibroin at the air-water and oil-water interfaces. The observed interfacial crystal structure incorporates an approximately threefold, irregular helical conformation that separates the hydrophobic alanine and hydrophilic serine residues to opposite sides of the interface. Since this conformation is strongly associated with conditions under which the fibroin exhibits surfactancy and provides a mechanism for the observed surfactancy, the threefold conformation is likely to be largely responsible for fibroin surfactant behavior.

While initial studies focused on the air-water interface, an oil-water interface was desired for comparison. At the water - hexane and water - chloroform interfaces the threefold conformation of silk fibroin is observed in large single crystalline lamellae, and in a banded morphology. The diffraction obtained from lamellar crystallites and from crystallized regions of the banded mesophase indicate that the silk is crystallizing in a structure incorporating a similar roughly threefold (6/2) helical chain conformation, as is observed for silk III at the air-water interface. However in contrast to the air-water interface, the crystallites at the water - hexane and water - chloroform interfaces are hydrated, resulting in a two chain orthorhombic unit cell.

CHAPTER 6

SERICIN AND FIBROIN

6.1 Summary

Transmission electron microscopy and electron diffraction have been used to observe single crystals of sericin in LB films made using an aqueous solution of fibroin and sericin regenerated from Bombyx mori cocoons. Several different single crystal zones were observed which nicely confirm the β conformation of crystalline sericin and also allow us to make definite assignments of unit cell parameters and reflection indices. The single crystal patterns for sericin are easily distinguished from reflections due to silk fibroin because some of the d-spacings are quite different, and because the sericin lamellae observed under these conditions are larger than the fibroin crystallites, leading to sharper reflections for sericin. Because the two crystal types are easy to distinguish, we have been able to determine that the sericin and silk crystals have very consistent relative orientations and a morphology that supports the epitaxial crystallization of sericin onto fibroin reported by Komatsu.[23]

6.2 Introduction

There are two major protein components in *Bombyx mori* silk fiber, fibroin and sericin. Fibroin forms strong semicrystalline fibrils that are held together by sericin, which has also been referred to as silk gum. Sericin may also play a role in the silk spinning process, acting as a low viscosity lubricant around the exterior of the fiber inside the insect as it spins silk. Fibroin forms the major crystalline component of silk fiber and its crystal structure in spun fibers has been extensively studied and reported.[1,4,5,8,10,11,13,16,17,28,29,39-42,49,54,56,58,138] In silk fibers the crystalline regions of the fibroin are in the β-sheet conformation, resulting in a monoclinic structure referred to as silk II.

Sericin is also known to crystallize, and has been reported to form crystals of the β-sheet conformation even in silk fibers without additional treatment.[7,65,69-71] The crystalline sericin in silk fibers occurs in the sericin layer adhering directly to fibroin, and the sericin crystals are believed to be somehow nucleated by the crystalline fibroin.[23,70,71] Sericin can also be made to crystallize in laboratory conditions, without fibroin present. While there have been several X-ray studies describing the crystallization of sericins and the major reflections (d-spacings) observed, little has been reported regarding the specifics of the sericin unit cell or even indices for the major reflections observed for this protein.[23,65,69-71] We have observed sericin crystallites in some preparations of thin films of aqueous silk, in mixtures of silk and sericin. These crystallites yield single crystal patterns indicating an orthorhombic crystalline structure,

as seen in Figure 43, just as would be expected from a β-sheet conformation. In addition, TEM morphological images, such as the one in Figure 44 demonstrate a tendency towards epitaxial.crystallization for sericin crystallites onto other sericin crystallites and onto silk fibroin crystallites present in the sample. Diffraction from regions containing crystals of both proteins, as shown in Figure 45 always reveals a faint ring for the small silk crystallites in the film and a sharper, more complete pattern from one or several larger single crystallites of sericin. The patterns from the two proteins, fibroin and sericin are usually in the same orientation relative to each other. Morphology images show silk film with striations or and large sericin lamellae projecting out from the film. The sericin lamellae always occur at the same angle to the striations in the silk film, suggesting epitaxial crystallization, and corroborating the diffraction evidence. Because the reflections for fibroin are significantly broadened and are also weak, it is difficult to say with certainty which crystalline form of fibroin is present However the strongest fibroin reflection at approximately 4.5 Å and the striated (banded) morphology strongly suggest that the fibroin is in the silk III conformation.
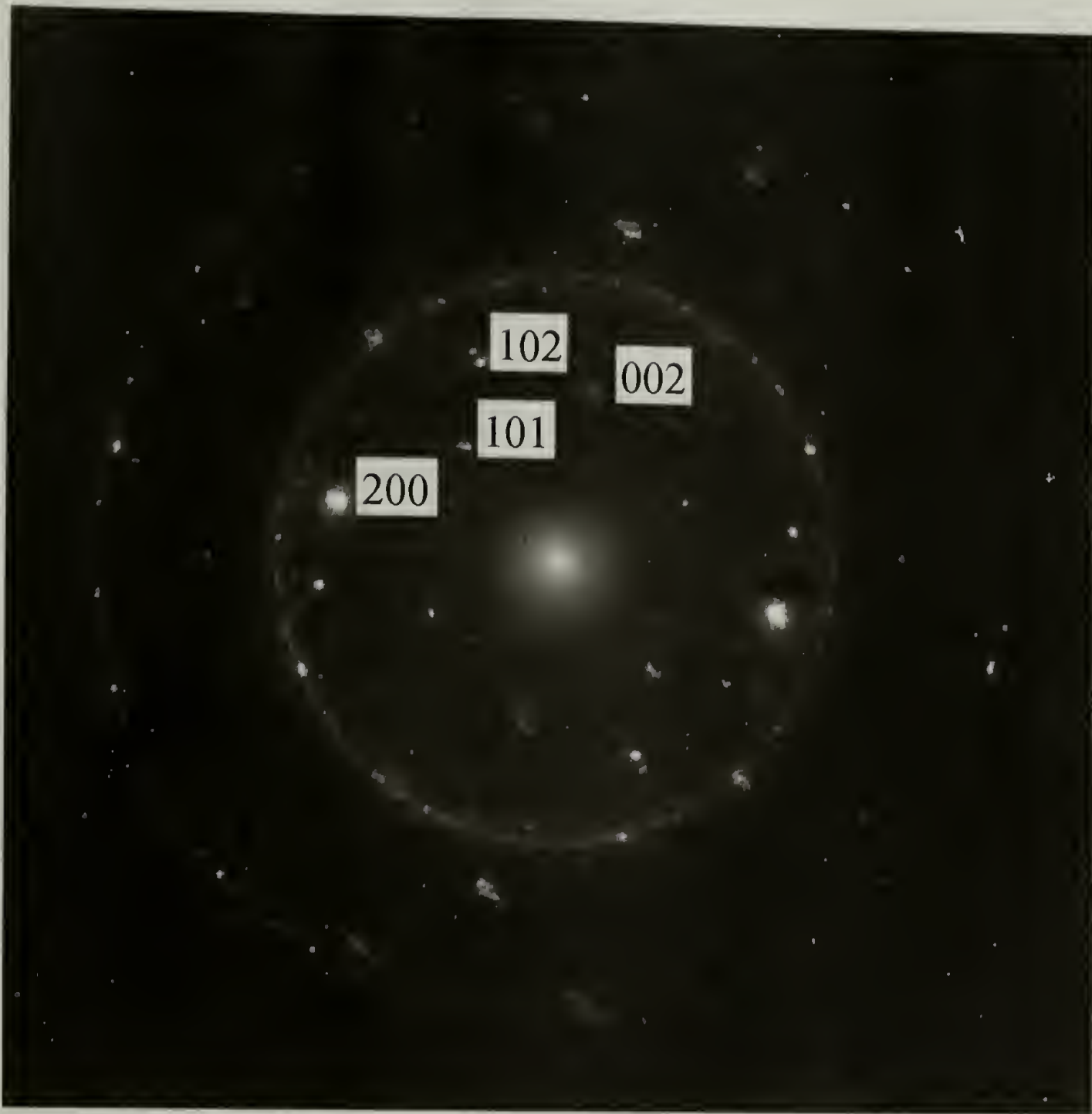
Figure 43: Sericin single crystal pattern, [010] zone. Note that for β-sheet sericin the chain axis is the *b* axis.

Figure 44: Sericin crystallites growing from an oriented fibroin film
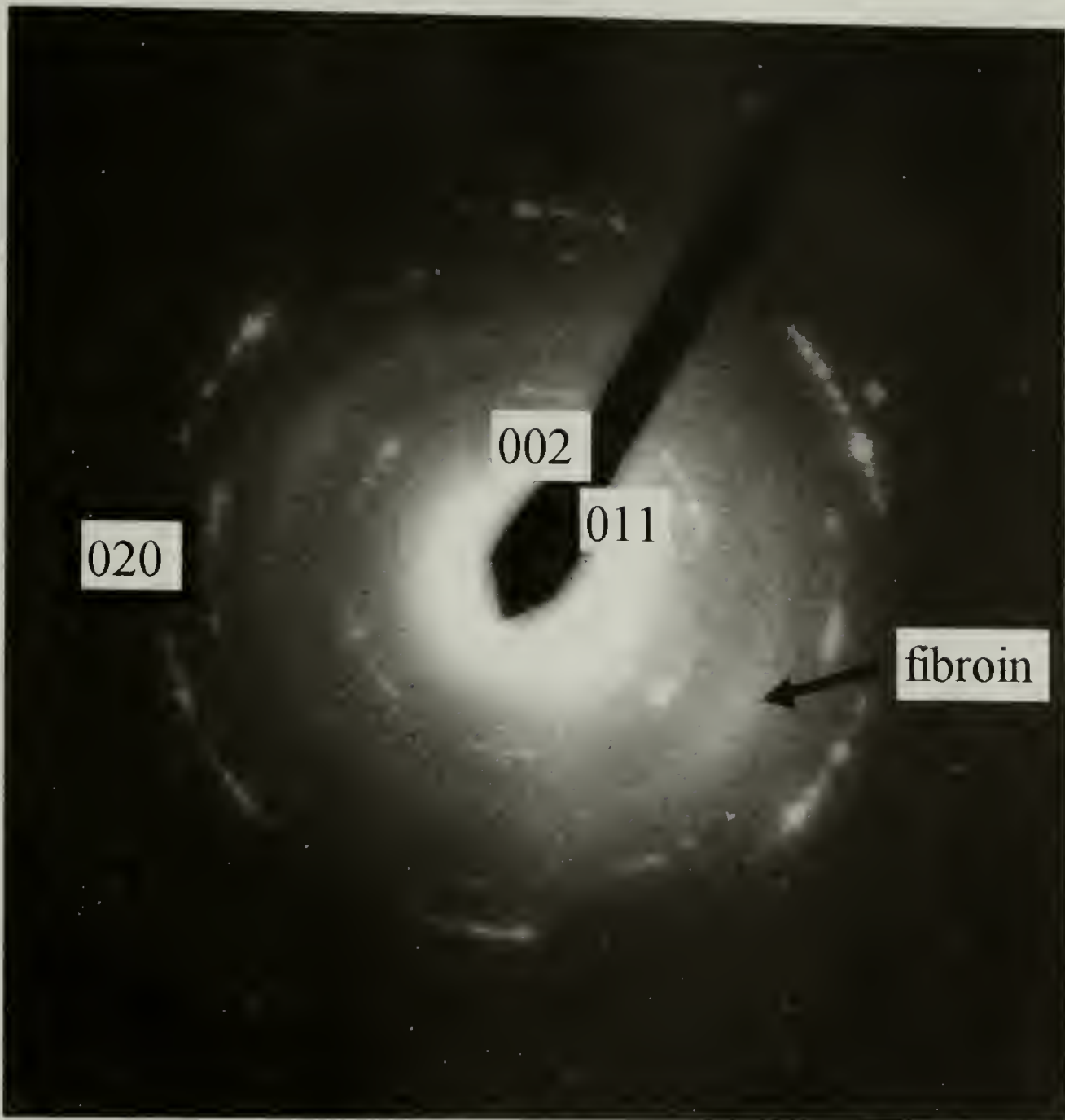
Figure 45: Diffraction pattern from the morphology in the previous Figure, [100] sericin zone axis. The sericin diffraction pattern is overlayed on an oriented polycrystalline fibroin diffraction pattern, labeled "fibroin".
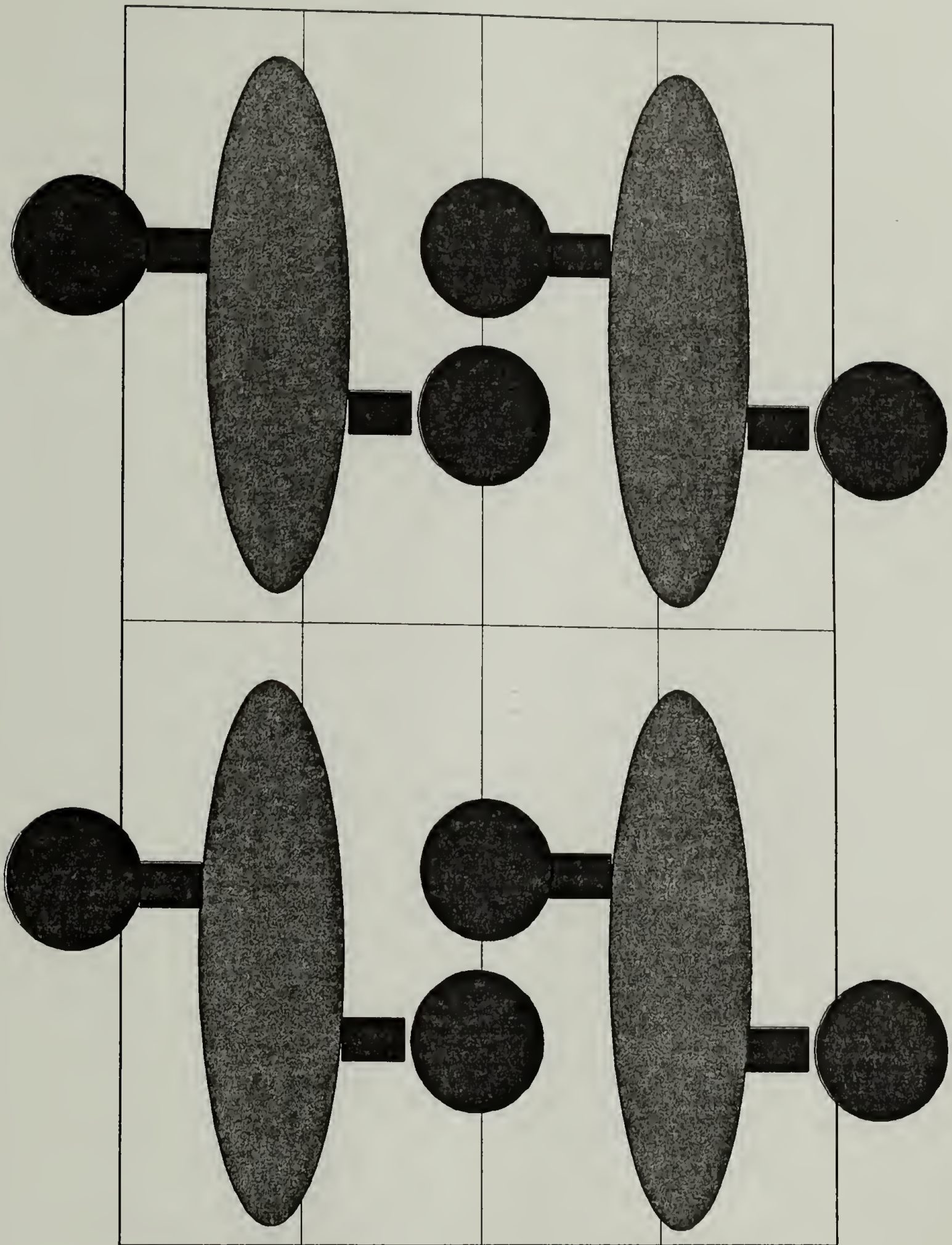
Figure 46: Asymmetric β–sheet. When sericin is in this arrangement the backbone atoms are not centered on the 002 planes. The protruding serine and threonine sidegroups midway between the β–sheets intensify the 004 reflection at the expense of the 002 reflection.
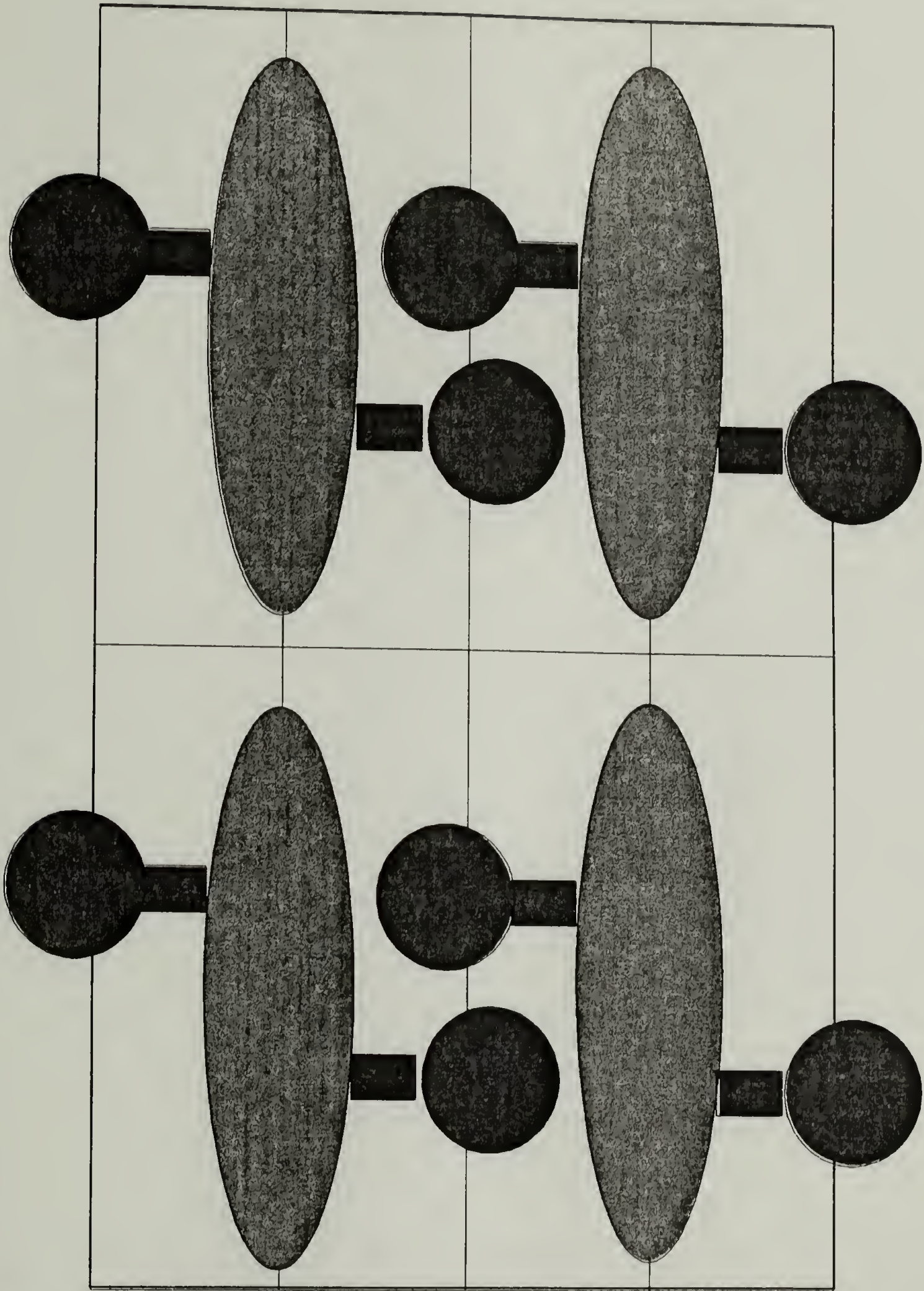
Figure 47: Symmetric sericin β–sheet with protruding serine and threonine groups. The strongest periodicity is in the (002) planes.

6.3 Preliminary Data

The films containing sericin in addition to fibroin behave differently from fibroin films from the air-water interface. The sericin-containing films form an ice layer readily at cryogenic temperatures, despite the high vacuum in the 2000FX electron microscope. When specimens are removed from the cryogenic holder, a layer of frost or dew often forms on the grids, effectively exposing the sample film to moisture. This was found to have little effect on the structure of the fibroin films, when they were subsequently viewed for the second and even third time in the electron microscope (under cryogenic conditions). The sericin-containing films exhibit marked changes in crystallinity when they are exposed to moisture, with sharper diffraction patterns appearing the second time a grid is viewed, after it has been exposed to moisture from dewing. These patterns have d-spacings that are the same as reported d-spacings for sericin, but cannot be indexed in a consistent fashion using the published crystallographic information available for sericin. The crystallites appear more beam stable than fibroin single crystals, but still degrade noticeably in the electron beam, suggesting that a protein salt complex may be responsible for the diffraction patterns. The hygroscopic nature of the material in the films and the observed moisture induced crystallization behavior have been reported for sericin.[7,58,66,67,69-71]

Single crystal diffraction patterns have also been obtained that match the spacings reported for β-sheet sericin and can be indexed on an orthorhombic unit cell which incorporates fiber period, intermolecular, and intersheet spacing distances values

that are very reasonable for a β-sheet crystal of serine. Komatsu reports X-ray crystallographic data characterizing the orthorhombic β-sheet structure of sericin, but only the fiber period is given for the unit cell.[23] Other reflections are reported and diagrammed, but are not indexed, labeled, or described. Nevertheless, assuming that his X-ray fiber diagram is drawn approximately to scale, the d-spacings and locations of the reflections that he observed in X-ray fiber diffraction are consistent with the d-spacings and symmetry of the single crystal diffraction patterns that we observe in the electron microscope.

6.3.1 The Unit Cell for β–sheet sericin

Two different, perpendicular diffraction zones have been observed for β-sericin, enabling all of the unit cell parameters to be determined. Diffraction from the [010] zone (looking down the chain axis) and from the [100] zone are shown in Figure 43 and Figure 45 respectively. Other zones are observed, but in these cases there appear to be multiple crystallites contributing to the diffraction pattern, and the contributions of the individual crystallites cannot be separated out reliably enough to analyze the patterns. In all cases the d-spacings observed match the major d-spacings reported for β-sheet sericin.

The sericin unit cell parameters are a = 9.2 Å, b = 6.8 Å, c = 11.6 Å; $\alpha = \beta = \gamma = 90°$. The intersheet distance of 11.6 Å (c axis) and the fact that the 001 reflection is not observed suggest a two-chain periodicity in the intersheet direction. The fiber repeat of 6.8 Å is shorter than the 7.0 Å observed in the silk II pleated β-sheet

structure. A similar relative compression is observed when comparing polyglycine I to silk II. Polyglycine I is in the fully extended β-sheet conformation with a fiber repeat of 7.05 Å, whereas silk II is in a slightly more compressed β-sheet conformation, with a fiber repeat of 6.99 Å. The fiber axis compression and pleating of the silk II β-sheet are ascribed to the presence of the alanine and serine sidechains. Sericin contains large proportions of serine, glutamic acid, and aspartic acid. While the crystallizable sequence involved in β-sheet crystallization of sericin is not well characterized, the large intersheet distance of 11.6 Å suggests large residues residing in the intersheet region. Incorporation of serine and larger residues into the β-sheet crystals of sericin would require further pleating of the β-sheets in the structure, is consistent with the shortening of the fiber repeat compared to silk II and polyglycine I, and is also consistent with the intersheet distance. For comparison, in polyglycine I, which incorporates a single chain in the intersheet direction, the intersheet distance is 3.7 Å, corresponding to a two chain periodicity of 7.4 Å. Silk II is an asymmetric β-sheet with an intersheet spacing of 9.4 Å for two sheets, reflecting an expansion of the intersheet distance between every other pair of sheets to incorporate alanine, serine, and possibly tyrosine.
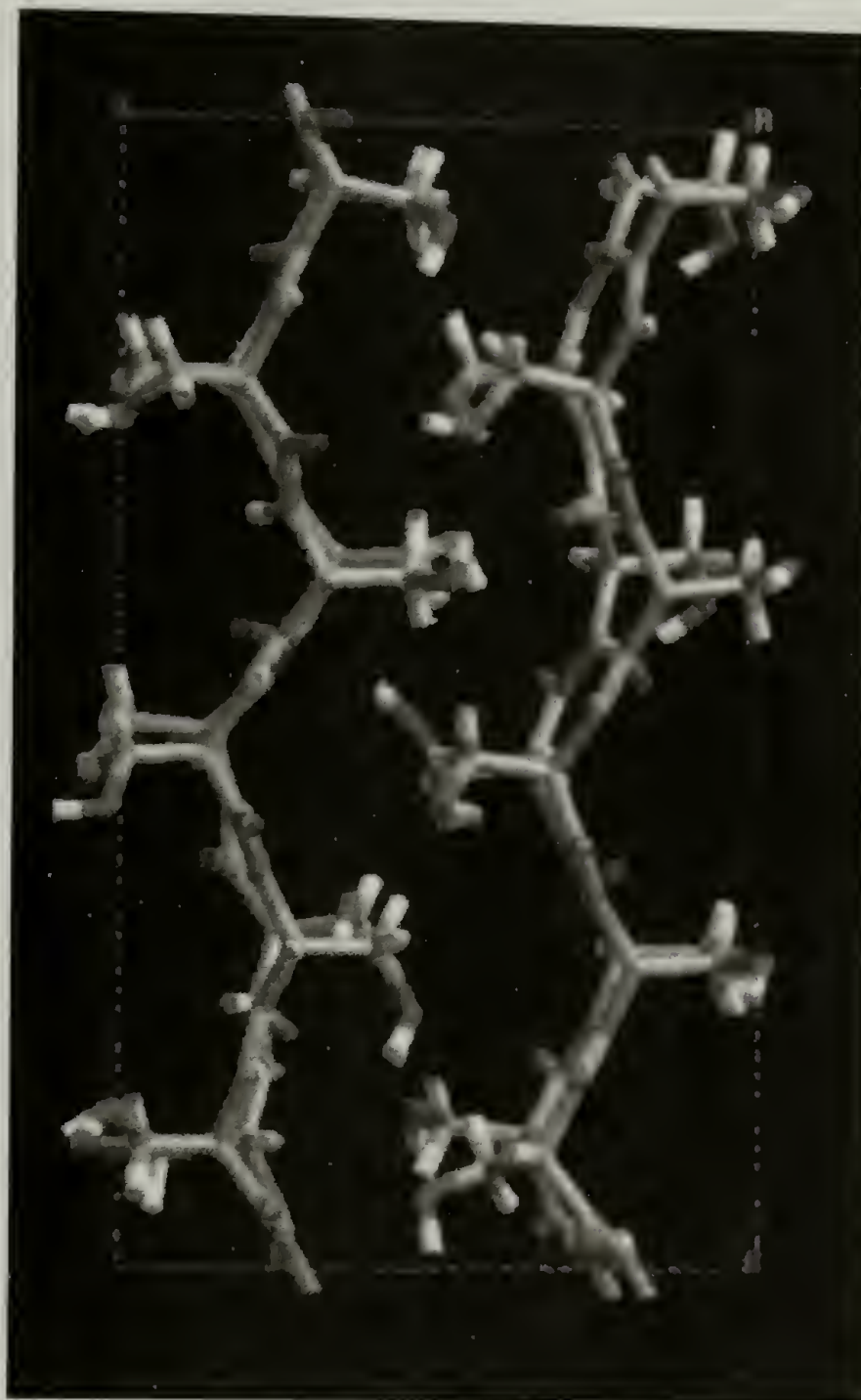
Figure 48: Side view of polyserine model.
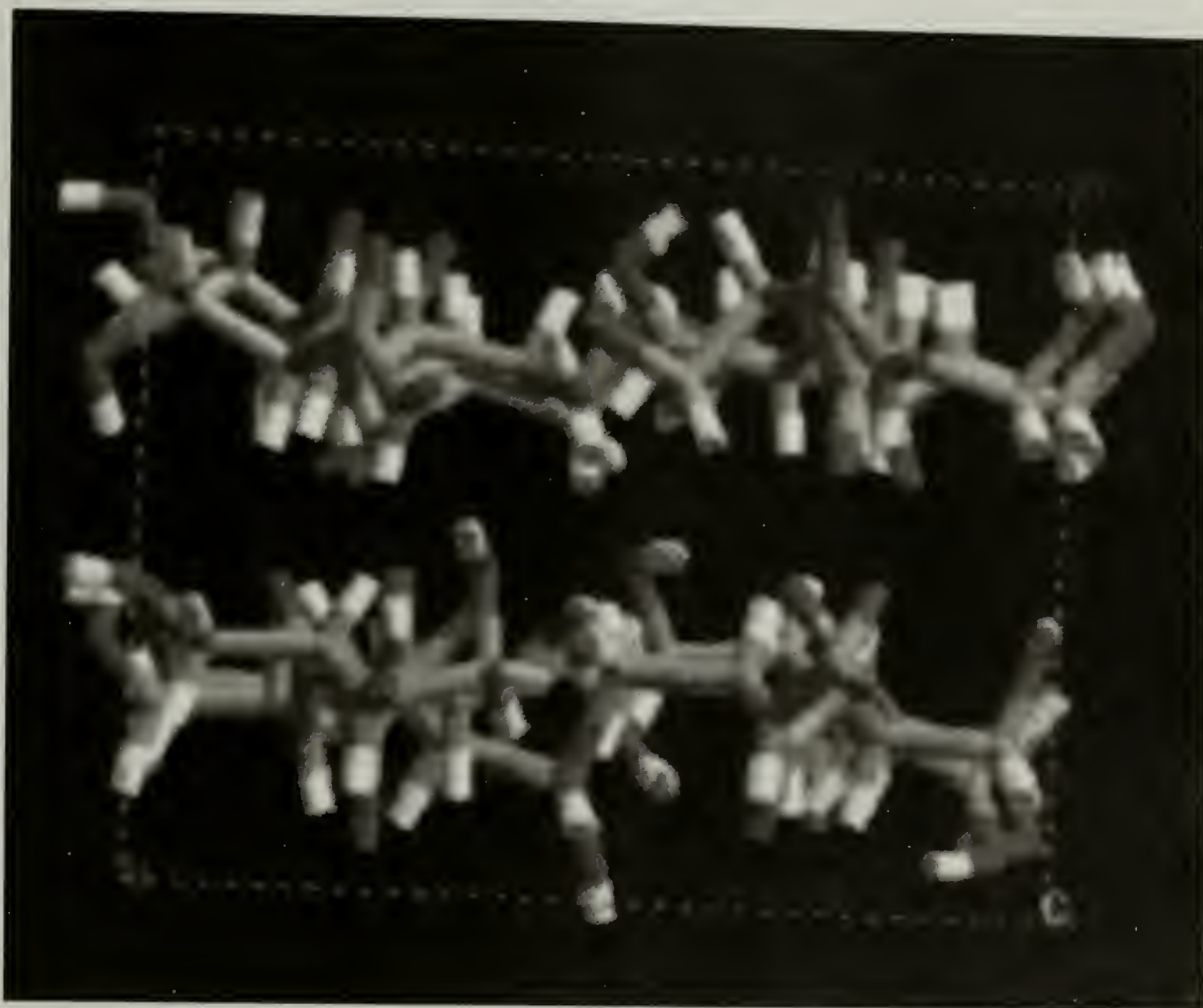
Figure 49: Top view of polyserine model. This view is looking down the chain axis which is the b axis in the sericin β–sheet crystal structure.
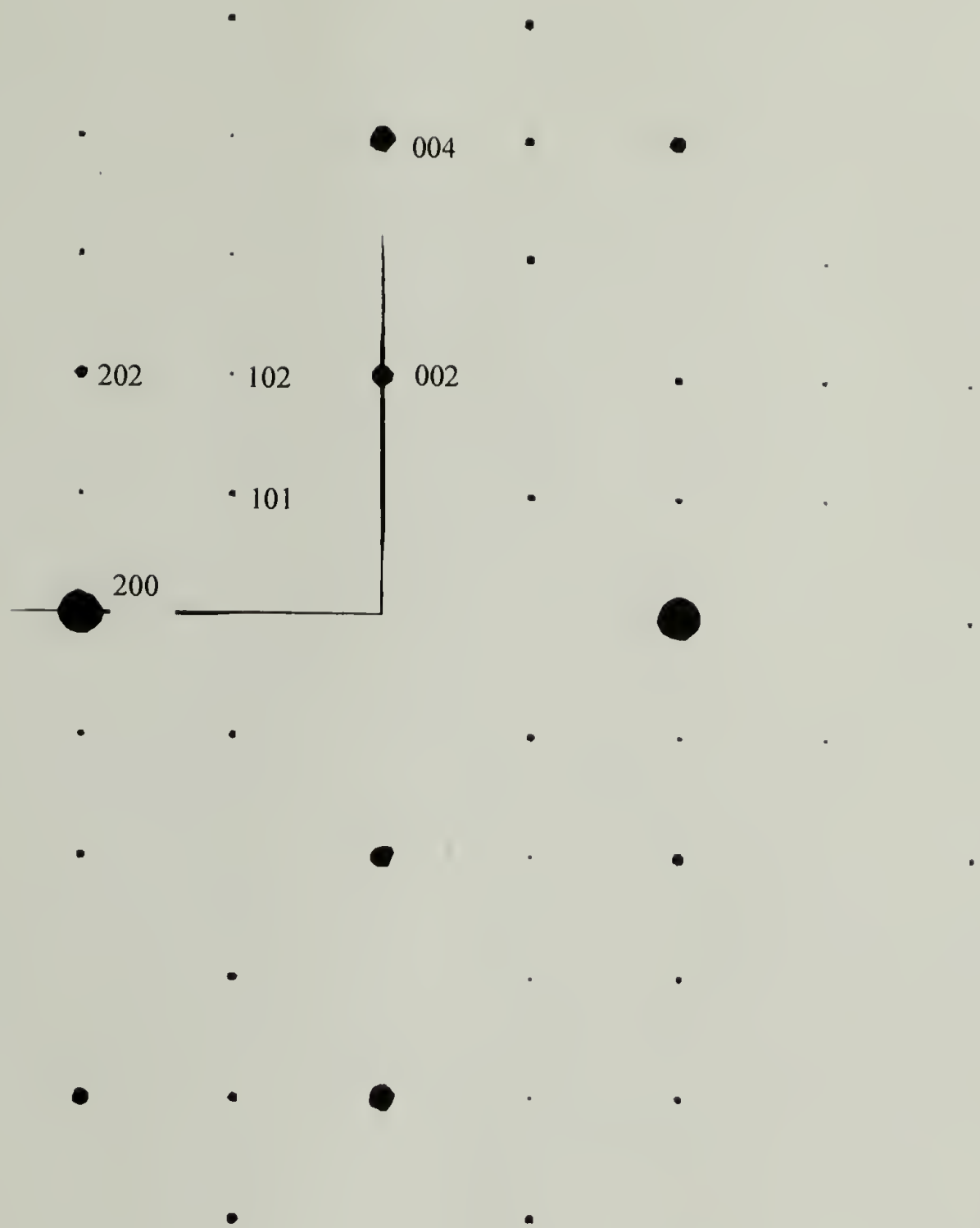
Figure 50: Simulated electron diffraction from a polyserine modes for β–sheet sericin. The simulated diffraction pattern is from the [010] zone axis. Intensification of he 004 reflection relative to the 002 reflection can be seen. This intensification is due oxygen and carbon from the serine residues' side groups protruding into the (004) planes.

The intersheet distance of 11.6 Å observed in sericin indicates that there are a large proportion of bulky residues present in the β-sheet crystal. A symmetric β-sheet structure with strongly scattering side-groups centered on the (004) planes, as shown in Figure 46, would result in strong 004 reflections and somewhat weakened 002 reflections, but can be ruled out because an asymmetric β-sheet is needed to obtain the observed two chain periodicity in the intersheet direction. In the experimentally observed diffraction pattern from the [010] axis (chain axis parallel to the electron beam), the 004 reflection is stronger than the 002 reflection. The strengthening of the 004 reflection at the expense of the 002 reflection results from the patterning and position of the large residues in the sericin sequence in the intersheet region.

There are two possible arrangements incorporating the larger residues. The first involves packing of large residues between every other pair of β-sheets in a highly asymmetric β-sheet structure. This situation would result in inclusion of two chains in the unit cell in the intersheet direction, but does not agree with the observed diffraction data. In the diffraction pattern from the [010] zone shown in Figure 43, the 004 reflection is significantly stronger than the 002, which is barely visible even after dodging. Odd 00l reflections are missing. The pattern of weak strong and absent reflections indicates that there is a high concentration of scattering atoms in the (004) planes, arranged periodically at an interval of 2.9 Å, canceling out most of the scattering from the (002). In a highly asymmetric β-sheet crystal, this would be achieved by having a close pair of sheets approximately 2.9 Å apart, and these pairs would be separated by 8.7 Å, with large residues protruding 2.9 Å into the larger intersheet distance. If the

168

distances between sheets were not extremely close to multiples of the interplanar distance between (004) planes, odd *00l* reflections would be observed. Since this situation requires a shorter distance which is less than the polyglycine I intersheet spacing, this arrangement is very unlikely.

The second possibility involves an asymmetric β–sheet arrangement with nearly equal distances between the sheets planes, as shown in Figure 46. The β–sheets would be approximately 5.8 Å apart and centered on the (002) planes. Large protruding residues such as serine and threonine would result in a high concentration of atoms with a higher scattering cross-section, such as oxygen and nitrogen, centered on the (004) planes. In this case, the 004 reflection would be intensified at the expense of the 002, and the *00l* reflections would still be absent. Assuming that the sericin can be approximated as polyserine, but may incorporate some larger residues an intersheet distance of 5.8 Å is reasonable. This distance is slightly larger than the longer intersheet spacing in silk II, which has predominantly alanine and serine between adjacent β-sheets alternating with predominantly glycine. There is a repetitive sequence in reported protein sequence data for sericins which incorporates a large number of serine and threonine residues. This sequence is (Val-Ser-Ser-Thr-Gly-Ser-Thr-Ser-Asn-Thr-Asp-Ser-Ser-Ser-Lys-Ser-Ala-Gly-Ser-Arg-Thr-Ser-Gly-Gly-Thr-Ser-Thr-Tyr-Gly-Tyr-Ser-Ser-Ser-His-Arg-Gly-Gly-Ser)$_{11}$. Threonine has one more methyl group than serine and has similar hydrogen bonding properties. In a β-sheet conformation, the sericin repetitive sequence has a higher proportion of threonine residues on one side of the β-sheet. Large stiff residues, namely histidine, tyrosine, and arginine are predominantly placed on the

other side of the chain. In an asymmetric β-sheet crystal, the repetitive sequence reported for sericin could thus account for the intersheet spacing and the intensification of the 004 reflection relative to the 002 reflection.

The interchain distance of 9.2 Å also indicates a two chain periodicity, however this could be due either to an antiparallel arrangement of chains in the interchain direction or to the pleating of the β-sheet, resulting in alternating displacements of chains relative to the (200) planes. The 020 and 040 reflections are clearly visible in Figure 45, favoring an arrangement of chains that are in registry in the 0k0 planes. Models using an antiparallel arrangement of chains have fairly weak relative intensity for the 020 reflection. Recall that the chain axis is the b axis. A parallel arrangement of chains within each β–sheet results in stronger (even) 0k0 reflections, but to reproduce the observed two chain periodicity in the [100], or interchain direction, there must be an additional factor, such as pleating present resulting in non-identical chains.

Models using polyserine to approximate sericin, and the unit cell determined using the single crystal diffraction evidence have been constructed to examine the possible variations in packing. Single crystal diffraction simulations and energies from molecular mechanics and quenched dynamics simulations have been used to evaluate the models. Simulated powder diffraction relative intensities were also calculated and compared to reported values. These data suggest which packing arrangements are most reasonable for β–sheets of sericin, but there are several caveats to be considered:

1) The exact sequence of residues present in the crystallites is not known and polyserine is only an approximate model.

2) Relative intensities from single crystal data depend on the orientation of the crystallite and may not be representative of the true diffracted intensity from the corresponding crystal planes.

3) Single crystal diffraction data are two dimensional sections of the reciprocal lattice, and a large number of different patterns are needed for a complete three-dimensional picture.

4) The X-ray powder diffraction data reported do not contain enough information to uniquely determine a crystal structure.

Even with these points in mind, a relative comparison of different packing possibilities is nevertheless feasible assuming that the sericin crystallites are a β-sheet structure and that the interchain and fiber axis repeats are similar to silk II.

Models were created that varied the displacement of polyserine chains away from the (200) and (002) planes, and the effects of parallel versus antiparallel packing were examined. Published X-ray diffraction data, single crystal electron diffraction patterns observed in thin films, and the energy per residue per mole were used to compare the structures. Coordinate files for the best structures obtained using this comparison are included in Appendix 1. Structures with a parallel arrangement of chains within each β–sheet, alternation of chain direction in the intersheet direction, a slight deviation from symmetric β–sheet packing, and a high degree of pleating, reproduce the observed single crystal diffraction data reasonably well, and are consistent with the reasoning detailed

earlier. The model for sericin using the polyserine sequence is shown in Figure 48 and Figure 49. A slight deviation from symmetric β-sheet packing was incorporated in the model and is evident, in Figure 49, a view looking down the chain axis (*b* axis). A simulated single crystal diffraction pattern from the [010] zone, looking down the chain axis, can be seen in Figure 50. The model used to calculate the simulated diffraction pattern incorporates a small displacement of the β–sheets away from the (002) planes of the crystal. In the polyserine model used, there is no physical rationale for such a displacement. However the repetitive sequence found in sericin would result in slightly asymmetric sterics in the β–sheet conformation, as described earlier. In diffraction simulations, the asymmetry in the intersheet spacing imposed in the model does in fact result in attenuation of the 004 reflection at the expense of the 002 reflection as can be seen from the simulated single crystal pattern in Figure 50. Comparison of the experimentally observed single crystal pattern shown in Figure 41 with the simulation in Figure 50 demonstrates that the major features of the observed diffraction are reproduced in the diffraction simulation from the polyserine model. Reflections such as the 101 and 103 also have approximately the correct relative intensity. Although the experimental data probably reflect additional effects such as hydration as well as chain placement, the reasonable match between experimental and simulated diffraction data suggest that the general features of the model used are a correct representation of the packing of sericin in the experimentally observed crystallites.

Diffraction patterns from LB films containing both fibroin and sericin are typically observed in the orientation shown in the experimentally observed diffraction

pattern shown in Figure 45. In the Figure, the reflections that can be indexed on the sericin unit cell are the sharp diffracted spots labeled as the 002, 020, and 011 reflections (and higher orders of these reflections). The d-spacings, indices and symmetry of the reflections in the sericin diffraction pattern indicate that the sericin diffraction pattern is approximately from the [100] zone. In addition to the (sharp) sericin single crystal diffraction pattern there is a more diffuse oriented polycrystalline pattern in the same Figure. This diffuse pattern is often observed with the sericin diffraction patterns from the [100] zone. The reflections from the diffuse diffraction pattern form a ring with increased intensity at two points 180° apart, indicating the presence of orientation. The intensity from the ring is centered at 4.5Å, the strongest d-spacing for fibroin in the silk III crystal structure, suggesting that silk III may also be present. Diffraction patterns identical to the diffuse pattern in Figure 45 are also observed in compressed LB films of pure fibroin. These films have a uniaxially oriented crystal texture with the chain axes of the crystallites perpendicular to the plane of the film as described in Chapters 2-5. In regions where the fibroin film has broken and rolled up at the edge of a film fragment, polycrystalline diffraction patterns such as the diffuse pattern Figure 45 are observed. The same biaxially oriented texture observed in the uncompressed fibroin films from the air-water interface can result in different appearances for diffraction patterns., depending on the angle at which the Ewald plane intersects the reciprocal space diffracted intensity distribution.

Reciprocal Space

Diffraction Pattern

a)

Ewald plane

b)
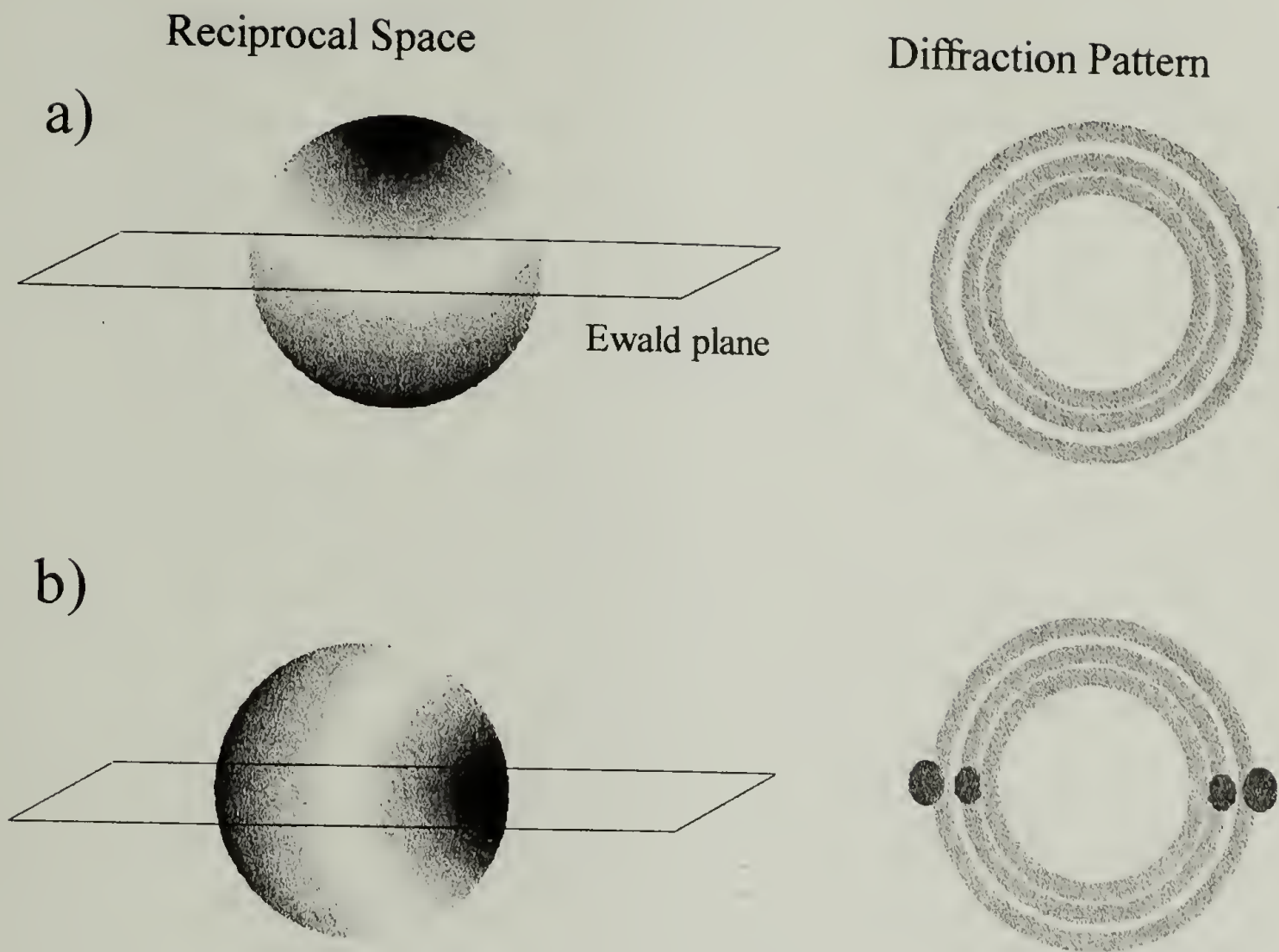
Figure 51: The diffraction pattern from a biaxial orientation of crystallites as seen at two different (perpendicular) orientations with respect to the incident beam. (a) Biaxial orientation observed in uncompressed air-water interfacial films of fibroin. The c axis is always in the plane of the sample film, and always perpendicular to the incident electron beam. The resulting diffraction pattern consists of polycrystalline rings with an even distribution of intensity in each ring. (b) Biaxial orientation observed at edges of 16.7 mN/meter LB films (uniaxially oriented) that have rolled up. The reciprocal space picture is the same, but the orientation of the crystalline texture with respect to the electron beam is different from the situation in (a). The resulting diffraction pattern has an uneven distribution of intensity, similar to that observed in the fibroin/sericin LB films.

As can be seen in Figure 51a, an orientation that places the c axis in the sample plane, but always perpendicular to the incident beam, results in a polycrystalline diffraction rings with an even distribution of intensity within each ring. A piece of uniaxially oriented LB film that has rolled up has the same biaxial orientation as shown in Figure 51b, but the orientation axes of the crystalline texture with respect to the electron beam are tilted 90° and the resulting diffraction pattern has two intensity maxima. In essence, the diffraction patterns from the uncompressed film and the rolled uniaxially oriented film represent two different intersections of the Ewald plane with the same diffracted intensity distribution.

Comparison with the situation in a rolled film shard suggests that the fibroin regions that yield diffraction patterns such as the one in Figure 45 have a biaxial orientation . Interestingly enough, the morphology image in Figure 44, corresponding to the diffraction pattern in Figure 45, reveals facetted lamellae protruding from a film region which possesses a striped or banded texture. The appearance of the film region supports the cholesteric-like orientation for the fibroin film suggested by the diffuse diffraction pattern in Figure 45. The lamellae protruding from the film region are parallel to each other, implying a preferred orientation between the sericin crystallites and the fibroin in the polycrystalline film. A preferred orientation between sericin and fibroin is also suggested by the repeated occurrence of diffraction patterns with the same sericin single crystal [100] axis diffraction pattern and fibroin oriented polycrystalline diffraction pattern. In separate studies of silk fibroin foams made by blowing air through purified

fibroin solutions, a banded or striped morphology similar to the film morphology in Figure 44 is observed and the corresponding diffraction patterns match a silk III hydrate structure in a cholesteric type of orientation. In fibroin films from an uncompressed air-water interfacial surface, similar banding is also observed and is associated with the silk III crystal structure and a left-handed threefold helical conformation (a 6/2 helix) of fibroin. The diffraction evidence from the films containing fibroin and sericin suggests epitaxial crystallization of the silk III fibroin structure and crystalline β–sheet sericin. Diffraction patterns from oriented fibroin and single crystals of sericin coexist and apparently retain a consistent orientation relative to each other. Morphology images, such as the one in Figure 44, further support epitaxial crystallization, with a preferred orientation evident between the striations or bands in the fibroin film and the sericin lamellae.

6.4 Conclusions and Suggested Experiments

In LB films prepared using a mixture of fibroin and sericin and a surface pressure of 16.7 mN/meter crystalline sericin and crystalline fibroin are observed. In some cases sericin single crystal diffraction patterns are observed with enough sharp crystalline reflections to deduce some general features of the unit cell and packing of crystalline sericin β-sheets. The sericin β-sheets are arranged asymmetrically within the crystallites, but the difference in intersheet spacings is small. The unit cell is orthorhombic with dimensions: 9.2 Å X 6.8ÅX11.6Å. The crystalline sericin is often observed with

crystalline fibroin, and there is evidence for a preferred orientation between the two types of crystallites, suggesting epitaxial crystallization. The fibroin appears to be in the threefold silk III conformation, but the diffraction patterns obtained for fibroin with sericin also present contain very few reflections and these are very diffuse, making the structure of the fibroin present uncertain.

Further LB experiments with films made by varying the proportions of sericin and fibroin and with films made using a variety of surface pressures would be useful in determining the influence of the sericin and fibroin crystal structures on each other.

CHAPTER 7

A COMPARISON OF FIBROIN INTERFACIAL AND BULK STRUCTURES

The monoclinic unit cell and deviation from a perfect hexagonal lattice for silk III at the air-water interface is supported by a variety of types of diffraction data. These include the single crystal diffraction evidence for a distortion away from hexagonal packing, the curvefitting of digitized data that support this distortion, and the presence of diffraction patterns from the aqueous-chloroform interface with sharp rings for the component reflections produced by a monoclinic unit cell. Additionally, highly oriented crystalline regions are occasionally observed in cast films made using aqueous solutions containing less than 1% fibroin by weight. In these instances, stress during drying may have induced orientation of the silk III crystallites. A large number of reflections are observed as separate localized arcs whereas in other film regions and samples broad overlapping diffraction rings are observed. These oriented patterns can be indexed on the silk III monoclinic unit cell, adding further confidence to the structural analysis presented in Chapter 3.

In addition to the supporting diffraction evidence, there are some features of the IR spectra reported for fibroin samples that are of interest with respect to the silk III structure. Many preparations of fibroin have been reported as mixtures of silk I and silk II using IR spectroscopic evidence, even though this method may not distinguish the silk

I crystal structure from amorphous silk fibroin. Unpublished IR data from the (liquid) air-water interface of aqueous fibroin in an LB trough from S.Riou and S. L. Hsu indicate similar absorbencies to fibroin preparations believed to contain a mixture of silk I and silk II. The possibility of a third crystal structure such as silk III, and the ability of IR spectroscopy to distinguish among closely related extended chain conformation such as the silk III helix and silk II β-sheet have not been considered. Infrared spectroscopy data reported for fibroin usually use amide I, II, and III bands to determine which structures are present. Silk I has reported amide I, II, and III absorbencies at 1650, 1535, and 1260 wavenumbers respectively whereas silk II has absorbencies at 1625, 1528, and 1235. Reported values for the silk I and II structures can vary by a few wavenumbers, especially in the case of silk I . The difficulties in distinguishing silk I from amorphous silk using spectroscopic techniques have caused some controversy over the IR spectrum of silk I. For comparison polyglycine I, a fully extended β-sheet structure, has amide bands similar to silk II, at 1630, 1530, and 1015 respectively. Polyglycine II, a crystal structure incorporating a threefold helical conformation, would be expected have absorbencies similar to threefold helical silk III. The values reported for polyglycine II are also somewhat similar to those generally attributed to silk I, at 1648, 1558, and 1026 wavenumbers respectively.

Using the assumption that the IR spectra are correlated to chain conformation, a comparison of the silk II, polyglycine II and silk III chain conformations yields some intriguing results. Silk I is not compared because its conformation is currently the

subject of some controversy. A single parameter, the rise per residue, can be used to combine the different conformational features into one easily comparable quantity. The rise per residue in silk II is 3.5 Å, whereas in fully extended polyglycine I the rise per residue is closer to 3.6 Å. Polyglycine II, is a fairly extended structure as well, with a rise per residue of 3.1 Å. Both the threefold helical structure and the β-sheet structure incorporate interchain hydrogen bonds that are perpendicular to the chain axis. Silk II also incorporates perpendicular hydrogen bonds. The conformation of the silk III threefold helix is best described as a left-handed 6/2 helix, making two full turns in six residues. The silk III structure involves a preferred location for serine, and the hydroxy group on the serine side chain can also form interchain hydrogen bonds. The change in the hydrogen bonding pattern around serine and the alternation of highly flexible glycine residues with larger residues, causes the silk III helical conformation to be irregular. The torsional angles and rise per residue vary depending on the location in the six residue repeated sequence (GAGAGS). The rise per residue for silk III, proceeding from amino to carboxy terminal in the sequence GAGAGS is 3.1Å, 3.3 Å, 3.1Å, 3.4 Å, 3.2 Å, and 3.4Å respectively, measured from the monoclinic silk III model. There are thus places along the silk III helix where the conformation locally resembles a β-sheet. Other portions of the chain have a local conformation the resembles a polyglycine II threefold helix.

It is thus questionable whether the silk III structure can be easily distinguished from the other structures formed by fibroin using IR spectroscopic methods without first

developing standard spectra for all three structures. In fact, in experiments with silk foams, IR spectroscopic data appear to indicate a mixture of silk I and silk II in samples where silk II X-ray and electron diffraction is never observed, only silk I and sometimes silk III. In the case of data gathered from the liquid surface of a protein solution , such as fibroin, the distinction between crystal structures is further complicated by the fact that in numerous studies of protein absorption to interfaces, an increase in β character was generally observed in the interfacial state as compared to the natured conformation. A notable exception is proteins that form helices with surfactant properties (only alpha helices were studied) . These proteins tended to have more alpha helical character at the interface. There is no reason to believe that the behavior of fibroin would not be analogous to the behavior reported for other proteins. Regions that do not have an available conformation that would impart surfactancy to the protein would contain more b conformation at the interface. The amorphous blocks of the fibroin chain would be expected to contain β-conformation at the interface. The crystallizable blocks may also contain some β-conformation, but with the sterically reasonable threefold helical conformation available imparting surfactancy to these regions, it is not a foregone conclusion that the β-conformation will be favored by the crystallizable fibroin segments at the air-water interface.
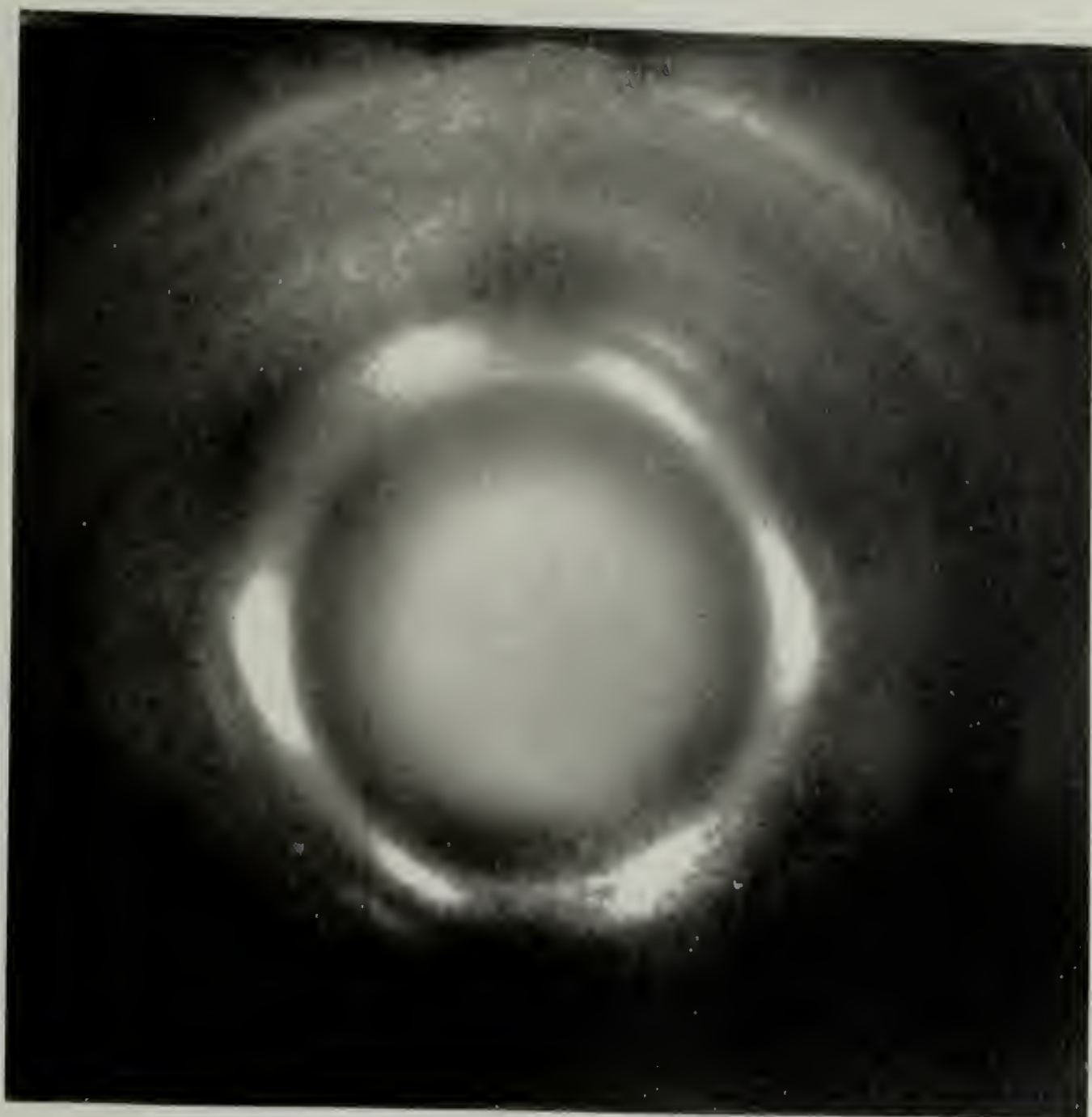
Figure 52: Diffraction pattern from an oriented region of a cast fibroin film.
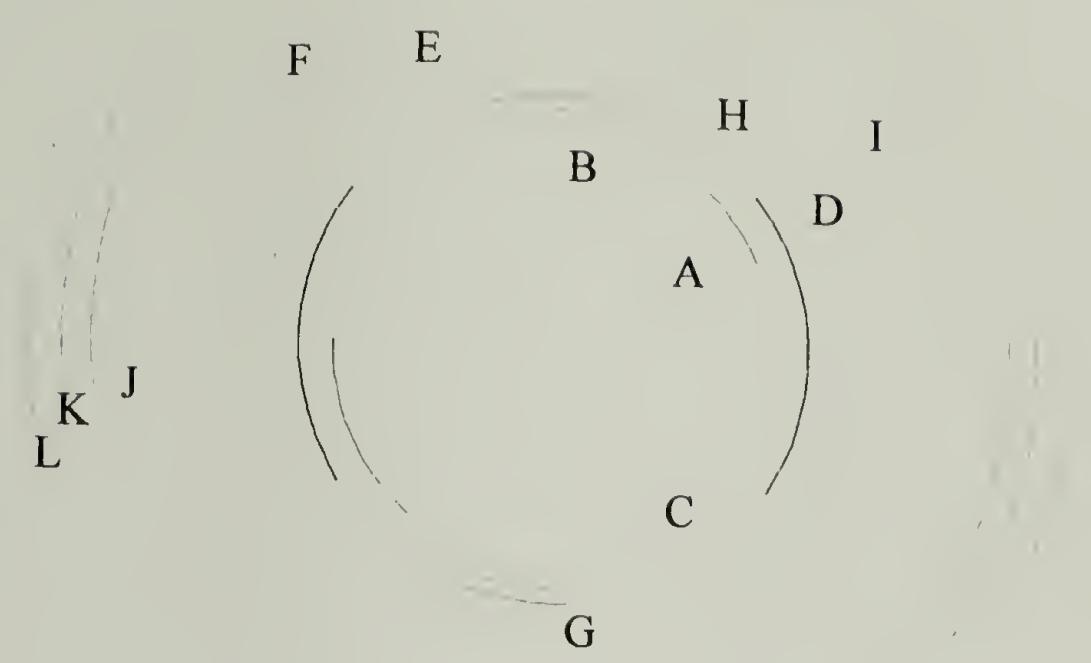
	D spacing	Index
A	5.58,4.88	111,112
B	4.56	200
C	4.45	020, some 021
D	4.34	2$\bar{2}$0
E	3.70	2$\bar{2}$2
F	3.19	224
G	4.23	202
H	3.82	204

	D spacing	Index
I	3.19 - 2.70	1$\bar{3}$2,006
J	2.34	2$\bar{2}$6
K	2.21	2$\bar{4}$4
L	2.07	404

Figure 53: Diffraction pattern schematic from an oriented region of a cast fibroin film.

Spectroscopic data do not distinguish which blocks of the fibroin molecule are in which conformation, nor do they reveal the crystallinity associated with a conformation in a particular preparation. Localized conformational and crystallographic information is however available through electron diffraction techniques. The silk II crystal structure is rarely observed in films of fibroin from the air-water interface. The only instance in which silk II has been observed is in films from surfaces compressed to 34 mN/meter and subsequently aged. In films from uncompressed surfaces or where a lower compression of 17 mN/meter was used silk III predominates, silk I is occasionally observed, and the silk II crystal structure is notably absent.

Also of interest is the fact that the distortion from hexagonal packing is more noticeable at cryogenic temperatures than in data taken at room temperature. In fact the magnitude of the distortion in g observed at cryogenic temperatures and the accompanying changes in lattice spacings are very similar to the behavior observed by Brack and Spach for poly(Ala-Gly-Gly) in crystals that incorporate a threefold helical conformation. In the case of poly(Ala-Gly-Gly), even though there is a higher proportion of glycine than in the (Gly-Ala-Gly-Ala-Gly-Ser) hexapeptide fibroin sequence, strict alternation of glycine and larger residues is not provided for in the sequence. As a consequence, a three-fold helical conformation of poly(Ala-Gly-Gly) would have a column of alanine residues running parallel to the helical axis and two columns of glycine residues. The disruption in the hydrogen bonding pattern is minimized if the helices pack into sheets, with the columns of alanine residues between

the sheets and the polyglycine II hydrogen - bonding pattern preserved within the sheets. Sheet-like packing in poly(Ala-Gly-Gly) should thus produce a distortion in the unit cell and create non-equivalent nearest neighbor interhelical packing directions. The strong reflections due to interhelical nearest neighbor packing in poly(Ala-Gly-Gly) occur at 4.30 Å and 4.63 Å. These d-spacings are both slightly larger than the 100 spacing of 4.15 Å observed in polyglycine II and are similar to the 4.34 Å and 4.56 Å observed for fibroin. Brack and Spach observed two slightly different structures for poly(Ala-Gly-Gly). One of these was believed to be a hydrate and incorporated a larger split in the interhelical nearest neighbor packing directions with d-spacings of 4.24 Å and 4.68 Å. In fibroin, distortions in hexagonal packing are slightly more pronounced at cryogenic temperature than at room temperature. The difference is especially evident in measurements of single crystal patterns. The similarity in behavior between fibroin and poly(Ala-Gly-Gly) and its hydrate suggest similarities in the unit cell packing and hydration mechanism and support the sheet-like packing of fibroin helices.

As the fibroin is hydrated further, the silk III hydrate structure becomes distorted to the extent that the packing is best described using an orthorhombic two chain unit cell. It is interesting to note that there is a corresponding change in crystallite shape, from apparently hexagonal crystal lamellae in the LB films, where hydration within the crystals is minimal if present, to a truncated triangular structure at the liquid-liquid interfaces made by aqueous fibroin with hexane and chloroform as seen in the two images on the left hand side of Figure 54.

185
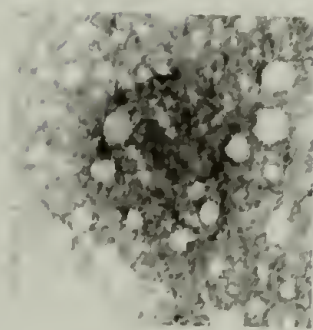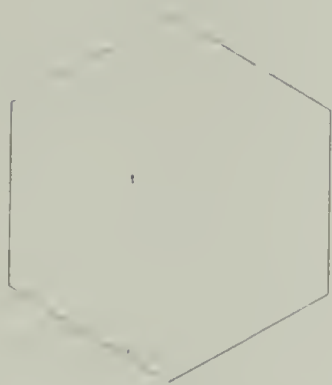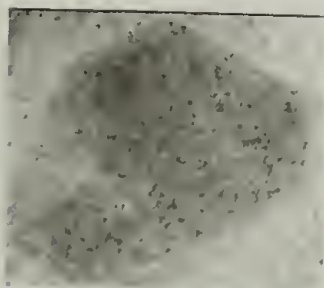
Silk III   Silk III hydrate   Silk I

Figure 54: There is a progression of crystalline habit from hexagonal to triangular when going from silk III to the silk III hydrate structure to silk I (left to right). Silk I is also believed to be a hydrate.

The reason for the change in crystalline habit is not clear, but it is apparent that as crystallites of the hydrated structure grow, three of the nearest neighbor packing directions are no longer equivalent. Since this loss of equivalence is coincident with hydration, a change in the hydrogen bonding pattern is a possible explanation for the change in the symmetry of the crystalline habits. In crystalline silk I prepared using a novel foaming technique lamellae are sometimes observed with a triangular habit shown on the right hand side of Figure 54. A triangular lamella is essentially a hexagonal structure with three major growth directions, resulting in a three-sided habit (a triangular lattice packing is not space filling). Since the silk I structure is also believed to be a hydrate of fibroin, the progression of crystalline habits suggest a progression from the silk III crystal structure to the silk III hydrate structure and the silk I structure. It is also possible that the progressive loss of symmetry in the crystal habit as silk III is compared to its hydrate and then to silk I is due to a progressive change in conformation and packing behavior. In this manner, analysis of the silk III crystal structure and its hydrates may help illuminate the silk I crystal structure and the silk I to silk II transition reported in the literature.

CHAPTER 8

CONCLUSIONS

A novel crystalline structure has been observed for Bombyx mori silk fibroin at the air-water interface and at aqueous-organic liquid-liquid interfaces. This structure, referred to as silk III incorporates a threefold helical conformation of the crystallizable portions of fibroin. The silk III crystal structure appears to be trigonal to a reasonable approximation, with an approximately hexagonal packing of chains. Subtle features of the diffraction data such as additional weak reflections and a deviation from perfect hexagonal symmetry in single crystal diffraction patterns, indicate that the silk III is monoclinic rather than hexagonal and that it is also non-primitive.

A regular arrangement of chains incorporating clustering of hydrophilic serine residues in the (110) planes of the monoclinic unit cell provides a very reasonable fit to the diffraction data. surfactancy of the threefold conformation provides a physical justification for models incorporating clustered serine residues. The crystallizable portions of fibroin incorporate a hexapeptide repeat, $(Gly-Ala-Gly-Ala-Gly-Ser)_n$, which would exhibit mild surfactancy in a threefold conformation due to the placement of hydrophobic alanine and hydrophilic serine on opposite sides of the helix. The presence of serine in the sequence also provides a side chain group that can hydrogen bond, resulting in a roughly threefold conformation that has a regular variation in torsional angles for the six residues in the crystallizable sequence.

188

The first studies of the silk III structure focused on compressed LB films, which have a uniaxially oriented crystalline texture which places the chain axes (c axes) of the silk III crystallites perpendicular to the plane of the sample film. If stabilization of a threefold helical conformation at the air-water interface is due to the surfactant properties of a threefold helix of the fibroin crystallizable sequence, the resulting helices should lie with their chain axes (c axes) in the plane of the sample film. For uncompressed fibroin films prepared at the air-water interface, an oriented crystalline texture is observed where the chain axes of the silk III helices are in the plane of the sample film. This orientation is evidenced in electron diffraction data as changes in the relative intensities of polycrystalline reflections. In the uncompressed films, the (hk0) reflections are attenuated, whereas (hkl) and (00l) reflections are intensified. Different LB trough compression and aging treatments can be used to alter the oriented crystalline textures present and the crystal structure present in the LB films.

The differences in oriented crystalline texture are accompanied by intriguing changes in film morphology observed in TEM images. In the compressed LB films, there are large, extremely thin, facetted crystalline lamellae. In the uncompressed films, these lamellae are not generally observed, and a fibrous film texture is much more evident than in the compressed LB films.

The threefold helical conformation is also observed for fibroin films prepared at interfaces between aqueous fibroin and organic solvents, notably hexane and chloroform.

In these films, a hydrated crystal structure predominates, accompanied by evidence of cholesteric liquid crystallinity. The silk III hydrate structure has a two chain orthorhombic unit cell, and is consistent with hydration of the monoclinic silk III crystal structure observed at the air water interface. Large single lamellae, approximately 500 nm across with threefold symmetry are observed having the silk III hydrate crystal structure. In addition regions with a banded morphology and a chain orientation typical of a cholesteric liquid crystal are also observed. In areas where the banded morphology has fully crystallized, diffraction evidence reveals a crystal structure identical to the silk III hydrate crystalline lamellae. Different regions of the banded mesophase were observed to possess different degrees of crystallinity using diffraction evidence. This strongly suggests that the mesophasic orientation and banded morphology formed prior to crystallization, which then occurred to varying degrees *in situ*. A better understanding of the interaction of mesophase formation and crystallization in fibroin and the subsequent influences of these two organization processes on morphology and crystal structure is needed. Such insight may help elucidate the origin of other crystalline biological structures that appear to incorporate a liquid crystalline type of orientation and organization with small crystallites or fibrils acting to define the director field.

The control offered by interfacial preparation methods over film morphology, oriented crystalline texture and crystal structure should lead to some useful applications. One potential application cited for interfacial fibroin films is as templates for epitaxial crystallization. In this vein, it is interesting to note that in compressed LB films made

with solutions containing both fibroin and sericin, evidence from electron diffraction and TEM images suggests epitaxial crystallization of the two proteins. In addition to observations of the behavior of fibroin and sericin, high quality single crystal diffraction data were obtained for sericin, enabling a better elucidation of its crystalline unit cell and some general packing features. Structural elucidation of sericin will be useful in further studies of its cocrystallization with fibroin, and to better understand the role played by sericin in the silk spinning process.

CHAPTER 9

SUGGESTED FUTURE WORK

9.1 Microfibrils and Mesophase Formation

Under different conditions, the morphology observed for films of fibroin obtained

using an air-water or and organic liquid-water interface varies. In the uncompressed

films from the air-water interface a fibrous morphology is observed, featuring fibrils with

a diameter of approximately 20 nm. At the hexane-water interface a banded morphology

is observed which has a repeat of approximately 1 μm and possesses characteristics of a

cholesteric liquid crystalline domain. Fibrils are not observed within the banded region in

films from the hexane-water interface. Since fibroin exhibits both fibril formation and

mesophasic ordering behavior, it would be a good candidate for a study of the

interaction between these two processes and the relationship between two possible levels

of supermolecular hierarchical order.

Other biological molecules involved in structural extracellular tissues have been

observed *in vivo* in three dimensional non-periodic structures that are liquid crystalline

analogues. In contrast to liquid crystals the biological analogues are solid structures in

which a director field, similar to that of a liquid crystal, is apparent. Collagen, a

structural protein, is among these biomolecules. These molecules form larger crystalline fibrils which then define the director field of the biological liquid crystalline analogue. It is currently not certain whether crystallization into fibrils occurs concurrently with mesophasic texture formation. It is also possible that the fibrils form before the liquid crystalline analogue morphology and are subsequently oriented, or that they are formed by crystallization in situ after the orientation field is established. In the case of fibroin the data indicate varying degrees of crystallinity for the cholesteric-like banded morphology with more crystallinity developed in the thin areas near the edges of the region. The fact that regions of high crystallinity coexist with regions of merely cholesteric order, without observable differences in the banded morphology, strongly suggests that the silk fibroin molecules crystallize *in situ* after the mesophasic texture is formed. Whether or not this is a general characteristic of fibrous biomolecules remains to be seen, but such a mechanism would clarify formation of liquid crystal-like textures in biological systems where the mesogen aspect ratio for fibrils is outside the limiting value of 10 calculated by Flory.

The occurrence of fibrils and of a banded mesophase can be reconciled using the hypothesis that the mesophase forms first and mesophasic orientation is a fast process, and that fibril formation is slower. In addition, the data suggest that different interactions are involved in the two processes. There are a number of features in the experimental data that, when taken together, suggest this hypothesis. These features point to further experiments that could be used to either modify (or disprove) the hypothesis or put it on firmer footing.

In the uncompressed films, fibrils are observed protruding from the edges in the films, protruding from cracks and dark fibril-shaped objects within the films. Within the continuous areas of film the fibrils appear to often be arranged in grains of locally oriented fibrils. In these grains, there are dark parallel 20 nm diameter lines which can be interpreted as fibrils. They have the same size and shape as the fibrils protruding from the ends of the film, and in some instances a single fibril can be traced from the interior of the film to the edge where its end protrudes. Within the grains of oriented fibrils there appear to be not one but three orientations preferred by the fibrils. These directions are at an apparent angle of 120° to each other. In some regions of the uncompressed film there a larger scale banded structure is apparent which looks similar to the banded mesophase observed at the hexane-water interface. This structure was shown in Chapter 4, Figure 26. In the uncompressed film from the air-water interface, the edge of the banded structure has fibrils protruding from it as seen in Figure 55. In the two uncompressed film images shown in Figure 55, there are dark fibrillar objects that can be seen in the interior of the film. These objects protrude from the edges of cracks, which also have a fibrous appearance. Edges and cracks in the structure from the hexane-water interface do not. Another contrast between the micron sized bands in the uncompressed films from the air-water interface and the banded morphology in films from the hexane-water interface is apparent in diffraction from these structures. At the hexane-water interface, diffraction indicates an orientation consistent with a cholesteric liquid crystal or closely analogous structure. In some regions the banded structure is not crystalline and only one or sometimes two arced diffraction rings are observed,

consistent with a non-crystalline mesophase. In other regions, crystallinity has occurred, resulting in uniaxially oriented crystalline regions with many reflections. This is not observed in the fibrillar films obtained at the uncompressed air-water interface, which have a consistent biaxially oriented crystalline texture and a smaller number of observable reflections.

If a mesophase orientation is responsible for the micron scale banding in both cases, air-water and hexane-water interfacial films, then there are a few possibilities for subsequent organization of the films. One is mesophase formation and solidification of the director texture (with some localized crystallization) before fibril formation can occur. This situation would be expected in instances where the interactions leading to fibril formation are screened or where the orientation and crystallization rate are enhanced through some agency such as external stress. At the water-hexane interface, silk acting as a surfactant can compatibilize the two liquids, hexane and water, by surrounding small clusters of hexane molecules with hydrophobic silk regions. Such behavior is observed for surfactant layers of lecithin at water-organic liquid interfaces.[89,90,105] The presence of hexane rich areas in the interfacial region would be expected to shield hydrophobic van der Waals interactions between silk molecules. Since fibril formation in fibrous proteins such as collagen is believed to be due to hydrophobic interactions, it is not unreasonable to expect analogous behavior in fibroin.[3,154-160] Shielding of the interactions between fibroin molecules would prevent fibrillar aggregation and allow the cholesteric analogue texture to develop more fully.

Since fibril formation results in large objects that are difficult to align, absence or damping of the competing fibril formation process would result in a better developed cholesteric morphology, perhaps covering larger regions, which could solidify without interference from fibril formation. In the water-hexane interfacial films, there is no evidence for fibrils within the banded morphology, and there are no fibrils protruding from cracks within the banded morphology either. In regions of the film that are not banded fibrils are sometimes observed protruding from the edges of the film, but they are less numerous than the fibrils in the uncompressed air-water interfacial films and they are smaller, with a typical diameter of 5-10 nm.

A micron-scale banded morphology has also been observed in air-blown foams of aqueous fibroin, although the banding is not as clear and well developed.[161] In crystalline regions of banded morphology, the silk III hydrated crystal structure is observed in the same orientation relative to the bands as at the hexane-water interface. In this instance, external stress would be expected to align the fibroin molecules, speeding solidification and crystallization of mesophasic regions before fibril formation could occur. A test of this explanation might have implications for understanding the natural silk spinning process and the role of liquid crystallinity in silk spinning.
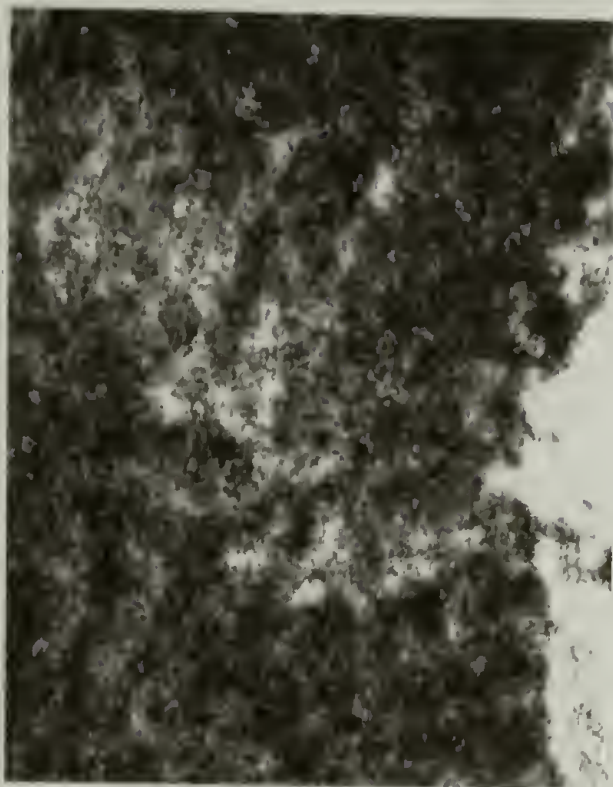
Figure 55: Morphology images from an uncompressed fibroin film prepared at the airwater interface. Dark objects which appear to be fibrils are apparent in the interior of the films as well as protruding from edges.

The following hypotheses need to be tested:

1. Cholesteric organization of fibroin is faster than fibril formation.

2. The two processes compete.

3. The two processes depend on different types of interactions, for example shape versus van der Waals forces.

4. The two types of interactions can be separately damped by changing the physical and chemical environment of the fibroin molecules.

Several experiments can be used to test these hypothesis. The first point should be addressable by varying the air-flow rate used in foaming identical solutions of fibroin (same batch and age). X-ray data can be used to characterize the relative proportions of the crystal structures present, and electron microscopy studies will reveal any changes in the amount and structure of banded or fibrillar morphologies present. If we assume that cholesteric organization is the faster process, then slower air-flow rates should result in less crystallinity in the banded morphology regions and in some fibril formation. Higher rates should produce more banding. Addition of small amounts of methanol or other agents known to produce the silk II structure would be expected to interfere with mesophase formation by changing the shape and packing behavior of the molecules. However since the β–sheet structure is stabilized by intermolecular hydrogen bonds and is difficult to observe except in a crystalline state, interpreting the influence of this conformation in terms of a preference for fibril formation or mesophase organization may be difficult. Similarly since a hydrophobic solvent interface results in more banding, comparison of the morphologies observed at a larger variety of liquid organic

solvent-water interfaces may be useful in determining whether hydrophobic organic liquids do in fact screen the interactions that result in fibril formation. Studies of aqueous silk fibroin solution gelation may also be useful to elucidate the processes involved in supermolecular aggregates and organized structures formed by fibroin.

9.2 Other Water- Organic Interfaces

In addition to the experiments using the liquid - liquid aqueous-hexane interface to prepare fibroin films, other aqueous organic interfaces were examined, and preliminary characterization results were obtained for the resulting fibroin films. The data presented in Chapter 5 on the banded mesophase are from an aqueous hexane interface made with a 5% aqueous fibroin solution. Some data from chloroform interfacial experiments with 1% aqueous fibroin solutions were also reported. Preliminary data were also obtained for interfacial films prepared using other organic solvents and different concentrations of fibroin. A series of interfacial experiments was performed using hexane, dodecane, and chloroform to form interfaces with 1% and 0.5% aqueous fibroin solutions.
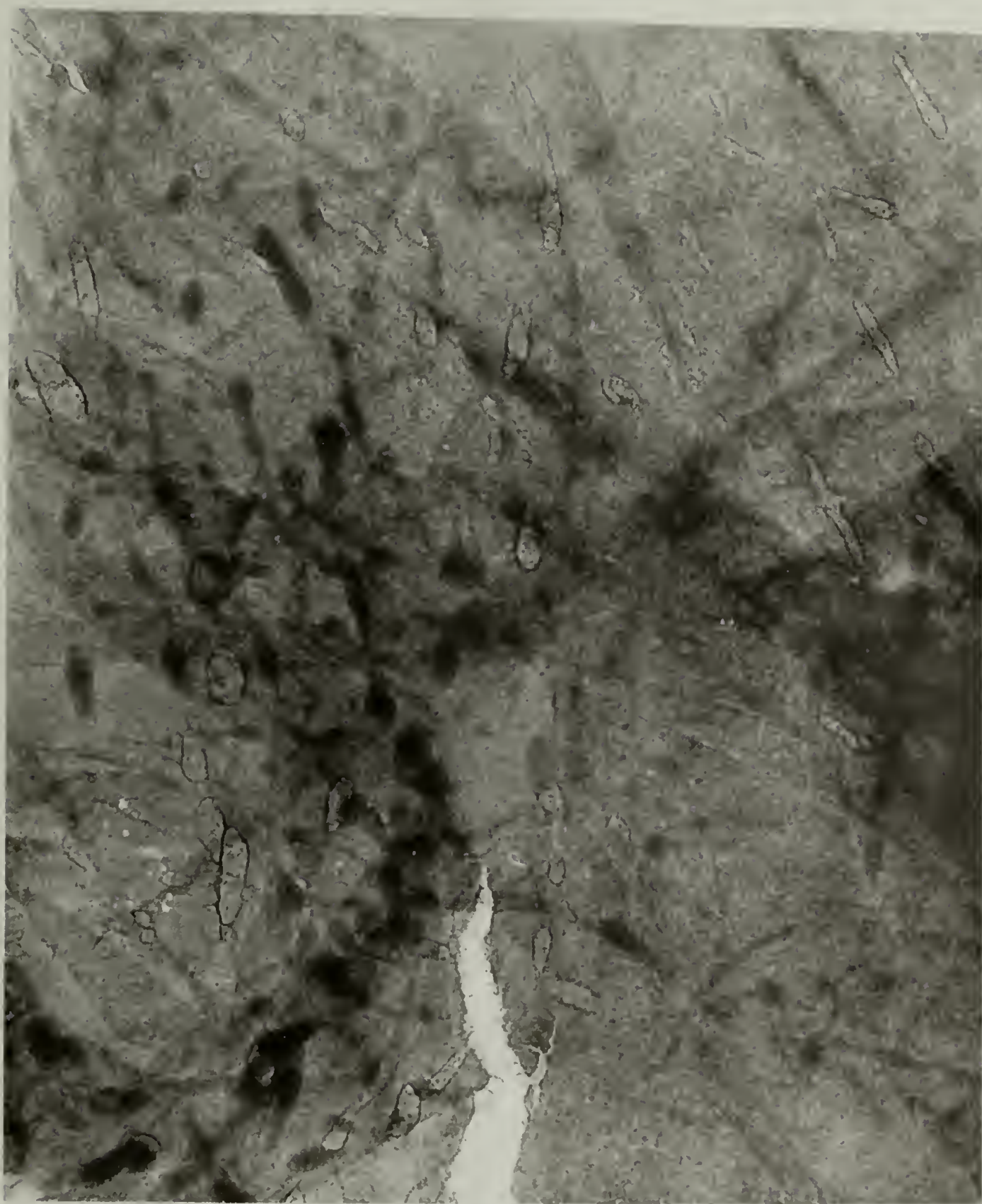
Figure 56: Morphology image from a fibroin film at the aqueous-chloroform, liquid-liquid interface.

Figure 57: Diffraction pattern from a fibroin film prepared at the aqueous-chloroform liquid-liquid interface.

The dodecane interfacial films had diffuse electron diffraction patterns with one or sometimes two rings. The corresponding film morphologies appeared featureless. The hexane and chloroform interfacial films were more intriguing, containing large regions where the films possessed a mesh-like pattern., such as the one shown in Figure 56. The dark oblong objects in the morphology image do not appear to be arranged at random. Instead they appear to locally prefer a small number of directions. The diffraction pattern in Figure 57 contains very few reflections, but is arced, indicating orientation within the film region. Further characterization of these and other aqueous - organic interfacial films would elucidate the structure and orientation of these films and perhaps point to methods for controlling morphology using interfacial chemistry.

In addition to the banded morphology and lamellar hydrate crystallites observed at the 5% aqueous fibroin - hexane interface there are also regions that possess a spongy appearance as shown in Figure 58. Within these regions are numerous facetted, apparently hexagonal objects which are probably crystalline lamellae. Diffraction from a spongy region, shown in Figure 59, revels a set of sharp arcs indicating a polycrystalline texture with a high degree of orientation. The d-spacings appear to be in good agreement with the monoclinic silk III structure, but further data are needed to make a definite structural determination. A more detailed survey of aqueous organic interfaces using higher concentrations of aqueous fibroin, roughly 3-6 %, might reveal more about the structure and orientation of the spongy morphology. The hexagonal habit of the lamellae and *d*-spacings observed in the diffraction pattern strongly suggest the

monoclinic silk III structure is present. Since the diffraction pattern obtained from the spongy region is much sharper than the monoclinic silk III diffraction obtained from LB films and cast films, additional insight into the detailed structure of the silk III polymorph may be possible. Further study of the structures in films prepared using aqueous organic interfaces with more concentrated solutions of fibroin may also help determine the variables which influence the hydration behavior of fibroin in the films. Comparisons of morphologies formed by silk III hydrated structures and dehydrated (or less hydrated) structures under similar conditions could also be made.
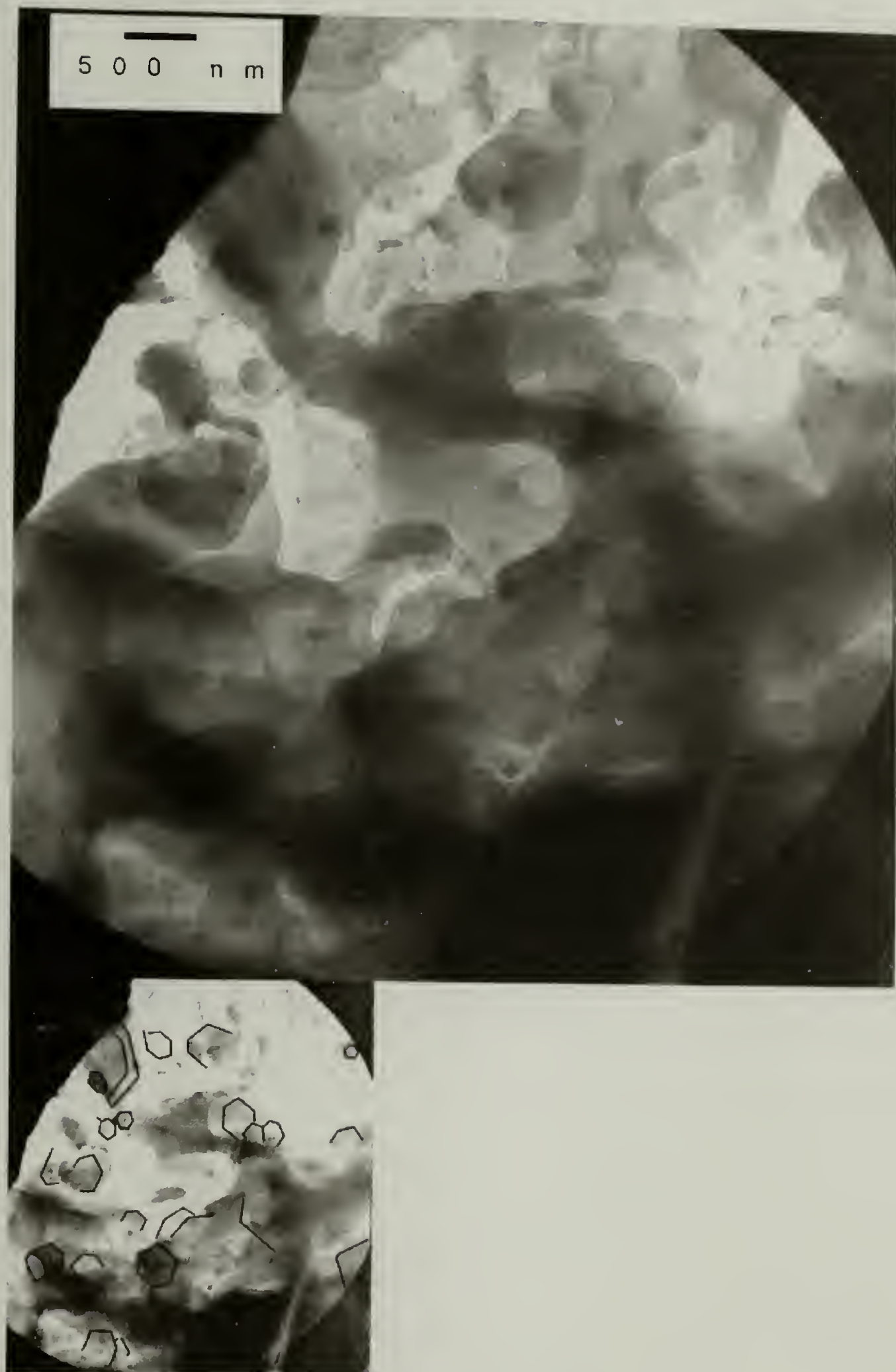
Figure 58: Spongy morphology observed at the 5% aqueous fibroin-hexane interface. The inset shows the location of facetted apparently hexagonal objects which are probably the lamellae responsible for the diffraction observed.

Figure 59: Diffraction pattern from a spongy region of a fibroin film prepared at the 5% aqueous fibroin-hexane interface.

APPENDIX A

ATOMIC COORDINATES FOR THE SILK III CRYSTAL STRUCTURES

206

Atomic coordinates are provided from representative superlattice structures for the trigonal crystalline unit cell model in Chapter 2, the monoclinic model in Chapter 3, and the hydrated model in Chapter 5. Atomic locations are given in fractional coordinates. Complete structural files, including bonding information and calculated atomic partial charges are not included here because of the large variety of possible mutually incompatible file formats used to represent such information. However these files can be obtained in either .cssr (Molecular Simulations - Cerius) or .pdb (Brookhaven Protein Databank) format, and will be made available upon request.

A.1 Trigonal model (superlattice)

a = 9.64; b = 9.64; c = 18.84 (2 x 2 x 2 supercell)

α = 90; β = 90; γ = 120;   Spacegroup = 1P1

208 atoms;   Created by CERIUS

1	C1	-0.14054	0.62321	0.32953
2	H2	-0.26607	0.5781	0.34264
3	C3	-0.26744	0.35874	0.16429
4	C4	-0.3928	0.28895	0.14854
5	H5	-0.45694	0.21715	0.1885
6	H6	-0.43398	0.36299	0.1367
7	H7	-0.40253	0.22268	0.10598
8	H8	-0.22258	0.27803	0.17718
9	C9	0	0.5	0
10	H10	0.08047	0.62529	0.01349
11	O11	-0.00284	0.51842	0.26015
12	O12	-0.16155	0.60478	0.09594
13	O13	-0.24495	0.36339	-0.06927
14	C14	-0.12739	0.52465	0.26817
15	C15	-0.16779	0.47181	0.1035
16	C16	-0.11197	0.48795	-0.06104
17	N17	-0.25343	0.45419	0.22559
18	H18	-0.33745	0.47171	0.23782
19	N19	-0.09656	0.41718	0.06085
20	H20	-0.1143	0.31539	0.07269

21	N21	-0.05886	0.60802	0.39684
22	H22	0.04199	0.69723	0.40535
23	C23	-0.13826	0.62995	-0.17046
24	C24	-0.08275	0.7556	-0.18572
25	H25	-0.09026	0.819	-0.14541
26	H26	0.03241	0.79698	-0.19775
27	H27	-0.13948	0.76608	-0.22804
28	H28	-0.26379	0.58484	-0.15735
29	C29	-0.26515	0.36548	-0.3357
30	H30	-0.22029	0.28477	-0.32281
31	C31	-0.00229	0.49326	0.5
32	C32	0.06779	0.43803	0.48418
33	H33	0.13887	0.44481	0.52429
34	H34	0.13483	0.49553	0.44194
35	H35	0.07819	0.61855	0.51348
36	O36	-0.00056	0.52516	-0.23985
37	O37	-0.15927	0.61152	-0.40405
38	O38	-0.24723	0.35665	0.43072
39	C39	-0.12511	0.53139	-0.23182
40	C40	-0.1655	0.47855	-0.39649
41	C41	-0.11425	0.48121	0.43896
42	N42	-0.25115	0.46093	-0.2744
43	H43	-0.33517	0.47845	-0.26217
44	N44	-0.09884	0.41044	-0.43916
45	H45	-0.11658	0.30865	-0.42732
46	N46	-0.05657	0.61476	-0.10315

47	H47	0.04427	0.70397	-0.09464
48	C48	-0.13013	0.1366	-0.66455
49	C49	-0.18677	0.20548	-0.64894
50	H50	-0.18029	0.2764	-0.68907
51	H51	-0.30168	0.13064	-0.63687
52	H52	-0.13033	0.27262	-0.60657
53	H53	-0.00489	0.21818	-0.67771
54	C54	0.00001	0	-0.50001
55	H55	-0.04365	-0.12524	-0.51316
56	C56	-0.26673	-0.13014	-0.33547
57	C57	-0.33562	-0.25567	-0.31985
58	H58	-0.40653	-0.32008	-0.35998
59	H59	-0.26077	-0.29572	-0.30777
60	H60	-0.40275	-0.26634	-0.27748
61	H61	-0.34832	-0.08648	-0.34862
62	O62	-0.26536	-0.1067	-0.59554
63	O63	-0.10794	0.13736	-0.43099
64	O64	-0.0213	-0.02006	-0.26645
65	C65	-0.1417	0.02489	-0.60349
66	C66	-0.10014	0.01156	-0.43894
67	C67	-0.15502	-0.02999	-0.2744
68	N68	-0.01537	0.08131	-0.56103
69	H69	0.06785	0.18334	-0.57317
70	N70	-0.17004	-0.11476	-0.39649
71	H71	-0.15122	-0.19796	-0.40862
72	N72	-0.21144	0.03991	-0.73196

73	H73	-0.31279	0.02718	-0.74024
74	C74	-0.13013	0.13659	-0.16455
75	H75	-0.00489	0.21818	-0.17771
76	C76	0	0	0
77	C77	0.12553	0.05663	0.01561
78	H78	0.18996	0.05015	-0.02452
79	H79	0.16559	0.17154	0.02768
80	H80	-0.04366	-0.12524	-0.01317
81	C81	-0.26673	-0.13014	-0.83546
82	H82	-0.34831	-0.08648	-0.84863
83	O83	-0.10794	0.13736	0.06901
84	O84	-0.02129	-0.02006	-0.76646
85	C85	-0.14629	0.03767	-0.10204
86	C86	-0.10014	0.01156	0.06106
87	C87	-0.15502	-0.02999	-0.7744
88	N88	-0.01537	0.08131	-0.06103
89	H89	0.06784	0.18334	-0.07316
90	N90	-0.17004	-0.11476	-0.8965
91	H91	-0.15122	-0.19796	-0.90862
92	N92	-0.21144	0.03991	-0.23195
93	H93	-0.3128	0.02717	-0.24024
94	C94	0.36278	0.1295	0.32908
95	C95	0.41911	0.255	0.31347
96	H96	0.41241	0.3193	0.3536
97	H97	0.53407	0.29534	0.3014
98	H98	0.23747	0.0855	0.34225

99	C99	0.23327	-0.13724	0.16452
100	H100	0.27726	-0.21856	0.17769
101	C101	0.5	0	0
102	C102	0.56917	-0.05633	-0.01561
103	H103	0.64018	-0.04963	0.02451
104	H104	0.49458	-0.17129	-0.0277
105	H105	0.63624	0.00024	-0.05799
106	H106	0.58131	0.1253	0.01315
107	O107	0.49856	0.02203	0.26007
108	O108	0.34061	0.10802	0.09553
109	O109	0.25461	-0.13592	-0.06901
110	C110	0.37467	0.02954	0.26801
111	C111	0.33322	-0.02538	0.10348
112	C112	0.38815	-0.0119	-0.06108
113	N113	0.24837	-0.04065	0.22557
114	H114	0.16496	-0.02213	0.2377
115	N115	0.40342	-0.08149	0.06101
116	H116	0.38488	-0.18342	0.07314
117	N117	0.44426	0.1144	0.39648
118	H118	0.5455	0.2032	0.40478
119	C119	0.36277	0.1295	-0.17092
120	H120	0.23747	0.08551	-0.15775
121	C121	0.23731	-0.12545	0.66452
122	C122	0.11181	-0.19463	0.64892
123	H123	0.0475	-0.26563	0.68904
124	H124	0.07145	-0.12003	0.63684

125	H125	0.10133	-0.26168	0.60654
126	H126	0.28129	-0.20677	0.67768
127	C127	0.5	-0.00001	0.5
128	H128	0.58132	0.1253	0.51316
129	O129	0.49856	0.02203	-0.23992
130	O130	0.34465	0.11981	0.59552
131	O131	0.25461	-0.13593	0.43098
132	C132	0.37467	0.02955	-0.23198
133	C133	0.33726	-0.0136	0.60346
134	C134	0.38816	-0.01191	0.43894
135	N135	0.24837	-0.04065	-0.27443
136	H136	0.16496	-0.02212	-0.2623
137	N137	0.40343	-0.08149	0.56103
138	H138	0.38489	-0.18342	0.57316
139	N139	0.44426	0.1144	-0.10353
140	H140	0.5455	0.2032	-0.09522
141	C141	0.37414	0.64076	0.83546
142	H142	0.49976	0.72055	0.82229
143	C143	0.5	0.5	0
144	C144	0.62538	0.55458	0.01561
145	H145	0.68898	0.54667	-0.02452
146	H146	0.66743	0.66983	0.02769
147	H147	0.63475	0.49733	0.05798
148	H148	0.45417	0.37438	-0.01316
149	C149	0.23338	0.37413	0.16454
150	H150	0.15359	0.41996	0.15138

151	O151	0.23523	0.39779	0.90446
152	O152	0.39611	0.64107	0.06901
153	O153	0.47851	0.48018	0.23355
154	C154	0.36034	0.52814	0.89652
155	C155	0.40118	0.5138	0.06106
156	C156	0.346	0.47297	0.22561
157	N157	0.4865	0.58248	0.93897
158	H158	0.57098	0.68381	0.92684
159	N159	0.32936	0.38763	0.10351
160	H160	0.34623	0.30317	0.09138
161	N161	0.29166	0.54478	0.76806
162	H162	0.19112	0.53405	0.75976
163	C163	0.37414	0.64076	0.33546
164	C164	0.31955	0.71155	0.35106
165	H165	0.2043	0.63835	0.36314
166	H166	0.37681	0.77819	0.39344
167	H167	0.49976	0.72055	0.32229
168	C168	0.5	0.5	0.5
169	H169	0.45417	0.37438	0.48684
170	C170	0.23338	0.37413	0.66454
171	C171	0.16259	0.24876	0.68015
172	H172	0.09107	0.18517	0.64002
173	H173	0.23579	0.20672	0.69223
174	H174	0.09595	0.23939	0.72252
175	H175	0.15359	0.41996	0.65138
176	O176	0.23523	0.39779	0.40446

177	O177	0.39611	0.64107	0.56901
178	O178	0.47851	0.48018	0.73355
179	C179	0.36034	0.52814	0.39652
180	C180	0.40118	0.5138	0.56106
181	C181	0.346	0.47297	0.7256
182	N182	0.4865	0.58248	0.43897
183	H183	0.57098	0.68381	0.42684
184	N184	0.32936	0.38763	0.60351
185	H185	0.34623	0.30317	0.59138
186	N186	0.29166	0.54478	0.26806
187	H187	0.19112	0.53405	0.25976
188	H188	-0.09203	0.73258	0.31598
189	H189	0.0608	0.45174	-0.01402
190	H190	-0.37428	0.30493	-0.34967
191	H191	0.10926	0.0495	-0.48615
192	H192	-0.17963	0.19635	-0.15071
193	H193	-0.3265	-0.23939	-0.82161
194	O194	-0.27775	-0.08533	-0.08891
195	H195	0.12404	-0.19725	0.15067
196	H196	0.412	0.23874	-0.18477
197	H197	0.56001	-0.04923	0.48614
198	H198	0.32643	0.70218	0.84931
199	H199	0.17195	0.265	0.17839
200	H200	0.60913	0.54771	0.51385
201	O201	0.13941	-0.01674	0.0707
202	H202	0.25312	0.02502	0.08102

203	O203	0.34558	0.26863	0.25839
204	H204	0.34848	0.37164	0.26023
205	O205	-0.02819	0.28861	0.4681
206	H206	-0.06873	0.22561	0.51174
207	O207	0.32984	0.80453	0.2989
208	H208	0.44286	0.87892	0.28822

A.2 Monoclinic model

a = 20.4; b = 20.4; c = 19.683

α = 90; β = 90; γ = 116.546; Spacegroup 1P1

832 atoms;   Created   by CERIUS

1	C1	0.4369	0.50349	0.04872
2	H2	0.42466	0.44558	0.04832
3	C3	0.32134	0.47525	0.20652
4	C4	0.2484	0.40551	0.21194
5	H5	0.21266	0.40759	0.17184
6	H6	0.25848	0.35788	0.20375
7	H7	0.22127	0.4001	0.26091
8	H8	0.30884	0.52143	0.19891
9	O9	0.37573	0.57051	0.09175
10	O10	0.41811	0.48564	0.28189
11	C11	0.3877	0.51539	0.09599
12	C12	0.35935	0.48957	0.27303
13	N13	0.3612	0.46836	0.1484
14	H14	0.37106	0.42599	0.14735
15	C15	0.35667	0.54001	0.38745
16	H16	0.41447	0.57942	0.38222
17	C17	0.34779	0.48664	0.44096
18	N18	0.39792	0.51735	0.48957

19	H19	0.42989	0.56892	0.48389
20	N20	0.32872	0.51262	0.32236
21	H21	0.28133	0.50971	0.3137
22	H22	0.49211	0.5344	0.06892
23	H23	0.32745	0.56995	0.40612
24	O24	0.29693	0.4225	0.44683
25	C25	0.29316	0.1754	0.20967
26	H26	0.30073	0.12583	0.20636
27	C27	0.41099	0.25544	0.0567
28	C28	0.48258	0.2487	0.04658
29	H29	0.52313	0.28839	0.07994
30	H30	0.47655	0.1939	0.05881
31	H31	0.50424	0.26331	-0.00488
32	H32	0.4269	0.3132	0.06773
33	C33	0.34132	0.28681	0.88699
34	H34	0.28431	0.24537	0.88957
35	O35	0.3502	0.29763	0.16679
36	O36	0.30168	0.17885	-0.00647
37	O37	0.44142	0.28831	0.82792
38	C38	0.34195	0.23272	0.16291
39	C39	0.36124	0.23662	-0.003
40	C40	0.37467	0.27339	0.82586
41	N41	0.37543	0.21188	0.11404
42	H42	0.37237	0.16319	0.11586
43	N43	0.38133	0.28801	0.94804
44	H44	0.42967	0.3291	0.95491

45	N45	0.3007	0.20015	0.27804
46	H46	0.26403	0.21477	0.29276
47	C47	0.36088	0.2067	0.71987
48	C48	0.29451	0.13303	0.70432
49	H49	0.2499	0.14435	0.68665
50	H50	0.30578	0.10264	0.66414
51	H51	0.27745	0.09868	0.7496
52	H52	0.40375	0.19418	0.74065
53	C53	0.45053	0.24304	0.54888
54	H54	0.47056	0.30245	0.54375
55	C55	0.35636	0.24439	0.3887
56	C56	0.36553	0.3247	0.38323
57	H57	0.34508	0.33779	0.42966
58	H58	0.32939	0.32594	0.34263
59	H59	0.30132	0.21037	0.40874
60	O60	0.39286	0.30841	0.64315
61	O61	0.33603	0.14979	0.5062
62	O62	0.41351	0.20969	0.29918
63	C63	0.39365	0.24871	0.65762
64	C64	0.396	0.20501	0.49484
65	C65	0.35793	0.2145	0.3203
66	N66	0.42538	0.21932	0.61602
67	H67	0.42377	0.17326	0.63023
68	N68	0.40934	0.2427	0.4358
69	H69	0.45991	0.27552	0.42601
70	N70	0.33572	0.24249	0.76903

71	H71	0.28739	0.23951	0.76266
72	C72	0.39558	0.53733	0.87158
73	C73	0.38934	0.6095	0.88395
74	H74	0.33421	0.6032	0.87523
75	H75	0.4048	0.63047	0.9359
76	H76	0.42819	0.65113	0.85048
77	H77	0.45208	0.5521	0.85653
78	O78	0.32514	0.43518	0.94468
79	C79	0.3817	0.49471	0.93485
80	N80	0.43571	0.5281	0.98098
81	H81	0.47361	0.5761	0.96845
82	N82	0.34863	0.49602	0.81805
83	H83	0.29725	0.48459	0.82276
84	C84	0.41078	0.48436	0.54777
85	C85	0.493	0.52159	0.56218
86	H86	0.52345	0.52942	0.51505
87	H87	0.50894	0.57496	0.5863
88	H88	0.39512	0.42644	0.53901
89	C89	0.311	0.43264	0.71237
90	H90	0.27924	0.37641	0.72977
91	O91	0.37671	0.55404	0.6246
92	O92	0.4357	0.48647	0.75235
93	C93	0.37343	0.49307	0.6083
94	C94	0.36941	0.47291	0.76133
95	N95	0.34146	0.43348	0.64777
96	H96	0.35182	0.3928	0.63819

97	C97	0.1098	0.53663	0.39705
98	C98	0.11245	0.61206	0.38899
99	H99	0.10772	0.6307	0.43928
100	H100	0.16555	0.65126	0.36957
101	H101	0.05241	0.49894	0.40496
102	C102	0.08064	0.46143	0.22046
103	H103	0.11005	0.42832	0.22479
104	C104	0.14541	0.51151	0.04002
105	C105	0.2093	0.49114	0.03199
106	H106	0.24906	0.51844	0.0718
107	H107	0.19116	0.43251	0.03735
108	H108	0.23765	0.50989	-0.0167
109	H109	0.17069	0.57107	0.04595
110	O110	0.18672	0.49994	0.33066
111	O111	0.13013	0.5772	0.16225
112	O112	0.05375	0.42793	0.96511
113	C113	0.13166	0.51175	0.3343
114	C114	0.10774	0.50884	0.16043
115	C115	0.09788	0.49347	0.97831
116	N116	0.08571	0.49967	0.28228
117	H117	0.04937	0.51717	0.29122
118	N118	0.10974	0.47544	0.10196
119	H119	0.08729	0.42249	0.10302
120	N120	0.15018	0.53794	0.45665
121	H121	0.18213	0.58461	0.47729
122	C122	0.07568	0.53803	0.86809

123	H123	0.01946	0.49615	0.865
124	C124	0.12514	0.47313	0.70515
125	C125	0.05688	0.39824	0.69509
126	H126	0.05154	0.37761	0.64308
127	H127	0.05826	0.35706	0.72939
128	H128	0.00626	0.40224	0.70534
129	H129	0.16695	0.46079	0.72847
130	C130	0.2048	0.48285	0.53227
131	H131	0.25387	0.53504	0.52694
132	O132	0.16919	0.5048	0.84169
133	O133	0.1671	0.57019	0.61767
134	O134	0.08617	0.41885	0.48068
135	C135	0.11823	0.51745	0.82014
136	C136	0.15919	0.50945	0.63973
137	C137	0.14492	0.47705	0.48693
138	N138	0.10047	0.5102	0.75325
139	H139	0.05968	0.52	0.74042
140	N140	0.18418	0.47081	0.60124
141	H141	0.17813	0.42473	0.62024
142	N142	0.09882	0.54775	0.9377
143	H143	0.11686	0.59703	0.9553
144	C144	0.16231	0.2118	0.04595
145	H145	0.11801	0.15625	0.04292
146	C146	0.09062	0.22913	0.21584
147	C147	0.02182	0.15545	0.22278
148	H148	-0.01243	0.14581	0.17829

149	H149	0.0365	0.11061	0.22852
150	H150	-0.01119	0.15421	0.26638
151	H151	0.07272	0.26977	0.1981
152	C152	0.11414	0.31543	0.38576
153	H153	0.16966	0.36069	0.38445
154	O154	0.14699	0.30775	0.09744
155	O155	0.18964	0.25896	0.29709
156	O156	0.04581	0.20689	0.44809
157	C157	0.1466	0.24647	0.10468
158	C158	0.1296	0.25775	0.28026
159	C159	0.10527	0.26343	0.44029
160	N160	0.13032	0.21274	0.16541
161	H161	0.13636	0.16861	0.17244
162	N162	0.09335	0.28267	0.32116
163	H163	0.04429	0.27139	0.30705
164	N164	0.16929	0.25181	0.98403
165	H165	0.21164	0.2998	0.98173
166	C166	0.16677	0.23867	0.53991
167	C167	0.24498	0.27675	0.56941
168	H168	0.2572	0.33405	0.57565
169	H169	0.24745	0.25381	0.61931
170	H170	0.15413	0.18219	0.52778
171	C171	0.06651	0.18083	0.70638
172	H172	0.02777	0.20422	0.70295
173	C173	0.14288	0.27955	0.87096
174	C174	0.13409	0.35004	0.88109

175	H175	0.15243	0.3735	0.93143
176	H176	0.1677	0.3903	0.84336
177	H177	0.07749	0.3405	0.8752
178	H178	0.20078	0.29532	0.86125
179	O179	0.11925	0.29742	0.61333
180	O180	0.19091	0.23961	0.74681
181	O181	0.05975	0.17566	0.9351
182	C182	0.12016	0.23801	0.59774
183	C183	0.12391	0.22222	0.75631
184	C184	0.12049	0.23218	0.93237
185	N185	0.09661	0.18046	0.6412
186	H186	0.10813	0.14024	0.63284
187	N187	0.10061	0.24343	0.81277
188	H188	0.04893	0.23259	0.81518
189	N189	0.16375	0.27947	0.48075
190	H190	0.2074	0.32452	0.47072
191	H191	0.23747	0.161	0.19338
192	H192	0.34171	0.33969	0.88
193	H193	0.49667	0.23255	0.54165
194	H194	0.27506	0.45861	0.70966
195	H195	0.02292	0.42383	0.2124
196	H196	0.07909	0.58939	0.84949
197	H197	0.21765	0.43899	0.51762
198	H198	0.21363	0.20948	0.05646
199	H199	0.0753	0.33622	0.39743
200	H200	0.03686	0.12463	0.72464

201	O201	0.50612	0.47343	0.60596
202	H202	0.55522	0.50376	0.6241
203	O203	0.05184	0.60786	0.34941
204	H204	0.04014	0.64302	0.37487
205	O205	0.43889	0.3835	0.37407
206	H206	0.46177	0.39313	0.4197
207	O207	0.29981	0.27153	0.52855
208	H208	0.34011	0.28547	0.56137
209	C209	0.42583	0.97866	0.0567
210	H210	0.40186	0.91946	0.05084
211	C211	0.31139	0.96262	0.21146
212	C212	0.24632	0.88852	0.22835
213	H213	0.20081	0.88121	0.19618
214	H214	0.26011	0.84362	0.21982
215	H215	0.22858	0.886	0.28069
216	H216	0.29029	1.00073	0.19497
217	O217	0.38552	1.05927	0.10787
218	O218	0.42313	1.01115	0.27517
219	C219	0.38504	0.99737	0.10786
220	C220	0.35691	0.99652	0.27159
221	N221	0.35025	0.94906	0.15774
222	H222	0.35439	0.90348	0.16065
223	C223	0.34806	1.04429	0.38688
224	H224	0.40658	1.08248	0.38392
225	C225	0.33339	0.99194	0.44284
226	N226	0.38575	1.01588	0.48991

227	H227	0.42088	1.06728	0.48591
228	N228	0.3214	1.00944	0.32361
229	H229	0.26973	0.99381	0.31874
230	H230	0.48153	0.99753	0.07594
231	H231	0.31849	1.07582	0.39937
232	O232	0.28032	0.92856	0.44544
233	C233	0.31162	0.6748	0.23659
234	H234	0.34136	0.64203	0.24057
235	C235	0.41589	0.76682	0.07512
236	C236	0.48508	0.75616	0.06317
237	H237	0.52707	0.79163	0.09805
238	H238	0.47473	0.69896	0.06902
239	H239	0.50794	0.77283	0.01241
240	H240	0.43253	0.82255	0.09397
241	C241	0.38621	0.82437	0.89965
242	H242	0.3274	0.80712	0.90005
243	O243	0.35212	0.7943	0.19017
244	O244	0.31797	0.70862	-0.00333
245	O245	0.46111	0.77709	0.84796
246	C246	0.34681	0.73094	0.1832
247	C247	0.37777	0.7626	0.01022
248	C248	0.40222	0.78145	0.8482
249	N249	0.37444	0.71436	0.12621
250	H250	0.36753	0.66446	0.11908
251	N251	0.41137	0.81853	0.96612
252	H252	0.4604	0.85593	0.97647

253	N253	0.30711	0.70511	0.3017
254	H254	0.27015	0.72289	0.30408
255	C255	0.35323	0.69977	0.74661
256	C256	0.27611	0.63487	0.75323
257	H257	0.25791	0.62667	0.80615
258	H258	0.23661	0.64575	0.72401
259	H259	0.27607	0.58456	0.73484
260	H260	0.39321	0.67992	0.75739
261	C261	0.43958	0.73524	0.56865
262	H262	0.47394	0.79377	0.5728
263	C263	0.35782	0.7692	0.40817
264	C264	0.37439	0.8445	0.37686
265	H265	0.38938	0.88624	0.41634
266	H266	0.32524	0.84091	0.35105
267	H267	0.30205	0.74638	0.42736
268	O268	0.36505	0.78703	0.66393
269	O269	0.33314	0.66445	0.50435
270	O270	0.40895	0.70339	0.34758
271	C271	0.37334	0.73144	0.67745
272	C272	0.39041	0.72283	0.51097
273	C273	0.35653	0.71945	0.35283
274	N274	0.40573	0.70478	0.63263
275	H275	0.40159	0.65704	0.6422
276	N276	0.40699	0.77616	0.46423
277	H277	0.4531	0.8202	0.46939
278	N278	0.35382	0.75039	0.79784

279	H279	0.31176	0.76034	0.79866
280	C280	0.39181	1.03761	0.88652
281	C281	0.3912	1.11209	0.89649
282	H282	0.42134	1.14989	0.85556
283	H283	0.33508	1.10495	0.89684
284	H284	0.4165	1.13624	0.94477
285	H285	0.44836	1.05107	0.87531
286	O286	0.31433	0.92903	0.94887
287	C287	0.37218	0.98893	0.94678
288	N288	0.42645	1.01292	0.99352
289	H289	0.46851	1.05902	0.98329
290	N290	0.34796	1.00367	0.82921
291	H291	0.29777	0.99746	0.82985
292	C292	0.39697	0.97456	0.54314
293	C293	0.47552	0.98415	0.53692
294	H294	0.48211	0.96389	0.48759
295	H295	0.51347	1.04245	0.5408
296	H296	0.36075	0.91605	0.5376
297	C297	0.31879	0.94906	0.71925
298	H298	0.28096	0.89359	0.73559
299	O299	0.40103	1.06399	0.6241
300	O300	0.43832	0.99351	0.76799
301	C301	0.3842	0.99784	0.61062
302	C302	0.37257	0.98284	0.77349
303	N303	0.35125	0.94405	0.6568
304	H304	0.35035	0.89664	0.64845

305	C305	0.10521	1.02974	0.40063
306	C306	0.11903	1.10714	0.37731
307	H307	0.14257	1.14612	0.41841
308	H308	0.15773	1.12699	0.33531
309	H309	0.04794	1.00188	0.41496
310	C310	0.06186	0.93162	0.23027
311	H311	0.08925	0.89688	0.23635
312	C312	0.15342	1.00971	0.06062
313	C313	0.227	1.00668	0.04936
314	H314	0.26812	1.04714	0.08228
315	H315	0.22156	0.95225	0.06104
316	H316	0.2472	1.01991	-0.00262
317	H317	0.16759	1.06666	0.07403
318	O318	0.17247	0.9835	0.33293
319	O319	0.1168	1.052	0.18359
320	O320	0.05121	0.94257	-0.01482
321	C321	0.11451	0.99004	0.34121
322	C322	0.09848	0.98483	0.17536
323	C323	0.11127	0.99808	-0.00342
324	N324	0.06043	0.96508	0.29417
325	H325	0.01851	0.9754	0.3019
326	N326	0.1165	0.96193	0.11675
327	H327	0.10633	0.91057	0.11279
328	N328	0.14602	1.02992	0.46053
329	H329	0.19201	1.07287	0.47044
330	C330	0.1198	1.05918	-0.11479

331	H331	0.06158	1.04524	-0.11466
332	C332	0.09505	0.94236	0.72949
333	C333	0.02175	0.87377	0.71514
334	H334	0.01509	0.85837	0.66139
335	H335	0.02009	0.8273	0.7436
336	H336	-0.02422	0.88393	0.72962
337	H337	0.13498	0.9226	0.74405
338	C338	0.1807	0.97805	0.55592
339	H339	0.22395	1.03474	0.55915
340	O340	0.19405	1.01082	0.83471
341	O341	0.1291	1.04954	0.65653
342	O342	0.07244	0.91112	0.48982
343	C343	0.13414	1.01434	0.8349
344	C344	0.12546	0.98714	0.66745
345	C345	0.12922	0.97021	0.50007
346	N346	0.08502	0.9822	0.78495
347	H347	0.04114	0.98871	0.78404
348	N348	0.14682	0.95225	0.62084
349	H349	0.13337	0.90123	0.62999
350	N350	0.14445	1.05387	-0.048
351	H351	0.19245	1.09149	-0.03494
352	C352	0.19699	0.74278	0.0637
353	H353	0.1818	0.68476	0.06475
354	C354	0.07328	0.72609	0.20962
355	C355	-0.0007	0.65788	0.21373
356	H356	-0.03721	0.658	0.17369

357	H357	0.00784	0.60886	0.20915
358	H358	-0.02617	0.65665	0.26248
359	H359	0.06177	0.77189	0.19626
360	C360	0.12386	0.80664	0.3865
361	H361	0.18131	0.81797	0.38479
362	O362	0.14477	0.81771	0.09981
363	O363	0.17484	0.74281	0.2815
364	O364	0.03531	0.70167	0.44345
365	C365	0.14903	0.75851	0.10807
366	C366	0.11567	0.74732	0.2739
367	C367	0.09246	0.76114	0.44662
368	N368	0.11337	0.71144	0.15779
369	H369	0.11722	0.66581	0.16025
370	N370	0.08928	0.77401	0.32395
371	H371	0.04035	0.76849	0.31884
372	N372	0.19233	0.76737	0.99743
373	H373	0.22985	0.81498	0.98413
374	C374	0.13358	0.73213	0.55462
375	C375	0.21121	0.73717	0.55951
376	H376	0.25364	0.79204	0.54684
377	H377	0.22361	0.72135	0.60924
378	H378	0.09525	0.67566	0.54109
379	C379	0.06433	0.68163	0.73147
380	H380	0.02306	0.70162	0.7305
381	C381	0.14874	0.78189	0.89177
382	C382	0.1467	0.85619	0.90226

383	H383	0.17431	0.89002	0.85843
384	H384	0.09098	0.85057	0.90591
385	H385	0.17608	0.88396	0.94776
386	H386	0.20514	0.795	0.87726
387	O387	0.10739	0.80019	0.63929
388	O388	0.18999	0.74664	0.76485
389	O389	0.08097	0.67706	0.96704
390	C390	0.10835	0.74135	0.62134
391	C391	0.12313	0.72603	0.77883
392	C392	0.13567	0.73703	0.95417
393	N393	0.0914	0.68449	0.66463
394	H394	0.10279	0.64509	0.64933
395	N395	0.10143	0.74592	0.83673
396	H396	0.05105	0.73871	0.83997
397	N397	0.13559	0.78123	0.50195
398	H398	0.17581	0.83013	0.50011
399	H399	0.25655	0.63877	0.21943
400	H400	0.41416	0.88169	0.88356
401	H401	0.47413	0.70977	0.55391
402	H402	0.28675	0.97941	0.71226
403	H403	0.00583	0.8969	0.21384
404	H404	0.1497	1.11567	-0.13236
405	H405	0.20705	0.94403	0.54139
406	H406	0.25281	0.77286	0.08281
407	H407	0.12041	0.85788	0.39413
408	H408	0.03765	0.62498	0.74849

409	O409	0.48848	0.94177	0.58783
410	H410	0.53845	0.97107	0.60235
411	O411	0.05129	1.10483	0.35483
412	H412	0.04642	1.14225	0.38253
413	O413	0.43172	0.86438	0.32794
414	H414	0.46933	0.85395	0.35057
415	O415	0.2089	0.68556	0.51095
416	H416	0.25337	0.67961	0.51605
417	C417	0.93806	0.48829	0.04871
418	H418	0.90778	0.42901	0.04505
419	C419	0.83152	0.48725	0.20677
420	C420	0.761	0.41566	0.21573
421	H421	0.72588	0.40705	0.17209
422	H422	0.77578	0.37063	0.21924
423	H423	0.72985	0.41462	0.26122
424	H424	0.81753	0.53078	0.19208
425	O425	0.89385	0.57162	0.08785
426	O426	0.92906	0.50845	0.28254
427	C427	0.89985	0.51364	0.09736
428	C428	0.87048	0.51197	0.27193
429	N429	0.87105	0.47217	0.1535
430	H430	0.87464	0.42598	0.1555
431	C431	0.86774	0.56236	0.38535
432	H432	0.92754	0.58941	0.38408
433	C433	0.84336	0.50452	0.43775
434	N434	0.88801	0.52515	0.49157

435	H435	0.91886	0.57744	0.49342
436	N436	0.84059	0.53858	0.31897
437	H437	0.79285	0.53494	0.30913
438	H438	0.99212	0.50203	0.06956
439	H439	0.85097	0.60342	0.40065
440	O440	0.79035	0.44183	0.43129
441	C441	0.81972	0.19985	0.20257
442	H442	0.83711	0.15641	0.20125
443	C443	0.92293	0.27744	0.04245
444	C444	0.99198	0.26581	0.03692
445	H445	1.02761	0.29431	0.07938
446	H446	0.97984	0.20811	0.03858
447	H447	1.02085	0.28846	-0.01017
448	H448	0.9401	0.3357	0.05305
449	C449	0.86802	0.31399	0.87121
450	H450	0.80859	0.28697	0.8771
451	O451	0.87529	0.32308	0.16114
452	O452	0.81945	0.20047	0.97329
453	O453	0.95402	0.2866	0.81258
454	C454	0.86167	0.25683	0.15296
455	C455	0.87743	0.25915	0.98029
456	C456	0.88961	0.27877	0.81727
457	N457	0.88553	0.23378	0.09847
458	H458	0.87859	0.18341	0.09836
459	N459	0.90254	0.31282	0.93339
460	H460	0.95061	0.35252	0.94227

461	N461	0.82551	0.22595	0.27046
462	H462	0.78436	0.23365	0.28719
463	C463	0.85212	0.19686	0.72023
464	C464	0.77759	0.13279	0.70139
465	H465	0.76573	0.0858	0.73405
466	H466	0.73334	0.14822	0.70697
467	H467	0.77641	0.11528	0.64885
468	H468	0.8903	0.17613	0.73703
469	C469	0.94072	0.2424	0.54938
470	H470	0.9617	0.3021	0.54954
471	C471	0.86976	0.26229	0.38152
472	C472	0.88746	0.34483	0.37718
473	H473	0.88107	0.36396	0.42721
474	H474	0.84823	0.35144	0.34334
475	H475	0.81139	0.23041	0.39252
476	O476	0.89486	0.30744	0.65748
477	O477	0.82323	0.1648	0.50169
478	O478	0.93863	0.23554	0.29672
479	C479	0.88884	0.24366	0.66226
480	C480	0.88778	0.21363	0.49285
481	C481	0.88132	0.23968	0.3138
482	N482	0.91067	0.21165	0.61371
483	H483	0.89925	0.16088	0.62191
484	N484	0.91161	0.25022	0.43403
485	H485	0.96404	0.27622	0.42998
486	N486	0.83848	0.23594	0.77334

487	H487	0.79013	0.23293	0.77584
488	C488	0.90023	0.54053	0.87486
489	C489	0.87721	0.59977	0.89637
490	H490	0.89464	0.64157	0.85787
491	H491	0.81791	0.57641	0.90221
492	H492	0.90348	0.62666	0.94393
493	H493	0.95708	0.56833	0.85805
494	O494	0.83472	0.434	0.94028
495	C495	0.89225	0.49351	0.93456
496	N496	0.94528	0.51967	0.98265
497	H497	0.98724	0.56661	0.97442
498	N498	0.85507	0.49916	0.81982
499	H499	0.80377	0.48789	0.82276
500	C500	0.89956	0.48174	0.54337
501	C501	0.98298	0.50763	0.55263
502	H502	1.01104	0.52858	0.5049
503	H503	1.00581	0.55236	0.5893
504	H504	0.87536	0.42454	0.52867
505	C505	0.81723	0.42357	0.71783
506	H506	0.79549	0.3676	0.73626
507	O507	0.87443	0.5456	0.63203
508	O508	0.9405	0.48413	0.75789
509	C509	0.86866	0.48554	0.60995
510	C510	0.87457	0.47055	0.76594
511	N511	0.84403	0.42638	0.65022
512	H512	0.85451	0.38668	0.63621

513	C513	0.61949	0.5397	0.39463
514	C514	0.62925	0.61651	0.37321
515	H515	0.64227	0.6522	0.41752
516	H516	0.67435	0.64299	0.33737
517	H517	0.56204	0.50974	0.40806
518	C518	0.57247	0.42651	0.23329
519	H519	0.6048	0.39627	0.23771
520	C520	0.65439	0.51014	0.06358
521	C521	0.73149	0.5179	0.04887
522	H522	0.76596	0.54344	0.09294
523	H523	0.73028	0.4648	0.03785
524	H524	0.75515	0.55451	0.00621
525	H525	0.66273	0.5645	0.07766
526	O526	0.69195	0.51346	0.31469
527	O527	0.60468	0.54426	0.1844
528	O528	0.54975	0.42644	0.00023
529	C529	0.62926	0.50286	0.33349
530	C530	0.60272	0.48175	0.17881
531	C531	0.60422	0.48841	0.00407
532	N532	0.5699	0.45791	0.29754
533	H533	0.52148	0.44645	0.31517
534	N534	0.6269	0.46347	0.12112
535	H535	0.62694	0.41614	0.11702
536	N536	0.66227	0.54018	0.4538
537	H537	0.70886	0.58257	0.46289
538	C538	0.57912	0.52882	0.89296

539	H539	0.52422	0.48313	0.89703
540	C540	0.60592	0.44789	0.73106
541	C541	0.53966	0.37453	0.71405
542	H542	0.53802	0.35909	0.66076
543	H543	0.54097	0.33104	0.74494
544	H544	0.48893	0.3782	0.72418
545	H545	0.64863	0.43523	0.75173
546	C546	0.69602	0.50087	0.55659
547	H547	0.72312	0.5597	0.56499
548	O548	0.68411	0.52958	0.84195
549	O549	0.64387	0.55621	0.6674
550	O550	0.58468	0.42708	0.49505
551	C551	0.6174	0.51398	0.83713
552	C552	0.64121	0.49355	0.67122
553	C553	0.64359	0.48556	0.49972
554	N554	0.57944	0.48234	0.78095
555	H555	0.53194	0.4815	0.7745
556	N556	0.6654	0.46584	0.61994
557	H557	0.65717	0.41583	0.62599
558	N558	0.61974	0.53885	0.95472
559	H559	0.66638	0.58314	0.95781
560	C560	0.7158	0.26876	0.04749
561	H561	0.71033	0.2131	0.04938
562	C562	0.57969	0.22169	0.19569
563	C563	0.52372	0.13994	0.19267
564	H564	0.51879	0.11672	0.14178

565	H565	0.54256	0.10902	0.2259
566	H566	0.47019	0.13319	0.20977
567	H567	0.55338	0.25644	0.1832
568	C568	0.59906	0.28711	0.38319
569	H569	0.6588	0.32089	0.3808
570	O570	0.64972	0.33085	0.08651
571	O571	0.67859	0.26384	0.2709
572	O572	0.53577	0.17948	0.45394
573	C573	0.66149	0.27594	0.09336
574	C574	0.61103	0.2443	0.26527
575	C575	0.58505	0.24396	0.44688
576	N576	0.63229	0.22693	0.14445
577	H577	0.64575	0.18738	0.14447
578	N578	0.57116	0.24919	0.31901
579	H579	0.51904	0.23003	0.31231
580	N580	0.71373	0.29	0.97818
581	H581	0.75309	0.336	0.96319
582	C582	0.65046	0.24132	0.55298
583	C583	0.73425	0.27679	0.54946
584	H584	0.75446	0.33608	0.5507
585	H585	0.75891	0.26062	0.59176
586	H586	0.63166	0.18298	0.54316
587	C587	0.57507	0.17896	0.72639
588	H588	0.5267	0.18876	0.72698
589	C589	0.6575	0.30154	0.87702
590	C590	0.63697	0.36296	0.89667

591	H591	0.66355	0.40711	0.86065
592	H592	0.57751	0.34326	0.89608
593	H593	0.65692	0.38499	0.94725
594	H594	0.71459	0.32726	0.85892
595	O595	0.63168	0.31037	0.64216
596	O596	0.69484	0.23396	0.77444
597	O597	0.60191	0.1965	0.95064
598	C598	0.62704	0.24985	0.62132
599	C599	0.62976	0.22481	0.77608
600	C600	0.65523	0.25724	0.93672
601	N601	0.60444	0.18854	0.65969
602	H602	0.60689	0.14586	0.64059
603	N603	0.60969	0.25895	0.82429
604	H604	0.55936	0.25102	0.82484
605	N605	0.62884	0.27574	0.50037
606	H606	0.65563	0.32851	0.49753
607	H607	0.76241	0.17611	0.1863
608	H608	0.88503	0.37045	0.85618
609	H609	0.98669	0.23104	0.54039
610	H610	0.77238	0.43872	0.71762
611	H611	0.51689	0.38733	0.21925
612	H612	0.57389	0.57779	0.8796
613	H613	0.7391	0.48675	0.5401
614	H614	0.76967	0.30571	0.0674
615	H615	0.57399	0.32402	0.39091
616	H616	0.55713	0.12214	0.74104

617	O617	0.9976	0.44926	0.57635
618	H618	1.03081	0.44354	0.54392
619	O619	0.56273	0.60891	0.34135
620	H620	0.54713	0.6387	0.37046
621	O621	0.95911	0.39103	0.35093
622	H622	0.99442	0.38256	0.37919
623	O623	0.75116	0.25453	0.48614
624	H624	0.76297	0.21383	0.49663
625	C625	0.94516	1.00856	0.04941
626	H626	0.93262	0.95063	0.04965
627	C627	0.81882	0.98077	0.19815
628	C628	0.75255	0.90693	0.211
629	H629	0.71352	0.89439	0.16986
630	H630	0.77039	0.86404	0.21442
631	H631	0.72394	0.90674	0.25783
632	H632	0.79862	1.02084	0.18681
633	O633	0.89508	1.08312	0.09408
634	O634	0.92397	1.00307	0.26575
635	C635	0.8966	1.02247	0.09565
636	C636	0.8638	1.00546	0.26025
637	N637	0.85833	0.97108	0.14231
638	H638	0.86344	0.92569	0.14171
639	C639	0.86734	1.057	0.37533
640	H640	0.92507	1.06918	0.37607
641	C641	0.83406	1.00632	0.43245
642	N642	0.86845	1.03514	0.49125

643	H643	0.88231	1.08647	0.49742
644	N644	0.83686	1.03082	0.31044
645	H645	0.78921	1.02716	0.30298
646	H646	1.0004	1.03945	0.06913
647	H647	0.86288	1.10734	0.38479
648	O648	0.78785	0.94019	0.42653
649	C649	0.81465	0.68782	0.22387
650	H650	0.83822	0.64939	0.22475
651	C651	0.9154	0.76953	0.0602
652	C652	0.9857	0.75965	0.05082
653	H653	1.02453	0.7933	0.0888
654	H654	0.97723	0.70335	0.05588
655	H655	1.01088	0.77888	0.00082
656	H656	0.93197	0.82662	0.07552
657	C657	0.87832	0.81977	0.88627
658	H658	0.81953	0.80042	0.88698
659	O659	0.86255	0.80733	0.17779
660	O660	0.81863	0.70112	0.98224
661	O661	0.9548	0.77179	0.84409
662	C662	0.84869	0.74152	0.16902
663	C663	0.87529	0.75896	0.99591
664	C664	0.89564	0.77585	0.83731
665	N665	0.87238	0.72106	0.1123
666	H666	0.86272	0.6701	0.10755
667	N667	0.90431	0.81641	0.95312
668	H668	0.9496	0.85748	0.96768

669	N669	0.81894	0.71872	0.29047
670	H670	0.77534	0.72419	0.30237
671	C671	0.8599	0.69758	0.73455
672	C672	0.79237	0.62232	0.73335
673	H673	0.79833	0.58634	0.77086
674	H674	0.74309	0.62811	0.74514
675	H675	0.7839	0.59491	0.68429
676	H676	0.9067	0.68809	0.74827
677	C677	0.93372	0.71701	0.55782
678	H678	0.97811	0.77299	0.55974
679	C679	0.86191	0.76838	0.40078
680	C680	0.87497	0.84807	0.38986
681	H681	0.85913	0.86746	0.43506
682	H682	0.84272	0.85209	0.34705
683	H683	0.80396	0.73533	0.41221
684	O684	0.87696	0.79024	0.65152
685	O685	0.82116	0.65337	0.49685
686	O686	0.93992	0.75396	0.31839
687	C687	0.87768	0.73081	0.66629
688	C688	0.8822	0.70885	0.50282
689	C689	0.87623	0.74327	0.33462
690	N690	0.90122	0.69482	0.62289
691	H691	0.8939	0.64659	0.63526
692	N692	0.9043	0.76364	0.45746
693	H693	0.95326	0.8037	0.46358
694	N694	0.84895	0.74169	0.78554

695	H695	0.80315	0.74542	0.78358
696	C696	0.88752	1.03463	0.87512
697	C697	0.852	1.08676	0.87138
698	H698	0.79932	1.06382	0.8967
699	H699	0.88779	1.1384	0.8959
700	H700	0.8422	1.09735	0.81948
701	H701	0.94588	1.0667	0.86319
702	O702	0.8268	0.93901	0.95639
703	C703	0.88176	0.99863	0.94157
704	N704	0.94106	1.03208	0.98199
705	H705	0.98046	1.07716	0.9647
706	N706	0.85051	0.98313	0.82379
707	H707	0.79741	0.95916	0.8262
708	C708	0.89623	1.0002	0.5384
709	C709	0.9807	1.04235	0.53924
710	H710	1.00293	1.06655	0.49006
711	H711	0.99704	1.08724	0.57524
712	H712	0.88267	0.94478	0.52197
713	C713	0.83352	0.9234	0.71553
714	H714	0.81955	0.86649	0.72762
715	O715	0.861	1.04814	0.63477
716	O716	0.95085	0.99561	0.76838
717	C717	0.86991	0.99609	0.60873
718	C718	0.88215	0.96863	0.76997
719	N719	0.86302	0.93918	0.6482
720	H720	0.87819	0.904	0.62923

721	C721	0.61762	1.01425	0.39433
722	C722	0.62657	1.09253	0.40325
723	H723	0.61526	1.10072	0.45589
724	H724 .	0.68245	1.13288	0.39183
725	H725	0.56031	0.97744	0.40469
726	C726	0.56761	0.93612	0.2163
727	H727	0.59347	0.89948	0.2184
728	C728	0.65311	1.01142	0.04492
729	C729	0.72182	0.99828	0.04109
730	H730	0.76388	1.03599	0.07557
731	H731	0.71134	0.94226	0.05341
732	H732	0.74388	1.0097	-0.00989
733	H733	0.67209	1.06969	0.05586
734	O734	0.69202	1.01078	0.30045
735	O735	0.62426	1.05733	0.16966
736	O736	0.5608	0.93318	-0.03244
737	C737	0.62989	0.99799	0.32312
738	C738	0.60262	0.98987	0.16152
739	C739	0.60877	0.99707	-0.01845
740	N740	0.56993	0.96565	0.28312
741	H741	0.52294	0.96026	0.30062
742	N742	0.61497	0.96632	0.10106
743	H743	0.60134	0.91463	0.09767
744	N744	0.6629	1.00648	0.44611
745	H745	0.71429	1.04146	0.44668
746	C746	0.58699	1.04892	-0.12502

747	H747	0.52987	1.00899	-0.12152
748	C748	0.6012	0.94782	0.72095
749	C749	0.53185	0.87579	0.70731
750	H750	0.53757	0.84868	0.66224
751	H751	0.5211	0.83772	0.74903
752	H752	0.48395	0.88515	0.69895
753	H753	0.64125	0.93265	0.74449
754	C754	0.69627	0.97478	0.5506
755	H755	0.73571	1.03297	0.5496
756	O756	0.68489	1.03098	-0.1716
757	O757	0.62866	1.04145	0.6373
758	O758	0.57608	0.90698	0.49872
759	C759	0.62099	1.02404	-0.17818
760	C760	0.63592	0.98668	0.65747
761	C761	0.64022	0.95861	0.49757
762	N762	0.58009	0.98896	0.76771
763	H763	0.53235	0.98832	0.7638
764	N764	0.66785	0.95531	0.61734
765	H765	0.66227	0.90837	0.63254
766	N766	0.61895	1.0553	-0.05894
767	H767	0.65316	1.10443	-0.04444
768	C768	0.65987	0.69155	0.04858
769	H769	0.61533	0.63711	0.04065
770	C770	0.56635	0.70731	0.20969
771	C771	0.49341	0.6385	0.22295
772	H772	0.45371	0.63297	0.1842

773	H773	0.5021	0.58958	0.22247
774	H774	0.46955	0.64013	0.27208
775	H775	0.55249	0.74939	0.19052
776	C776	0.60384	0.80323	0.37982
777	H777	0.66376	0.83193	0.37864
778	O778	0.6463	0.78965	0.10155
779	O779	0.67658	0.76005	0.27552
780	O780	0.53198	0.69152	0.43812
781	C781	0.63667	0.72448	0.10415
782	C782	0.61001	0.74413	0.27044
783	C783	0.58165	0.75637	0.44042
784	N784	0.6027	0.68489	0.15892
785	H785	0.6018	0.63767	0.16582
786	N786	0.57394	0.76088	0.31957
787	H787	0.52109	0.74045	0.31739
788	N788	0.67558	0.73126	0.98602
789	H789	0.72648	0.76715	0.9791
790	C790	0.63258	0.73435	0.54514
791	C791	0.71052	0.74032	0.53816
792	H792	0.73561	0.76682	0.49048
793	H793	0.74827	0.77024	0.57974
794	H794	0.59278	0.67802	0.53389
795	C795	0.58415	0.68262	0.7273
796	H796	0.53571	0.69257	0.72913
797	C797	0.6515	0.78616	0.88711
798	C798	0.63766	0.85313	0.90205

799	H799	0.66287	0.89533	0.86301
800	H800	0.57879	0.83495	0.90385
801	H801	0.66004	0.87632	0.95114
802	H802	0.71034	0.80593	0.88065
803	O803	0.63017	0.8057	0.63859
804	O804	0.70639	0.74246	0.77384
805	O805	0.55853	0.67903	0.94441
806	C806	0.6187	0.74401	0.61623
807	C807	0.64081	0.73177	0.77423
808	C808	0.62453	0.72738	0.94032
809	N809	0.60795	0.68824	0.65893
810	H810	0.62076	0.65005	0.64289
811	N811	0.61568	0.75835	0.82495
812	H812	0.56653	0.75346	0.82142
813	N813	0.62271	0.78113	0.49652
814	H814	0.65066	0.83344	0.50161
815	H815	0.75695	0.6568	0.21011
816	H816	0.90369	0.87635	0.86815
817	H817	0.95765	0.68058	0.54487
818	H818	0.78262	0.92903	0.71615
819	H819	0.51018	0.90303	0.2021
820	H820	0.58992	1.10111	-0.14148
821	H821	0.72526	0.94297	0.53703
822	H822	0.70877	0.68736	0.06513
823	H823	0.58357	0.84401	0.38482
824	H824	0.56711	0.62657	0.74474

247

825	O825	1.00998	0.99289	0.55736
826	H826	1.05005	1.01879	0.58948
827	O827	0.57502	1.10216	0.36149
828	H828	0.56261	1.13479	0.38755
829	O829	0.95003	0.89379	0.37644
830	H830	0.96716	0.86329	0.35003
831	O831	0.69633	0.66585	0.53821
832	H832	0.74202	0.66362	0.52604

A.3 Hydrated orthorhombic model

a = 6.100; b = 10.400; c = 27.200

α = 90.000; β = 90.000; γ = 90.000; Spacegroup = 1 P 1

244 atoms; Created by CERIUS

1	N1	0.98483	0.14428	0.3631
2	C2	0.93507	0.0155	0.35402
3	C3	0.85139	-0.00346	0.30476
4	O4	0.97325	-0.0176	0.26789
5	N5	0.60686	0.08407	0.46943
6	C6	0.53458	0.20852	0.45349
7	C7	0.71282	0.28313	0.43044
8	O8	0.8904	0.29493	0.45322
9	N9	0.70531	0.32994	0.38396
10	C10	0.89308	0.35959	0.35574
11	C11	0.98955	0.24602	0.33257
12	O12	1.04155	0.24044	0.28855
13	C13	1.0821	-0.09339	0.37101
14	H14	0.80066	0.01445	0.3794
15	H15	0.61041	0.01167	0.44773
16	C16	0.34913	0.18122	0.41879
17	H17	0.47212	0.26724	0.48388
18	H18	0.86415	0.43626	0.33019
19	H19	1.0061	0.39313	0.38183

20	H20	0.98881	-0.18289	0.36728
21	O21	1.27355	-0.09754	0.34181
22	H22	1.12407	-0.08275	0.40956
23	H23	0.56903	0.34043	0.36664
24	H24	0.97476	0.17306	0.39634
25	N25	0.82823	0.12424	0.12474
26	C26	0.80862	0.03932	0.0847
27	C27	0.71106	0.07666	0.51274
28	O28	0.62051	0.13295	0.54752
29	N29	0.63657	0.01477	0.30179
30	C30	0.53489	0.06742	0.26025
31	C31	0.53	0.20451	0.26249
32	O32	0.48284	0.2581	0.30129
33	N33	0.56159	0.28874	0.22734
34	C34	0.63834	0.27366	0.17905
35	C35	0.65567	0.15083	0.15337
36	O36	0.49548	0.07672	0.15388
37	H37	0.76739	-0.05626	0.09684
38	H38	0.96654	0.03995	0.06764
39	H39	0.54489	-0.00163	0.32982
40	H40	0.62157	0.02888	0.22904
41	H41	0.36207	0.04276	0.25876
42	C42	0.5201	0.37321	0.14683
43	H43	0.80195	0.29793	0.18434
44	H44	0.53498	0.37528	0.23745
45	H45	0.96563	0.16526	0.13256

46	H46	1.22015	-0.08928	0.30709
47	H47	0.23774	0.1176	0.43683
48	H48	0.26802	0.27251	0.41275
49	H49	0.40335	0.13911	0.38412
50	H50	0.53102	0.47213	0.15721
51	H51	0.56937	0.36472	0.10868
52	H52	0.3528	0.35119	0.14884
53	N53	0.82221	0.18074	-0.13134
54	C54	0.89359	0.06599	-0.15187
55	C55	0.94132	0.09721	-0.20378
56	O56	1.1288	0.0971	-0.22243
57	N57	0.73031	0.15099	0.01202
58	C58	0.63662	0.26667	-0.00492
59	C59	0.78536	0.31999	-0.04151
60	O60	0.97589	0.28124	-0.04923
61	N61	0.71524	0.40596	-0.07366
62	C62	0.61993	0.36987	-0.11881
63	C63	0.62365	0.23396	-0.13235
64	O64	0.457	0.17219	-0.1455
65	C65	1.07061	0.00995	-0.11745
66	H66	0.75731	-0.00004	-0.15211
67	H67	0.87388	0.1234	0.00059
68	C68	0.40536	0.24956	-0.02555
69	H69	0.62427	0.33402	0.02553
70	H70	0.4549	0.39914	-0.11505
71	H71	0.69468	0.42154	-0.14852

72	H72	1.13164	0.0828	-0.09246
73	O73	0.97708	-0.09486	-0.08951
74	H74	1.21547	-0.01816	-0.13802
75	H75	0.6956	0.49181	-0.0603
76	H76	0.93362	0.23118	-0.11564
77	N77	1.05079	0.07019	0.55873
78	C78	0.92885	0.01947	0.51783
79	C79	0.64658	0.08885	0.05012
80	O80	0.44976	0.10016	0.06176
81	N81	0.75117	0.12381	-0.22706
82	C82	0.69219	0.10691	-0.27709
83	C83	0.8388	0.05007	-0.31376
84	O84	0.76821	-0.03859	-0.34071
85	N85	1.02823	0.11046	0.67598
86	C86	1.19964	0.06605	0.64301
87	C87	1.13526	0.00206	0.59609
88	O88	1.15037	-0.11725	0.59109
89	H89	1.01697	0.0417	0.485
90	H90	0.91064	-0.08335	0.51839
91	H91	0.63145	0.14582	-0.2048
92	H92	0.54674	0.04592	-0.27589
93	H93	0.6459	0.20035	-0.29026
94	C94	1.35374	0.18089	0.63544
95	H95	1.29582	-0.00565	0.66235
96	H96	1.04784	0.19333	0.69188
97	H97	1.07476	0.16162	0.55965

98	H98	1.019	-0.17168	-0.1096
99	H99	0.29078	0.20883	0.00005
100	H100	0.34352	0.34224	-0.03756
101	H101	0.41271	0.18369	-0.05621
102	H102	1.51707	0.1463	0.62861
103	H103	1.35392	0.24146	0.66807
104	H104	1.30905	0.24081	0.6041
105	O105	0.29672	0.82531	0.49481
106	H106	0.21151	0.83836	0.52483
107	H107	0.43298	0.83136	0.51463
108	O108	0.21524	0.91276	0.10205
109	H109	0.32772	0.95812	0.12099
110	H110	0.12516	0.88678	0.12927
111	N111	0.36572	0.63759	0.31798
112	C112	0.49268	0.61045	0.27621
113	C113	0.34888	0.56582	0.23723
114	O114	0.3305	0.62371	0.19725
115	N115	0.52125	0.59878	0.52589
116	C116	0.67258	0.56202	0.48901
117	C117	0.57778	0.56732	0.43986
118	O118	0.69525	0.59943	0.40383
119	N119	0.36165	0.5434	0.43639
120	C120	0.23596	0.60155	0.39942
121	C121	0.30027	0.55425	0.35146
122	O122	0.32113	0.43693	0.3452
123	C123	0.59854	0.74253	0.26729

124	H124	0.60859	0.53586	0.28491
125	H125	0.54486	0.68239	0.5404
126	C126	0.88317	0.64359	0.49482
127	H127	0.71548	0.46384	0.49591
128	H128	0.26379	0.70382	0.40312
129	H129	0.06577	0.585	0.40546
130	H130	0.72742	0.74376	0.24056
131	O131	0.43221	0.82811	0.24971
132	H132	0.67351	0.77268	0.30206
133	H133	0.29537	0.49722	0.46362
134	H134	0.33534	0.72705	0.32414
135	N135	0.08275	0.69833	0.05882
136	C136	-0.13207	0.72969	0.04436
137	C137	0.3746	0.52512	0.54949
138	O138	0.31728	0.4171	0.53346
139	N139	0.20288	0.47259	0.24721
140	C140	-0.01851	0.49999	0.23485
141	C141	-0.07059	0.52987	0.18342
142	O142	-0.19336	0.62045	0.17147
143	N143	0.01233	0.45503	0.14805
144	C144	0.01227	0.48092	0.09657
145	C145	0.14972	0.59102	0.08122
146	O146	0.35184	0.58986	0.08617
147	H147	-0.24249	0.71657	0.0753
148	H148	-0.13756	0.82799	0.03173
149	H149	0.24499	0.4014	0.26929

150	H150	-0.11824	0.42212	0.24714
151	H151	-0.05581	0.58456	0.25617
152	C152	0.07633	0.36098	0.06683
153	H153	-0.15462	0.50339	0.08662
154	H154	0.06586	0.37343	0.15955
155	H155	0.19828	0.75749	0.05255
156	H156	0.45316	0.83314	0.21395
157	H157	0.89318	0.72241	0.46793
158	H158	0.89694	0.68522	0.53068
159	H159	1.02243	0.57975	0.49148
160	H160	0.10869	0.39077	0.02953
161	H161	-0.05445	0.29294	0.06668
162	H162	0.21881	0.30852	0.07924
163	N163	0.356	0.58587	0.80205
164	C164	0.54826	0.66066	0.79637
165	C165	0.61182	0.70175	0.7468
166	O166	0.50793	0.78684	0.72494
167	N167	-0.08959	0.67367	0.9612
168	C168	0.07376	0.59135	0.94185
169	C169	0.05407	0.57248	0.88839
170	O170	-0.00622	0.66118	0.86115
171	N171	0.09551	0.45428	0.87087
172	C172	0.13757	0.40166	0.82256
173	C173	0.32448	0.45974	0.79532
174	O174	0.4616	0.39573	0.77142
175	C175	0.5302	0.78112	0.82954

176	H176	0.67752	0.604	0.81192
177	H177	-0.08321	0.75806	0.94828
178	C178	0.29798	0.64468	0.95713
179	H179	0.06808	0.49971	0.95956
180	H180	0.18463	0.30352	0.8278
181	H181	-0.01303	0.39134	0.80162
182	H182	0.50048	0.86894	0.80834
183	O183	0.3662	0.76811	0.86605
184	H184	0.68459	0.79413	0.84839
185	H185	0.08402	0.39132	0.89661
186	H186	0.23966	0.62693	0.81937
187	N187	0.50296	0.63569	0.62183
188	C188	0.31442	0.58218	0.5973
189	C189	-0.19817	0.65313	0.00283
190	O190	-0.30988	0.5554	0.0094
191	N191	0.79171	0.67273	0.72245
192	C192	0.95876	0.58082	0.73291
193	C193	0.9438	0.45525	0.7092
194	O194	1.10219	0.38004	0.7073
195	N195	0.75449	0.42277	0.68889
196	C196	0.70625	0.43251	0.63785
197	C197	0.69008	0.57084	0.62698
198	O198	0.85002	0.64286	0.62123
199	H199	0.23114	0.51534	0.62107
200	H200	0.20372	0.65906	0.58836
201	H201	0.80855	0.72554	0.69353

202	H202	1.11773	0.61301	0.72091
203	H203	0.9684	0.57576	0.77159
204	C204	0.83319	0.34862	0.60054
205	H205	0.54136	0.39866	0.63518
206	H206	0.63535	0.42618	0.71267
207	H207	0.5133	0.72716	0.62771
208	H208	0.41737	0.82424	0.89299
209	H209	0.29345	0.74909	0.95682
210	H210	0.33564	0.61478	0.99415
211	H211	0.42812	0.60974	0.93348
212	H212	1.00238	0.37412	0.59549
213	H213	0.75487	0.35312	0.56469
214	H214	0.83249	0.24966	0.61214
215	O215	-0.00244	0.82694	0.1825
216	H216	0.09217	0.7519	0.19156
217	H217	-0.00205	0.87932	0.21167
218	O218	0.61341	0.85273	0.40952
219	H219	0.51839	0.84254	0.43818
220	H220	0.50536	0.86029	0.38353
221	O221	0.46922	0.83625	0.03623
222	H222	0.47761	0.82188	0.07283
223	H223	0.3487	0.89941	0.03603
224	O224	0.13565	0.21232	0.19001
225	H225	0.12907	0.20785	0.22558
226	H226	0.27066	0.17012	0.18092
227	O227	0.12457	0.81916	0.69613

228	H228	0.00263	0.85932	0.67979
229	H229	0.21139	0.7904	0.66765
230	O230	0.51758	0.94441	0.93592
231	H231	0.51077	0.91215	0.96982
232	H232	0.67301	0.93517	0.92795
233	O233	0.02911	0.83115	0.79941
234	H234	0.0928	0.82234	0.76666
235	H235	0.1459	0.79888	0.82072
236	O236	0.53417	0.81856	0.14654
237	H237	0.687	0.81608	0.15751
238	H238	0.47283	0.74217	0.16208
239	O239	0.63048	0.84186	0.56637
240	H240	0.64759	0.90537	0.59258
241	H241	0.71838	0.77453	0.58204
242	O242	0.13513	0.99354	0.99416
243	H243	-0.00601	0.95902	0.98414
244	H244	0.23739	0.98122	0.96609

A.4 Polyserine model for crystalline sericin (β-sheet)

a = 9.2;   b = 11.6;   c = 20.4
α = 90;   β = 90;   γ = 90
264   atoms   Created by CERIUS
:

1	C1	0.72804	0.25013	0.23252
2	C2	0.35959	0.30339	0.22882
3	C3	0.26701	0.1863	0.06141
4	C4	0.71235	0.18378	0.11818
5	C5	0.79508	0.22231	0.29679
6	C6	0.26381	0.27026	0.1738
7	C7	0.3336	0.24602	1.00677
8	N8	0.73751	0.19938	0.99497
9	N9	0.78827	0.19631	0.17486
10	N10	0.2959	0.25211	0.28485
11	N11	0.33122	0.2246	0.12133
12	O12	0.6186	0.25823	0.11057
13	O13	0.92963	0.20866	0.3018
14	O14	0.12663	0.26687	0.17589
15	O15	0.44487	0.20131	0.98249
16	C16	0.76848	0.37925	0.217
17	C17	0.34321	0.43697	0.2273
18	C18	0.70594	0.66566	0.05443
19	C19	0.21711	0.67284	0.06942
20	C20	0.27316	0.82149	0.23315
21	C21	0.7726	0.74106	0.28605

22	C22	0.66361	0.72177	0.11802
23	C23	0.30413	0.71032	0.01156
24	C24	0.20482	0.75765	0.17967
25	N25	0.76104	0.75452	0.16486
26	N26	0.64906	0.72228	0.99307
27	N27	0.27975	0.72848	0.12574
28	N28	0.22891	0.76655	0.2931
29	O29	0.89453	0.69019	0.2856
30	O30	0.55893	0.78687	0.12184
31	O31	0.40426	0.64001	0.00785
32	O32	0.06867	0.75101	0.17672
33	C33	0.65723	0.54132	0.06841
34	C34	0.17403	0.5422	0.07716
35	H35	0.62415	0.21922	0.24279
36	H36	0.46108	0.28302	0.21148
37	H37	0.15853	0.2221	0.06195
38	C38	0.23222	0.05641	0.0606
39	C39	0.76864	0.13102	0.05733
40	H40	0.63876	0.20456	0.97983
41	H41	0.19466	0.22917	0.2815
42	O42	0.79824	0.45043	0.27155
43	H43	0.87057	0.38101	0.18928
44	H44	0.68407	0.41818	0.18734
45	O45	0.47432	0.4795	0.25094
46	H46	0.33863	0.4747	0.17874
47	H47	0.24811	0.46803	0.25494

48	H48	0.82402	0.66935	0.0525
49	H49	0.11272	0.71561	0.06135
50	H50	0.3853	0.80053	0.22284
51	C51	0.27506	0.95339	0.23256
52	C52	0.7276	0.80884	0.22789
53	H53	0.37811	0.74281	0.12617
54	H54	0.8627	0.73537	0.15801
55	H55	0.12307	0.75106	0.29772
56	O56	0.79164	0.48932	0.08813
57	H57	0.56954	0.53029	0.10464
58	H58	0.60866	0.50613	0.02524
59	O59	0.0915	0.51272	0.13552
60	H60	0.11023	0.51495	0.03459
61	H61	0.27294	0.49097	0.0768
62	C62	0.71556	0.00428	0.05625
63	H63	0.88564	0.13102	0.06468
64	H64	0.43291	0.22157	0.12306
65	C65	0.82639	0.91617	0.22532
66	H66	0.61513	0.83355	0.23286
67	H67	0.88851	0.16949	0.17564
68	H68	0.54937	0.74315	0.98703
69	C69	0.73242	0.31768	0.56719
70	C70	0.30461	0.34599	0.56901
71	C71	0.25248	0.21942	0.39713
72	C72	0.69828	0.20202	0.46601
73	C73	0.79549	0.28009	0.63036

74	C74	0.23756	0.30118	0.508
75	C75	0.35551	0.24051	0.34406
76	N76	0.709	0.18944	0.34712
77	N77	0.78696	0.24859	0.51255
78	N78	0.24434	0.29547	0.62753
79	N79	0.31189	0.24138	0.46183
80	O80	0.57913	0.17326	0.48819
81	O81	0.9312	0.26724	0.63285
82	O82	0.10771	0.32572	0.49366
83	O83	0.45892	0.30805	0.35358
84	C84	0.76688	0.44843	0.55615
85	C85	0.31349	0.47311	0.58023
86	C86	0.73021	0.67839	0.40101
87	C87	0.26392	0.65557	0.39416
88	C88	0.24813	0.84862	0.54448
89	C89	0.78933	0.81334	0.61117
90	C90	0.66829	0.74867	0.45561
91	C91	0.31949	0.72843	0.33996
92	C92	0.20541	0.76222	0.49516
93	N93	0.76172	0.80083	0.49684
94	N94	0.68996	0.73557	0.34075
95	N95	0.30312	0.71279	0.4537
96	N96	0.20738	0.81567	0.60953
97	O97	0.91909	0.79306	0.61828
98	O98	0.56076	0.81202	0.44719
99	O99	0.44372	0.72444	0.32031

100	O100	0.07124	0.75112	0.48526
101	C101	0.68389	0.54887	0.39846
102	C102	0.33018	0.53357	0.39006
103	H103	0.62035	0.30607	0.57451
104	H104	0.41335	0.32051	0.55992
105	H105	0.1529	0.2681	0.38664
106	C106	0.1999	0.09381	0.3864
107	C107	0.7562	0.13901	0.40819
108	H108	0.60973	0.21174	0.34414
109	H109	0.13863	0.29377	0.63222
110	O110	0.86649	0.50096	0.60047
111	H111	0.82282	0.45757	0.51012
112	H112	0.66685	0.49945	0.55524
113	H113	0.37761	0.47827	0.6249
114	O114	0.39531	0.53508	0.53363
115	H115	0.20181	0.50431	0.58643
116	H116	0.84795	0.68026	0.40691
117	H117	0.14528	0.65329	0.3881
118	H118	0.36615	0.83993	0.54154
119	C119	0.23935	0.97912	0.53737
120	C120	0.72396	0.8678	0.5534
121	H121	0.40251	0.73156	0.459
122	H122	0.86419	0.78389	0.48968
123	H123	0.1089	0.79601	0.61658
124	O124	0.81128	0.47982	0.40377
125	H125	0.60696	0.52453	0.43687

126	H126	0.63751	0.53012	0.35071
127	O127	0.20624	0.4656	0.37438
128	H128	0.37778	0.49095	0.43047
129	H129	0.40757	0.52806	0.3519
130	C130	0.71012	0.01274	0.40047
131	H131	0.87122	0.14927	0.41203
132	H132	0.40695	0.21672	0.47162
133	C133	0.7951	0.99108	0.55693
134	H134	0.60702	0.87156	0.55928
135	H135	0.89205	0.24916	0.50457
136	H136	0.61015	0.79145	0.34377
137	C137	0.82996	0.33878	0.90246
138	C138	0.30949	0.3478	0.89667
139	C139	0.24071	0.19802	0.73347
140	C140	0.71827	0.24177	0.80019
141	C141	0.84321	0.2618	0.96064
142	C142	0.22511	0.29783	0.84132
143	C143	0.32131	0.25176	0.67842
144	N144	0.70919	0.24665	0.68043
145	N145	0.82626	0.28056	0.83996
146	N146	0.25221	0.31135	0.96239
147	N147	0.29569	0.24127	0.79421
148	O148	0.58907	0.25068	0.81506
149	O149	0.96214	0.2702	0.98893
150	O150	0.10115	0.32645	0.82733
151	O151	0.43624	0.20891	0.66043

152	C152	0.7061	0.42274	0.91511
153	C153	0.3117	0.47916	0.88579
154	C154	0.75154	0.72483	0.72155
155	C155	0.25484	0.74703	0.71999
156	C156	0.32165	0.79659	0.90027
157	C157	0.74408	0.77436	0.94849
158	C158	0.67459	0.7766	0.77891
159	C159	0.29858	0.81587	0.66152
160	C160	0.24572	0.76606	0.8398
161	N161	0.75188	0.80201	0.83355
162	N162	0.70583	0.77555	0.66061
163	N163	0.31726	0.77869	0.78233
164	N164	0.24333	0.74996	0.9556
165	O165	0.87961	0.76107	0.95483
166	O166	0.56168	0.73341	0.79887
167	O167	0.42362	0.85772	0.65592
168	O168	0.11095	0.7455	0.83879
169	C169	0.71677	0.59416	0.72151
170	C170	0.30399	0.62215	0.70758
171	H171	0.92969	0.38951	0.89889
172	H172	0.42103	0.32097	0.89243
173	H173	0.12751	0.22379	0.72787
174	C174	0.23502	0.06617	0.73334
175	C175	0.76167	0.18715	0.73811
176	H176	0.60802	0.25314	0.67618
177	H177	0.15109	0.32209	0.97408

178	H178	0.65024	0.43954	0.87048
179	H179	0.62003	0.38718	0.94628
180	H180	0.35542	0.51664	0.92911
181	H181	0.38995	0.4999	0.8466
182	H182	0.86827	0.73792	0.72975
183	H183	0.13544	0.75083	0.72238
184	H184	0.42318	0.75124	0.90146
185	C185	0.32052	0.92836	0.9137
186	C186	0.6964	0.84591	0.89453
187	H187	0.40632	0.82169	0.78154
188	H188	0.85706	0.7909	0.83013
189	H189	0.13479	0.75735	0.9553
190	H190	0.66623	0.55843	0.67805
191	H191	0.82382	0.55625	0.71256
192	H192	0.22641	0.59587	0.67133
193	H193	0.28705	0.56709	0.75047
194	C194	0.71354	0.0586	0.7329
195	H195	0.88021	0.18525	0.73776
196	H196	0.39864	0.23329	0.80272
197	H197	0.58633	0.86812	0.89656
198	C198	0.78265	0.96455	0.89623
199	H199	0.91916	0.27227	0.82091
200	H200	0.60404	0.78712	0.65422
201	O201	0.45043	0.63477	0.69043
202	H202	0.49518	0.65366	0.73309
203	O203	0.17312	0.52936	0.86919

204	H204	0.09693	0.51907	0.90358
205	O205	0.77789	0.52342	0.9396
206	H206	0.80555	0.50645	0.98597
207	O207	0.68165	0.53836	0.78069
208	H208	0.74635	0.5691	0.81578
209	H209	0.82933	0.46068	0.35781
210	H210	0.80753	0.52414	0.13234
211	H211	-0.00735	0.49614	0.1205
212	H212	0.85231	0.524	0.2629
213	H213	0.16807	0.42849	0.41508
214	H214	0.93535	0.4375	0.60793
215	H215	0.48607	0.42725	0.28838
216	H216	0.31969	0.58234	0.50805
217	O217	0.83122	-0.05871	0.38451
218	H218	0.91661	-0.01928	0.40007
219	O219	0.83721	-0.01521	0.73028
220	H220	0.90099	0.00982	0.76587
221	O221	0.37824	0.02306	0.73528
222	H222	0.39685	0.01808	0.78117
223	O223	0.83648	-0.07328	0.05397
224	H224	0.92113	-0.03538	0.0726
225	O225	0.12059	0.03094	0.10758
226	H226	0.11601	0.09985	0.13472
227	O227	0.38237	0.9915	0.27782
228	H228	0.39247	1.07418	0.26995
229	O229	0.7528	1.00027	0.26006

230	H230	0.82722	1.05572	0.27024
231	O231	0.32066	1.04111	0.5845
232	H232	0.42276	1.0279	0.57308
233	O233	0.90501	1.01857	0.60431
234	H234	0.91106	1.10158	0.60796
235	O235	0.39968	1.00155	0.87194
236	H236	0.40993	1.07213	0.89781
237	O237	0.68962	1.06114	0.89628
238	H238	0.59583	1.03676	0.88156
239	H239	0.20874	0.96008	0.91367
240	H240	0.36972	0.94156	0.96162
241	H241	0.85055	1.01084	0.51164
242	H242	0.70726	1.05103	0.56665
243	H243	0.29211	1.00167	0.49234
244	H244	0.12723	1.00737	0.54027
245	H245	0.93306	0.90311	0.24761
246	H246	0.84427	0.95098	0.17663
247	H247	0.31313	0.98314	0.18521
248	H248	0.16843	0.9896	0.2426
249	H249	0.85307	0.97452	0.93849
250	H250	0.85776	0.97843	0.85572
251	H251	0.18388	0.03463	0.01385
252	H252	0.32923	0.00475	0.06958
253	H253	0.64296	-0.01101	0.01512
254	H254	0.6526	-0.01792	0.09972
255	H255	0.66284	-0.01179	0.44592

256	H256	0.62256	-0.00162	0.36516
257	H257	0.17432	0.03423	0.77566
258	H258	0.17965	0.03584	0.68944
259	H259	0.64603	0.03204	0.77392
260	H260	0.65099	0.04189	0.68849
261	O261	0.08997	0.06644	0.43142
262	H262	0.07929	0.13682	0.45731
263	H263	0.14835	0.0832	0.33892
264	H264	0.28887	0.03212	0.38945

APPENDIX B

TRIGONAL VERSUS ORTHORHOMBIC LATTICE

In the silk III crystal structure described in Chapters 2 a hexagonal arrangement of threefold helices results in a trigonal unit cell. In this instance the unit cell is primitive, and a representation of the two-dimensional lattice formed by the helices' positions (essentially the basal plane), the resulting single chain unit cell and a diffraction schematic are shown at the top of Figure 60. If two chains are included, while preserving the same hexagonal packing, a unit cell drawn on the same lattice would be twice as large. As shown at the bottom of the Figure, an orthorhombic two chain unit cell can be chosen that preserves the packing and translational periodicity. Single crystal diffraction from a unit cell containing two non-identical chains would have extra reflections, not observed in the primitive trigonal crystal structure.
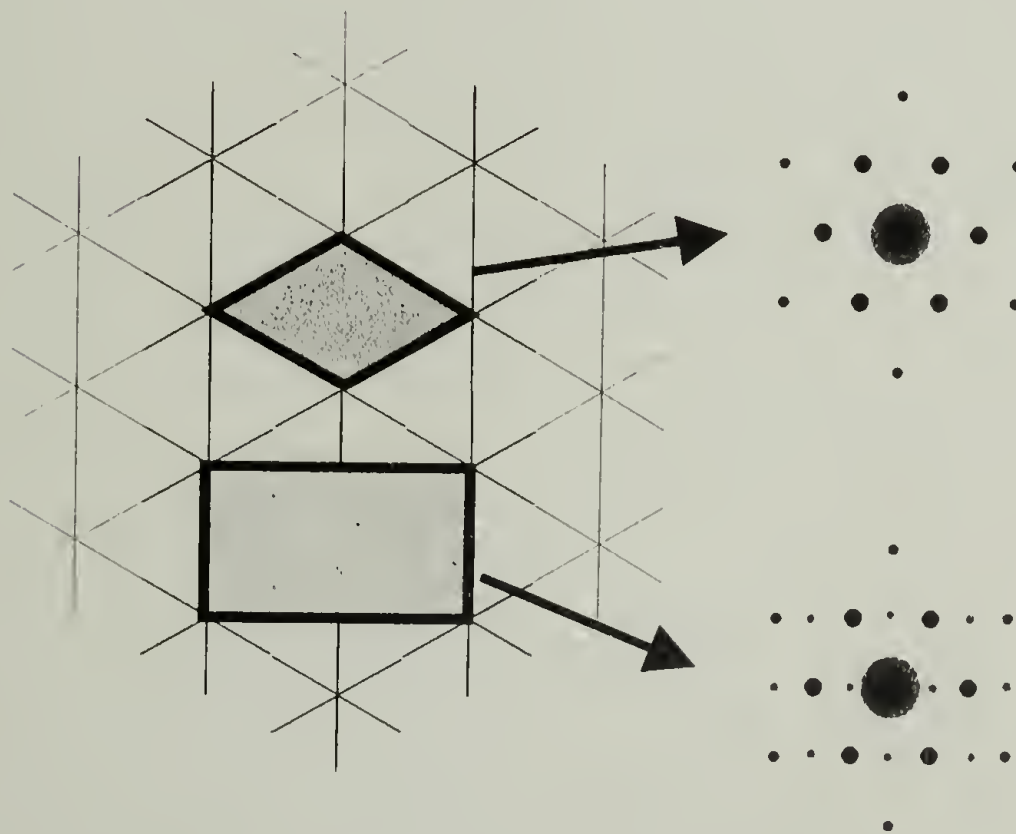
Figure 60: An orthorhombic unit cell can result if there are two chains per cell.

270

APPENDIX C

SAMPLE COMPOSITION

Amino acid analysis received on samples of solubilized regenerated silk were compared to published amino acid contents of *B.mori* fibroin and to the contents of the sericins. Since the higher molecular weight sericins are expected to be the last to dissolve in the degumming process the overall amino acid content for the sericion was considered in the sericin versus fibroin content calculation, and then the ratios were recalculated using the values published for high molecular weight sericin. While different names are given to the high molecular weight sericins, depending on how many different sericins were obtained using a particular fractionation method, for simplicity the charts refer to sericin 3 as the high molecular weight fraction and use Sprague's data.[6] As can be seen from the columns detailing published amino acid contents of fibroin; columns I, J, K, and L, there is some variation in reported amino acid contents for fibroin. "Method 1" refers to aqueous fibroin degummed using boiling water and $NaCO_3$; in "Method 2" SDS was also added. LiBr was used to solubilize the fibroin in both cases. "Method 1 LB" and "Method 2 LB" refer to samples that were degummed using methods 1 and 2 respectively, resolubilized, dialyzed to remove LiBr, and then placed into water in an LB trough. The resulting surface excess layers of fibroin were scraped off the surface of the LB trough onto a cover slip. The resulting films were only analyzable after resolubilization using boiling water and LiBr salt. Since all of the material was not resolubilized and because hot water and salt dissolve sericin much more readily than fibroin, any trace amounts of sericin in the film would be expected to be disproportionately represented in the subsequent amino acid analysis.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
	Method 1	Method 1LB	Method 2	Method 2 LB		Published cocoon		f-1	f-2	f-3	fibroin	s-1	s-2	s-3	sericin
gly	44.3	40.3	47.1	43.4		36.7		45.5	11.4	15.6	44.5	18	12.5	14.2	13.9
ala	24.9	22.1	27.5	24.7		23.4		29.5	13.7	8.5	29.3	6.2	6.9	6.2	4.6
val	1.7	2.1	1.7	2.1		2.4		1.8	5.5	3.1	2.2	3.7	2.8	1	3.2
leu	0.6	1.9	0.5	1.2		0.6		0.2	7	6.1	0.5	2.1	1.5	0.6	1.2
ile	1	0.7	1	0.2		0.7		0.3	6.4	4.5	0.7	1.6	1.3	0.4	0.7
ser	13.9	11.7	11.9	13.6		17.1		10.8	10.9	9.1	12.1	29.1	30.2	38.1	32.3
thr	1.6	2.4	0.9	1.5		2.8		1.5	3.4	4.8	0.9	8.5	8	4.1	8.4
asp	2.1	4.1	1.2	2.6		4.5		1	14.7	11.7	1.3	13	12.9	12.6	14.5
glu	1.3	2.9	1	2.3		2.3		1	8.9	6.4	1	5.4	5.2	10.7	4.8
lys	0.5	0.9	0.3	0.6		4.1		0.4	2.2	4.1	0.3	2.2	4.8	5.3	8.4
arg	0.8	1.7	0.5	1		1.1		0.3	4.1	3.8	0.5	4	3.9	3	2.3
his	0.5	1.6	0.3	0.6		0.4		0.7	2.2	3.3	0.2	1.6	1.5	0.8	1.6
tyr	4.2	3.6	4.6	3.3		4.8		6.2	3.1	9.5	5.2	3.3	6.8	2.1	2.6
phe	0.6	1	0.6	1.2		0.6		0.7	1.7	4.1	0.6	1.5	0.7	0.4	0.4
pro	0.6	1.9	0.5	1		0.5		0.4	3.4	2.9	0.3	0	0	0	0.4
met	0.2	0	0.1	0		0		0.2	0.3	0	0.1	0	0	0	0.1
cys	0	0	0	0		0		0.2	1.3	2.7	0.2	0	1	0.5	0.3
TOT	98.8	98.9	99.7	99.3	0	102	0	100.7	100.2	100.2	99.9	100.2	100	100	99.7

percent sericin
D fit

0.01	0.153736
0.02	0.142316
0.03	0.130896
0.04	0.122457
0.05	0.118832
0.06	0.115206
0.07	0.111581
0.08	0.110292 ***
0.09	0.112246
0.1	0.114602
0.075	0.11007 ***
0.11	0.119215

assumes sericin is impurity

in Method2 LB
solution

0.01	0.0928
0.02	0.0928
0.03	0.094
0.04	0.0952
0.05	0.0965
0.011	0.0927
0.012	0.0925
0.013	0.0924
0.014	0.0922
0.015	0.0921 **
0.016	0.0923
0.017	0.0924
0.018	0.0925
0.009	0.093
0	0.0943

percent sericin
in Method2 LB using only major residues
D fit   0.963

0.01	0.112	0.1	0.08
0.02	0.1081	0.11	0.083
0.03	0.1042	0.12	0.086
0.04	0.1004	0.13	0.09
0.05	0.0965	0.14	0.094
0.06	0.0926	0.15	0.098
0.07	0.0887	0.16	0.107
0.08	0.0849	0.17	0.116
0.09	0.081	0.18	0.125
0.1	0.0796 ***	0.19	0.134
0.005	0.1139	0.2	0.143
0.003	0.1147	0.21	0.152

assumes sericin is impurity

percent sericin 3 in Method2 LB

D fit

0.01	0.153505	0.1	0.1141
0.02	0.141853	0.11	0.115
0.03	0.130201	0.12	0.1171
0.04	0.120806	0.13	0.1203
0.05	0.115257	0.14	0.1247
0.06	0.110916 ***	0.15	0.1299
0.07	0.111611	0.16	0.135
0.08	0.112306	0.17	0.1402
0.09	0.113162	0.18	0.1454
0.1	0.114099	0.19	0.1507
0.11	0.115035		
0.12	0.1171		

assumes sericin is impurity

percent sericin 3 in Method2 LB using only major AA

D fit

0.01	0.1139	0.1	0.085
0.02	0.1084	0.11	0.086
0.03	0.103	0.12	0.088
0.04	0.0976	0.13	0.092
0.05	0.0922	0.14	0.097
0.06	0.0867	0.15	0.103
0.07	0.0822	0.16	0.11
0.08	0.0815 ***	0.17	0.121
0.09	0.083	0.18	0.132
0.1	0.0846	0.19	0.143
0.074	0.0805		
0.075	0.0807		

assumes sericin is impurity

percent sericin in Method 1 LB

D fit		solution	
0.01	0.260131	0.01	0.115
0.02	0.251304	0.02	0.1176
0.03	0.242477	0.03	0.1203
0.04	0.235349	0.04	0.1229
0.05	0.229171	0.05	0.1256
0.06	0.223236	0.06	0.1282
0.07	0.2173	0.07	0.1308
0.08	0.211365	0.08	0.1335
0.09	0.20543	0.09	0.1361
0.1	0.20455 ***	0.1	0.1388
0.11	0.207614	0.005	0.1137
0.12	0.210677	0.015	0.1163

assumes sericin is impurity

percent sericin in Method 1 LB using only major

D fit

0.01	0.1776	0.1	0.129
0.02	0.1713	0.11	0.129
0.03	0.165	0.12	0.13
0.04	0.1587	0.13	0.131
0.05	0.1524	0.14	0.133
0.06	0.1461	0.15	0.134
0.07	0.1397	0.16	0.136
0.08	0.1334	0.17	0.137
0.09	0.1282 ***	0.18	0.139
0.1	0.1286	0.19	0.141
0.074	0.1372		
0.075	0.1366		

assumes sericin is impurity

REFERENCES

1) Lotz, B.; Cesari, C. *Biochimie* **1979**, *61*, 205.

2. Shimura, K.; Kikuchi, A.; Ohtomo, K.; Katagata, Y.; Hyodo, A. *Journal of Biochemistry* **1976**, *80*, 693.

3. Viney, C.; Huber, A. E.; Dunaway, D. L.; Kerkham, K.; Case, S. T. *Optical Characterization of Silk Secretions and Fibers*; Viney, C.; Huber, A. E.; Dunaway, D. L.; Kerkham, K.; Case, S. T., Ed.; American Chemical Society: Washington D. C., 1994; Vol. 544, pp 120.

4. Kaplan, D.; Adams, W. W.; Farmer, B.; Viney, C. *Silk: Biology,Structure, Properties, and Genetics*; Kaplan, D.; Adams, W. W.; Farmer, B.; Viney, C., Ed.; American Chemical Society: Washington, D. C., 1994, pp 2-16.

5. Iizuka, E. *Journal of Applied Polymer Science: Applied Polymer Symposia* **1985**, *41*, 173-185.

6. Sprague, K. U. *Biochemistry* **1975**, *14*, 925-931.

7. Ishikawa, H.; Hirabayashi, K.; Hayakawa, T. *Sen-i Gakkaishi* **1969**, *25*, 425-430.

8. Marsh, R. E.; Corey, R. B.; Pauling, L. *Biochemica et Biophysica Acta* **1955**, *16*, 1 - 35.

9. Gamo, T.; Inokuchi, T.; Laufer, H. *Insect Biochemnistry* **1977**, *7*, 285 - 295.

10. Magoshi, J.; Magoshi, Y.; Nakamura, S. *Journal of Applied Polymer Science: Applied Polymer Symposium* **1985**, *41*, 187-204.

11. Magoshi, J.; Magoshi, Y.; Nakamura, S. *Mechanism of Fiber Formation of Silkworm*; Kaplan, D.; Adams, W. W.; Farmer, B.; Viney, C., Ed.; American Chemical Society: Washington, D. C., 1994, pp 292-310.

12. Fraser, R. D. B.; MacRae, T. P. *Silks.*; Fraser, R. D. B.; MacRae, T. P., Ed.; Academic Press: New York, 1973.

13. Kaplan, D. L.; Lombardi, S. J.; Muller, W. S.; Fossey, S. *Silks*; Byrom, D., Ed.; Stockton: New York, 1991

14. Mita, K.; Ichimura, S.; James, T. C. *J. Molecular Evolution* **1994**, *38*, 583-592.

15. Asakura, T.; Kaplan, D. L. *Silk Production and Processing*; Asakura, T.; Kaplan, D. L., Ed.; Academic Press: New York, 1994; Vol. 4, pp 1-11.

16. Kaplan, D. L.; Mello, C. M.; Fossey, S. A.; Arcidiacono, S.; Senecal, K.; Muller, W. S. *Silks*; McGrath, K.; Kaplan, D., Ed.; Birkhauser: Boston, 1997, pp 103-132.

17. Kaplan, D. L.; Fossey, S. A.; Viney, C.; Muller, W. S. *Self-organization (assembly) in biosynthesis of silk fibers - a hierarchical problem*; in "Hierarchically Structured Materials"; Aksay, 1. A.; Baer, E.; Sarikaya, M.; Tirrell, D. A., Ed., 1992; Vol. 255, pp 19-29.

18. Sinohara, H. *Insect Biochemistry* **1977**, *7*, 3.

19. Suzuki, Y.; Brown, D. D. *Journal of Molecular Biology* **1972**, *63*, 409-429.

20. Strydom, D. J.; Haylett, T.; Stead, R. H. *Biochemical and Biophysical Research Communications* **1977**, *79*, 932-938.

21. Ohshoma, Y.; Suzuki, Y. *Proceedings of the National Academy of Sciences, U. S. A.* **1977**, *74*, 5363-5367.

22. Manning, R. F.; Gage, L. P. *Journal of Biological Chemistry* **1978**, *253*, 2044-2052.

23. Komatsu, K. *Chemical and Structural Studies on Silk*; Komatsu, K., Ed.: Tokyo, 1985; Vol. I.

24. Lucas, F.; Shaw, J. T. B.; Smith, S. G. *Journal of Molecular Biology* **1960**, *2*, 339.

25. Gage, L. P.; Manning, R. F. *J. Biol. Chem.* **1980**, *255*, 9444-9450.

26. Garel, J. P.; Hentzen, D.; Schlegel, M.; Dirheimer, G. *Biochimie* **1976**, *58*, 1089-1100.

27. Hentzen, D.; Chevallier, A.; Garel, J. P. *Nature* **1981**, *290*, 267-269.

28. Kaplan, D. L.; Fossey, S. A.; Mello, C. M.; Arcidiacono, S.; Senecal, K.; Muller, W. S.; Stockwell, S.; Beckwitt, R.; Viney, C.; Kerkham, K. *Matls. Res. Soc. Bulletin* **1992**, *10*, 41-47.

29. Kaplan, D. L.; Adams, W. W.; Viney, C.; Farmer, B. *Silks: Materials Science and Biotechnology*; American Chemical Society: New York, 1994; Vol. 544.

30. Koga-ban, Y.; Suzuki, Y. *Development Growth and Differentiation* **1987**, *29*, 363 - 372.

31. Okamoto, H.; Ishiwaka, E.; Suzuki, Y. *Journal of Biological Chemistry* **1982**, *257*, 15192 - 15199.

32. Shibnev, V. A.; Chuvaeva, T. P.; Khalikov, S. K.; Poroshin, K. T. *Izv. Akad. Nauk SSSR , Ser Khim.* **1970**, *1970*, 2566-2570.

33. Suzuki, Y.; Gage, L. P.; Brown, D. D. *Journal of Molecular Biology* **1972**, *70*, 637-649.

34. Tsuijimoto, Y.; Suzuki, Y. *Cell* **1979**, *16*, 425-436.

35. Tsuijimoto, Y.; Suzuki, Y. *Cell* **1979**, *18*, 591-600.

36. Warwicker, J. O. *Journal of Molecular Biology* **1960**, *2*, 350.

37. Chevallier, A.; Garel, J. P. *Biochimie* **1979**, *61*, 245-262.

38. Gosline, J. M.; DeMont, M. E.; Denny, M. W. *Endeavour* **1986**, *10*, 37-43.

39. Magoshi, J.; Magoshi, Y.; Nakamura, S. *Polymer Communications* **1985**, *26*, 309.

40. Magoshi, J.; Kamiyama, S.; Nakamura, S. *Crystallization of Silk Fibroin Induced Under Various Conditions*; Magoshi, J.; Kamiyama, S.; Nakamura, S., Ed.: Tokyo, 1985; Vol. 1, pp 337.

41. Magoshi, J.; Mizuide, M.; Magoshi, Y.; Takahashi, K.; Kubo, M.; Nakamura, S. *Journal of Polymer Science, Polymer Physics Edition* **1979**, *17*, 515-520.

42. Magoshi, J.; Magoshi, Y. *Journal of Polymer Science, Polymer Physics Edition* **1977**, *15*, 1675-1683.

43. Mahoney, D. V.; Vezie, D. L.; Eby, R. K.; Adams, W. W.; Kaplan, D. *Aspects of the Morphology of Dragline Silk of Nephilia clavipes*; Mahoney, D. V.; Vezie, D. L.; Eby, R. K.; Adams, W. W.; Kaplan, D., Ed.; American Chemical Society: Washington D. C., 1994; Vol. 544, pp 196.

44. Palmer, J. M. *Journal of Morphology* **1985**, *186*, 195.

45. Yamaura, K.; Okumura, Y.; Ozaki, A.; Matsuzawa, S. *Journal of Macromolecular Science - Physics* **1985**.

46. Yamaura, K.; Okamura, Y.; Matsuzawa, S. *Journal of Macromolecular Science - Physics* **1982**, B(21., 49-69.

47. Willcox, P. J.; Gido, S. P.; Muller, W.; Kaplan, D. L. *Macromolecules* **1996**, *29*, 5106-5110.

48. Ambrose, E. J.; Bamford, C. H.; Eliiott, A.; Hanby, W. E. *Nature* **1951**, *167*, 264-265.

49. Kratky, O. *Monatshefte fur Chemie, Bd.* **1956**, *87*, 269-280.

50. Fraser, R. D. B.; MacRae, T. P.; Stewart, F. H. C. *Journal of Molecular Biology* **1966**, *19*, 580-582.

51. Takahashi, Y. *Crystal Structure of Silk of Bombyx mori*; Takahashi, Y., Ed.; American Chemical Society: Washington, D. C., 1994, pp 168-175.

52. Fraser, R. D. B.; MacRae, T. P.; Stewart, F. H. C.; Suzuki, E. *Journal of Molecular Biology* **1965**, *11*, 706-712.

53. Yoshimizu, H.; Asakura, T. *Journal of Applied Polymer Science* **1990**, *40*, 1745.

54. Ishida, M.; Asakura, T.; Yokoi, M.; Saito, H. *Macromolecules* **1990**, *23*, 88.

55. Bhat, N. V.; Ahirrao, S. M. *Journal of Polymer Science, Polymer Chemistry Edition* **1983**, *21*, 1273-1280.

56. Kratky, O.; Schauenstein, E.; Sekora, A. *Nature* **1950**, *165*, 319-320.

57. Demura, M.; Asakura, T. *J. Membrane Sci.* **1991**, *59*, 39-52.

58. Magoshi, J.; Nakamura, S. *Netsu Sokutei* **1987**, *14*, 58 - 69.

59. Saito, H.; Tabeta, R.; Asakura, T.; Iwanaga, Y.; Shoji, A.; Ozaki, T.; Ando, I. *Macromolecules* **1984**, *17*, 1405-1412.

60. Saito, H.; Iwanaga, Y.; Tabeta, R.; Narita, M. *Chemistry Letters, Japan* **1983**, 427-430.

61. Kricheldorf, H. R.; Muller, D.; Ziegler, K. *Polymer Bulletin* **1983**, *9*, 284-291.

62. Asakura, T.; Watanabe, Y.; Uchida, A.; Minagawa, H. *Macromolecules* **1984**, *17*, 1075-1081.

63. Asakura, T.; Watanabe, Y.; Itoh, T. *Macromolecules* **1984**, *17*, 2421-2426.

64. Asakura, T.; Kuzuhara, A.; Tabeta, R.; Saito, H. *Macromolecules* **1985**, *18*, 1841-1845.

65. Mercer, E. H.; Meyer, F. H. *Textile Research Journal* **1953**, *23*, 243 - 246.

66. Hirabayashi, K.; Arai, M. *Journal of Sericultural Science Japan* **1988**, *58*, 81 - 82.

67. Hirabayashi, K.; Tsukada, M.; Ishikawa, H.; Yasumura, S. *Sen-i Gakkaishi* **1974**, *30*, T459 - 464.

68. Tsukada, M.; Komoto, T.; Hirabayashi, K. *Sen-i Gakkaishi* **1983**, *39*, T227-232.

69. Tsukada, M.; Aoki, A. *Journal of Sericultural Science Japan* **1985**, *55*, 48-52.

70. Tsukada, M.; Nagura, M.; Ishikawa, H. *Molecular Conformation and Crystalline Structure in Model Compounds of Silk Sericin*; "Proceedings of the 7th International Wool and Textile Conference": Tokyo, 1985; Vol. 1, pp 383 - 388.

71. Kataoka, K. *Kobunshi Robunshu, Eng. Ed.* **1977**, *6*, 1 - 9.

72. Lotz, B.; Keith, H. D. *Journal of Molecular Biology* **1971**, *61*, 201-215.

73. Fossey, S. A.; Kaplan, D. L. *Molecular modeling studies on silk peptides*; Fossey, S. A.; Kaplan, D. L., Ed.; American Chemical Society: New York, 1994; Vol. 31, pp 1529-1541.

74. Oka, M.; Baba, Y.; Kagemoto, A.; Nakajima, A. *Polymer Journal* **1990**, *22*, 416-425.

75. Paterlini, M. G.; Freedman, T. B.; Nafie, L. A. *Biopolymers* **1986**, *25*, 1751-1765.

76. Tiffany, L.; Krimm, S. *Biopolymers* **1968**, *6*, 1379-1381.

77. Tiffany, L.; Krimm, S. *Biopolymers* **1969**, *8*, 347-359.

78. Traub, W.; Shmueli, U.; Suwalsky, M.; Yonath, A. *Some X-Ray Studies Concerning the Influence of Solvents on Polypeptide Structures*; Ramachandran, G. N., Ed.; Academic Press: New York, New York, 1967; Vol. 2, pp 785.

79. Traub, W. *Journal of Molecular Biology* **1969**, *43*, 479-485.

80. Rippon, W. B.; Walton, A. G. *Journal of the American Chemical Society* **1972**, *94*, 4319-4324.

81. Rippon, W. B.; Chen, H. H.; Anderson, J. M.; Walton, A. G. *Biopolymers* **1972**, *11*, 1411-1419.

82. Rippon, W. B.; Lowbridge, J.; Walton, A. G. *Biopolymers* **1977**, 1139-1151.

83. Brack, A.; Spach, G. *Biopolymers* **1972**, *11*, 563-586.

84. Loeb, G. I. *Journal of Colloid and Interface Science* **1969**, *31*, 572.

85. Loeb, G. I. *Journal of Colloid and Interface Science* **1968**, *26*, 236.

86. Weissbuch, I.; Berkovic, G.; Leiserowitz, L.; Lahav, M. *Journal of the American Chemical Society* **1990**, *112*, 5874 - 5875.

87. Malcolm, B. R. *Nature* **1962**, *4195*, 901.

88. Macritchie, F. *J. Coll. Int. Sci.* **1986**, *25*, 341-385.

89. Markin, V. S.; Volkov, A. G. *Adsorption Isotherms and the Structure of Oil/Water Interfaces*; Volkov, A. G.; Deamer, D. W., Ed.; CRC Press: Boca Raton, Florida, 1996, pp 63-75.

90. Markin, V. S.; Gugeshashvili, M. I.; Volkov, A. G.; Munger, G.; Leblanc, R. M. *Journal of Colloid and Interface Science* **1992**, *154*, 264-275.

91. Safinya, C. R. *Coll. Surf. A* **1997**, *128*, 183 - 195.

92. Malcolm, B. R. *Soc. Chem. Ind. London* **1965**, *19*, 102.

93. Malcolm, B. R. *Progress in Surface and Membrane Science* **1971**, *4*, 299.

94. Murray, B. S. *Coll. Surf. A* **1997**, *125*, 73-83.

95. Murray, B. S.; Nelson, P. V. *Langmuir* **1996**, *12*, 5973-5976.

96. Biridi, K. S. *Journal of Colloid and Interface Science* **1973**, *43*, 545.

97. Cheesman, D. F.; Davies, J. T. *Advan. Protein Chem.* **1954**, *9*, 439.

98. Jacuemain, D.; Wolf, S. G.; Leveiller, F.; Lahav, M.; Leiserowitz, L.; Deutsch, M.; Kjaer, K.; Als-Nielsen, J. *Journal of the American Chemical Society* **1990**, *112*, 7724 - 7736.

99. Magdassi, S.; Garti, N. *Surface Activity of Proteins*; Magdassi, S.; Garti, N., Ed.; Marcel Dekker: New York, 1991; Vol. 39, pp 289 -300.

100. Vroman, L. *Proteins in Blood Plasma at Interfaces*; Vroman, L., Ed.; Marcel Dekker: New York, 1991; Vol. 39, pp 137 - 152.

101. Wustneck, R.; Kragel, J.; Miller, R.; Wilde, P. J.; Clark, D. C. *Coll. Surf. A* **1996**, *114*, 255-265.

102. Zhang, W.; Gido, S. P.; Muller, W. S.; Fossey, S. A.; Kaplan, D. L. *Electron Microscopy Society of America, Proceedings* **1993**, 1216.

103. Valluzzi, R.; Gido, S.; Zhang, W.; Muller, W.; Kaplan, D. *Macromolecules* **1996**, *29*, 8606-8614.

104. Valluzzi, R.; Gido, S. P. *Biopolymers, in press* **1997**.

105. Shchipunov, Y. A. *Liquid/Liquid Interfaces and Self-Organized Assemblies of Lecithin*; Volkov, A. G.; Deamer, D. W., Ed.; CRC Press: Boca Raton, Florida, 1996, pp 295-315.

106. Miller, I. R. *Progress in Surface and Membrane Science* **1971**, *4*, 299.

107. Muller, W. S.; Samuelson, L. A.; Fossey, S. A.; Kaplan, D. L. *Langmuir* **1993**, *9*, 1857 - 1861.

108. Muller, W. S.; Samuelson, L. A.; Fossey, S. A.; Kaplan, D. *Formation and Properties of Silk Thin Films*; Kaplan, D.; Adams, W. W.; Farmer, B.; Viney, C., Ed.; American Chemical Society: Washington, D. C., 1994, pp 342-352.

109. Kanishi; Sakabe; Hiei *PI 16*; Hakone, Japan, 1985, pp 134.

110. David, W. I. F.; Matthewman, J. C. *Journal of Applied Crystallography* **1985**, *18*, 461 - 466.

111. Greaves, C. *Journal of Applied Crystallogreaphy* **1985**, *18*, 48 - 50.

112. Falls, A. H.; Wellinghoff, S. T.; Talmon, Y.; Thomas, E. L. *Journal of Materials Science* **1983**, *18*, 2752-2764.

113. Dowell, L. G.; Rinfret, A. P. *Nature* **1960**, *188*, 1144-1148.

114. Crick, A.; Rich, F. H. C. *Nature* **1955**, *176*, 780.

115. Schulz, G. E.; Schirmer, R. H. *Principles of Protein Structure*; Springer-Verlag New York Inc.: New York, 1979.

116. Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *Journal of Molecular Biology* **1963**, *7*, 95 - 99.

117. Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *175*, 593-595.

118. Rich, A.; Crick, F. H. C. *Nature* **1955**, *176*, 915-916.

119. Cowan, P. M.; McGavin, S. *Nature* **1955**, *178*, 501-502.

120. Lotz, B.; Keith, H. D. *Journal of Molecular Biology* **1971**, *61*, 195-200.

121. Ramachandran, G. N.; Sasisekharan, V.; Ramakrishnan, C. *Biochimica et Biophysica Acta* **1966**, *112*, 168-170.

122. Ramachandran, G. N.; Ramakrishnan, C.; Venkatachalam, C. M. *Structure of Polyglycine II with Direct and Inverted Chains*; in "Conformation of Biopolymers".; Ramachandran, G. N., Ed.; Academic Press: New York, New York, 1967; Vol. 2, pp 785.

123. Shibnev, V. A.; Khalikov, S. K.; M. P, F.; Poroshin, K. T. *Izv. Akad. Nauk SSSR , Ser Khim.* **1970**, *1970*, 2822-2823.

124. Tormo, J.; Puiggali, J.; Vives, J.; Fita, I.; Lloveras, J.; Bella, J.; Aymami, J.; Subirana, J. *Biopolymers* **1992**, *32*, 643-648.

125. Segal, D. M.; Traub, W.; Yonath, A. *Journal of Molecular Biology* **1969**, *43*, 519 - 527.

126. Yonath, A.; Traub, W. *Journal of Molecular Biology* **1969**, *43*, 461 - 477.

127. Atkins, T. A. *Crystal Structure by X-ray Diffraction*; Atkins, T. A., Ed.; Pergamon Press: Elmsford, NY, 1989; Vol. 1, pp 613 - 650.

128. IUCr *International Tables for Crystallography*; The Kynoch Press: Birmingham, England, 1962; Vol. III.

129. Taylor, R.; Kennard, O.; Versichel, W. *Acta Crystallographia* **1984**, *B40*, 280 - 288.

130. Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1994.

131. Baker, E. N.; Hubbard, R. E. *Progress in Biophysics and Molecular Biology* **1984**, *44*, 97 - 179.

132. Landau, E. M.; Popovitz-Biro, R.; Leiserowitz, L.; Lahav, M.; Sagiv, J. *Molecular Crystals and Liquid Crystals* **1986**, *134*, 323 - 335.

133. Landau, E. M.; Wolf, S. G.; Levanon, M.; Leiserowitz, L.; Lahav, M.; Sagiv, J. *Journal of the American Chemical Society* **1989**, *111*, 1436 - 1445.

134. Fossey, S. A.; Nemethy, G.; Gibson, K. D.; Scheraga, H. A. *Biopolymers* **1991**, *31*, 1529-1541.

135. Traub, W.; Yonath, A.; Segal, D. M. *Nature* **1969**, *221*, 914 - 917.

136. Padden, F. J.; Keith, H. D. *Journal of Applied Physics* **1965**, *36*, 2987.

137. Segal, D. M.; Traub, W. *Journal of Molecular Biology* **1969**, *43*, 487 - 496.

138. Happey, F.; Hyde, A. J.; Manogue, B. *Biopolymers* **1967**, *5*, 749-756.

139. Lahav, M.; Leiserowitz, L. *Journal of Physics D: Applied Physics* **1993**, *26*, B22 - B31.

140. Riou, S.; Hsu, S. L. *unpublished data* **1996**

141. Buren, A. R. v.; Marrink, S.-J.; Berendsen, H. J. C. *Journal of Physical Chemistry* **1993**, *97*, 9206-9212.

142. Carpenter, I. L.; Hehre, W. J. *Journal of Physical Chemistry* **1990**, *94*, 531-536.

143. Michael, D.; Benjamin, I. *Journal of Physical Chemistry* **1995**, *99*, 1530-1536.

144. Kyte, J.; Doolittle, R. F. *journal of Molecular Biology* **1982**, *157*, 105-132.

145. Poole, P. L.; Finney, J. L. *Bioploymers* **1983**, *22*, 255-260.

146. Thomas, E. L.; Wood, B. A. *Far. Disc. Chem. Soc.* **1985**, *79*, 229.

147. Wood, B. A.; Thomas, E. L. *Nature* **1986**, *324*, 655.

148. Giraud-Guille, M. M. *Calcif. Tissue Int.* **1988**, *42*, 167-180.

149. Giraud-Guille, M. M. *Mol. Cryst. Liq. Cryst.* **1987**, *153*, 15-30.

150. Bairati, A.; Garrone, R. *Biology of Invertebrate and Lower Vertebrate Collagens*; Bairati, A.; Garrone, R., Ed.; Plenum Publishing Company: New York, NY, 1985; Vol. 93.

151. Neville, A. C. *Biology of Fibrous Composites: Development Beyond the Cell Membrane*; Cambridge University Press: Cambridge, U.K., 1993.

152. Gaill, F. *Clloque de Physique* **1990**, *51*, c7-169 - c7-181.

153. Bunning, T. J.; Vezie, D. L.; Lloyd, P. F.; Haaland, P. D.; Thomas, E. L.; Adams, W. W. *Liq. Cryst.* **1994**, *16*, 769.

154. Ruggeri, A.; Motta, P. M. *Ultrastructure of the Connective Tissue Matrix*; Ruggeri, A.; Motta, P. M., Ed.; Martinus Nijhoff Publishers: Boston, MA, 1984; Vol. 3.

155. Rudall, K. M. *Comparative Biology and Biochemistry of Collagen*; Rudall, K. M., Ed.; Academic Press: London, 1968; Vol. 2, Part A, pp 83 - 137.

156. Miller, A. *Prog. Surf. Memb. Sci.* **1975**, *4*, p.299.

157. Mayne, R.; Burgeson, R. E. *Structure and Function of Collagen Types*; Mayne, R.; Burgeson, R. E., Ed.; Academic Press, Inc.: Orlando, Florida, 1987.

158. Ludtke, S. J.; He, K.; Heller, W. T.; Harroun, T. A.; Yang, L.; Huang, H. W. *Biochemistry* **1996**, *35*, 13723-13728.

159. Kucharz, E. J. *The Collagens:Biochemistry and Pathophysiology*; Springer-Verlag: Berlin, 1992.

160. Jones, E. Y.; Miller, A. *J. Mol. Biol.* **1991**, *218*, 209 - 219.

161. He, S. J., *unpublished data,* **1997**.