

## SELF-ORGANIZATION (ASSEMBLY) IN BIOSYNTHESIS OF SILK FIBERS - A HIERARCHICAL PROBLEM

DAVID L. KAPLAN\*, STEPHEN FOSSEY\*, CHRISTOPHER VINEY\*\*, WAYNE MULLER\*

\*Biotechnology Division, U.S. Army Natick Research Center, Natick, Massachusetts, 01760, USA

\*\*Center for Bioengineering, University of Washington, Seattle, Washington, USA

### ABSTRACT

In natural systems, structural macromolecules undergo prescribed recognition and assembly steps during synthesis and processing. These associations lead to more complex assemblies that exhibit useful multifunctional properties. Many of these processes are not well understood. Some aspects of these processes are presented using the fibrous protein polymer silk as an example. Issues such as polymer chain biosynthesis, chain interactions, processing into fibrils, and complex engineering into supra-assemblies are addressed and biochemical, spectroscopic and modeling studies are reviewed. Genetic level controls of chain composition, crystalline/amorphous domain distribution, chain aggregation, chain registry, silk I-silk II phase transitions, nematic liquid crystalline phase, loss of water, global molecular alignment, and solution spinning are some of the characteristics of this biological system that are addressed. Although some information is available at the molecular and macro-scale levels, a key issue is the paucity of information at the meso-scale level to fully understand the role of structural hierarchy in the silk fiber assembly process.

### INTRODUCTION

Efficient assembly processes inherent in biological systems are required for the formation of complex multifunctional materials. There are an impressive number of structures and systems that could be examined for insights into these processes, from complex membrane channel proteins and studies of microtubule assembly to flagellar motors. We will focus on assembly processes for biological fibers. Although fiber assembly would appear to be a simpler process when compared with some of the more complex systems mentioned above, very little detail is still understood. The possible role of hierarchy in this process will be described, where "hierarchy" [1] refers to structures with: (1) different scales of organization (e.g. molecular, nano-, micro-, meso-, macro-), (2) characterized by specific interactions between the different components, and (3) characterized by complex architecture to achieve desired functional properties.

Many different biological fibers could be examined for insight into these processes, including structural proteins (e.g., keratin, collagen, silk, troponin, elastin, actin) and structural polysaccharides (e.g., cellulose, chitin). We will focus on silks because of the structural hierarchy exhibited from the genetic level to the web or cocoon level, the unusual mechanical properties (Table I), the correlation between structure and function evident even by amino acid composition analysis of different silks from one species of spider (Table II), and the fact that this is the most common spun protein fiber in biological systems. Key to these observations is that the functional properties derive both from the primary structure and the processing conditions used to convert the polypeptide chains into fibers with a high degree of molecular order. Of major significance is that the mechanical properties of these fibers are achieved with processing conditions that appear relatively mundane (room temperature, aqueous solutions, and minimal draw), yet global alignment of the polymer chains is achieved and exceptional mechanical properties are realized [2]. To achieve comparable properties and alignment with synthetic polymers, more extreme processing conditions of temperature, solvent and draw are required. Additional significant features from a polymer perspective are the monodisperse, isotactic and stereoregular nature of the polymer.

An understanding of biological fiber assembly processes offers lessons that should have utility for both biological and synthetic process needs. Only through the detailed elucidation of

the biological processes can these ideas be applied in full. Some specific thoughts on the benefits to be gained from the study of biological systems related to fiber assembly include: (1) correlations of structure at the genetic level to function at the fiber level and approaches to tailor this relationship, (2) temporal factors of regulation where close coupling occurs between synthesis, processing and assembly to avoid chain entanglement and to control solubility, (3) identification of opportunities to improve upon mechanical properties of biological fibers that are presumably optimized within the bounds of evolution and survival (*in vivo*) but perhaps not optimized outside of these limits (*in vitro*), (4) cellular and genetic regulation for on-demand synthesis of monomers for polymerization instead of large scale storage of inventory, (5) control at all length scales (molecular to macro) to provide global control of final structure, (6) potential to develop fiber-assembly processes that operate at room temperature with water as the solvent, (7) solution spinning of fibers to obtain global alignment with minimal draw, (8) formation of fiber products with enhanced purity due to the absence of catalysts used in synthetic polymer processing, (9) incorporation of monodisperse and stereoregular polymer chains for greater homogeneity and predictability of properties in the spun product, and (10) formation of fibers with diameters as small as 0.01  $\mu\text{m}$  (some spiders), much finer than industrial-spun fibers.

The focus will be on silkworm cocoon silk (*Bombyx mori*) and spider silk (dragline from *Nephila clavipes*). The silkworm produces one type of silk (cocoon) at one stage in its lifecycle (fifth instar) which contains the proteins fibroin (structural) and sericin (family of glue-like proteins). The more evolutionary advanced orb-weaving spiders often produce many different silks, some throughout their lifecycle, with each silk originating from a different set of glands within the single species of spider and consisting primarily of a single protein. The total number of different silks produced by a given species of spider varies, with nine or more reported [3]. In Table II, the correlation between amino acid composition and silk function is illustrated. The major ampullate gland produces the structural silks for the orb frame, radii and dragline (the safety line and strongest of all the silk fibers); the flagelliform gland produces the viscid silk for prey capture; the aggregate gland produces an adhesive silk; the minor ampullate gland produces support fibers for the orb web; the cylindrical gland produces the cocoon silk; the aciniform gland produces silk for wrapping captured prey; and the piriform gland produces attachment silks to couple to environmental substrates [4]. Within a given species of spider, the general trend is that the proportion of short side chain amino acids increases as the strength requirements for the different silks increase, and conversely, the percentage of charged amino acids decreases. However, this trend is not universal, since spider dragline silk is stronger than silkworm cocoon silk (Table I), yet contains a lower percentage of short side chain amino acids. Silkworm and major ampullate glands generally contain three relatively distinct regions, the posterior region where fibroin is synthesized by the epithelial cells lining the gland, the middle region where the protein is stored and in the case of the silkworm the sericin is synthesized, and the anterior region which leads to the spinneret where the protein is spun into a fiber.

Table I. Mechanical Properties of Silks and Other Fibers [see specific references in 5].

Fiber	Elongation (%)	Modulus (N/m <sup>2</sup> )	Strength (N/m <sup>2</sup> )	Energy to Break (J/kg)
Spider Silk (dragline of <i>Nephila clavipes</i> )	10-32	1-30X10 <sup>9</sup>	3-18X10 <sup>8</sup>	3-10X10 <sup>4</sup>
Silkworm silk (cocoon of <i>Bombyx mori</i> )	15-35	5X10 <sup>9</sup>	6X10 <sup>8</sup>	7X10 <sup>4</sup>
Nylon	18-26	3X10 <sup>9</sup>	5X10 <sup>8</sup>	8X10 <sup>4</sup>
Cotton	6-7	6-11X10 <sup>9</sup>	3-7X10 <sup>8</sup>	5-15X10 <sup>3</sup>
Kevlar	4	1X10 <sup>11</sup>	4X10 <sup>9</sup>	3X10 <sup>4</sup>
Steel	8	2X10 <sup>11</sup>	2X10 <sup>9</sup>	2X10 <sup>3</sup>

<sup>1</sup>Note: compiled from literature values. Fiber diameters, water content, relative humidity and other conditions vary so comparisons should be made with these caveats in mind.

Table II. Relationship between amino acid composition and silk function [calculated from 6].

Silk Gland Type & Silk Function	Predominant Amino Acids (>10%)	Short Side Chain Amino Acids (%) <sup>1</sup>	Polar Amino Acids (%) <sup>2</sup>	Charged Amino Acid (%) <sup>3</sup>
Large Ampullate Gland - dragline/frame silks	Gly (37.2), Ala (17.6) Pro (15.8), Glu (11.5)	62.3	25.9	7.0
Small Ampullate Gland - frame/scaffolding silks	Gly (42.8), Ala (36.8)	84.6	16.7	4.2
Flagelliform Gland - sticky spiral silks	Gly (44.2), Pro (20.5)	55.5	16.9	6.1
Cylindrical(tubuliform) - egg cocoon silk	Ser (27.6), Ala (24.4)	60.7	49.7	11.7
Aciniform Gland - swathing silk	Ser (15.0), Gly (13.9) Ala (11.3), Leu (10.1)	40.3	47.2	15.0
Piriform Gland - attachment silk	Ser (14.8), Asp (10.5) Glu (10.4)	52.2	57.9	28.7
Aggregate Gland - sticky threads/glue silk	Gly (14.5), Pro (10.8)	27.5	48.7	23.4

<sup>1</sup>gly, ala, ser; <sup>2</sup>asp, thr, ser, glu, tyr, lys, his, arg; <sup>3</sup>asp, glu, lys, his, arg

We will review the state of limited knowledge about the key steps involved in the assembly of silk fibers and the role of hierarchy in this process, including: (1) biosynthesis and the genetic level controls over structure and assembly, (2) processing steps related to extracellular transport, protein folding and initial associations between polymer chains to form fibrils, (3) assembly processes involved in fiber formation, and (4) processing into supra-assemblies or highly engineered structures. Related information on silk chemistry, properties and genetics has recently been reviewed [5].

#### POLYMER CHAIN BIOSYNTHESIS

Polypeptide synthesis is initiated in epithelial cells lining the lumen of the silk glands in both silkworm and spiders. The protein polymer chains are encoded by highly repetitive genes containing domains responsible for crystalline and amorphous peptide segments. The complete gene structure for any silk encoding gene is unknown. Only partial sequencing of silkworm genes has been reported, with the predominant information available for the 5'-end and 5'flanking regions, and little information on the 3'-end and overall core (repetitive) structure. Only about 2 kb of the entire 21 kb genomic clone (including 16 kb fibroin gene plus the 5' and 3' non-coding flanking regions) for silkworm cocoon silk has been sequenced [7]. The crystalline domains have been estimated to be 1 kb to 2 kb in size, with periodic interspersed amorphous domains of about 220 base pairs [8]. The amorphous domains in spider silk were postulated to be about 1.4 kb [4]. The frequency, size and distribution of the crystalline and amorphous domains will play a major role in the protein polymer chain interactions and the resulting mechanical properties of fibers spun from these silk proteins. About 2.4 kb of cDNA sequence has been reported from the 3' end of a spider dragline silk clone containing a 34 amino acid crystalline repeat which is not highly conserved [9], while other clones are still being characterized [10,11]. The size of the crystalline domains containing this repeat are unknown. The ratio of introns (noncoding) to exons (coding) for the silk genes appears low when compared, for example, with the structural protein elastin from human and bovine sources, where a 15:1 ratio of intron to exon occurs in the 3' end of genomic clones. The exons, which vary in size and are less than 70 base pairs, are interrupted with long intron domains which translates to an elastin gene size over 40 kb and an mRNA of 3.5 kb [12], vs. about an 16 kb gene and 11 kb mRNA for silkworm and spider dragline silk.

The silkworm silk gene is present only as a single copy per haploid complement, despite the high level of protein expression required during cocoon formation. Strong transcriptional control over the fibroin gene and a very stable mRNA encoding the protein provide the regulation needed to produce protein at high levels during cocoon spinning in the silkworm. Preliminary data indicate similar genetic level controls for the dragline silk gene in the spider (unpublished observations).

Upon translation, the main protein polymer chains in both silkworm cocoon and spider dragline silks are over 300 kd [13,14]. Silkworm cocoon silk also contains a 25 kd protein and a family of sericin proteins which range from 20 kd to 310 kd. There is evidence for a single disulfide linkage between the small and large peptide chains in silkworm cocoon silk [15]. Little evidence for a multiplicity of proteins in a specific type of spider silk has been reported, although confounding results occur when proteins from different glands are not carefully segregated during excision or controlled silking, or proteases are not inhibited during protein preparation. The crystalline and amorphous domains within the protein are reflected by amino acid composition and based on genetic level encoding regions. The crystalline domains are characterized by a high percentage of the short side chain amino acids glycine, alanine and serine (3:2:1 ratio in silkworm cocoon silk), with the total percentage of short side chain amino acids in *B. mori* silkworm cocoon silk over 80%, while in the spider *N. clavipes* dragline silk the corresponding value is around 62% [16]. X-ray diffraction data indicate a degree of crystallinity between 62% and 66% for *B. mori* cocoon silk fibroin. Chymotrypsin digestion of silkworm silk leads to a precipitate which comprises 60% by mass of the total protein. The precipitate represents the crystalline domains which remain uncut and insoluble after exposure to the enzyme.

From protein analysis, the consensus crystalline domain of *B. mori* cocoon silk consists of the 59mer repeat: GAGAGSGAAG[SG(AG)<sub>n</sub>]<sub>8</sub>Y, where n is 2 [17]. Comparative data for spider silks is only beginning to be generated, although repeats similar to the silkworm sequence, and related smaller repeats such as GQGAG, AAVAQAQAGAGA, GAGHGA, GYGPG, GAGRG, GAGQG, and GYGGLG have been reported [9, unpublished observations]. For the amorphous domains which contain a higher percentage of bulkier amino acid side chains, more limited data is available, with silkworm silk sequences: GAGAGAGY, G(G<sub>3</sub>A<sub>2</sub>V)Y, GAGY, G(G<sub>2</sub>AD)Y, GVGY, GAGY, SGY, GPY, and others [18-20], and related sequences for spider silks [unpublished observations]. GAGAGAGY, G(G<sub>3</sub>A<sub>2</sub>V)Y, and GAGY accounted for 14.4%, 14.1% and 11.3% by weight of the soluble fraction, and 5.8%, 5.6%, and 4.5% of the total silk, respectively, from *B. mori*. [summarized in 21].

### POLYMER CHAIN INTERACTIONS

Chain interactions and registry are presumed to initiate intracellularly, prior to export into the lumen of the gland. Associations between the crystalline domains could initiate the aggregation process as the polymer chains are exported from the cell. This premise is supported with observations of the self-assembly of synthetic peptides modeled after the crystalline domains, which readily aggregate if the Gly-Ala repeat is >8, indicating a minimum domain size of 12.87 Å [unpublished observations]. The size and distribution of these recognition elements will provide associations through hydrophobic interactions due to the predominance of short side chain amino acids and hydrogen bonding from the peptide bonds. The silk I conformation, a soluble form of silk, exists at this intermediate stage in the lumen of the posterior region of the gland. This contrasts with the silk II conformation, which is present in the insoluble silk fiber.

The conformation of fibroin does not appear to depend on its concentration in water, provided the solution is not subjected to shear [21]. In addition, a hydropathy search indicates very few large domain hydrophilic or hydrophobic residues in representative sequences of *N. clavipes* dragline and *B. mori* cocoon silk [22,23]. This finding would support the contention that silk protein conformations are not especially sensitive to the polymer concentration in water. The conformation of silk within the gland has been probed by <sup>13</sup>C-

NMR and shown to be similar to that of silk I [24]. Additionally, although the statistical coil forms and silk I can be distinguished by X-ray diffraction they cannot be distinguished by IR spectroscopy and it is difficult by  $^{13}\text{C}$ -NMR since the chemical shifts are the same [25]. The  $^{13}\text{C}$ -NMR data show a line broadening in the statistical coil form as compared with that of silk I. These data suggest that the statistical coil form adopts a conformation distributed around the silk I conformation and the chains do not adopt a regular crystalline packing [24]. The inability to make a distinction between the silk I peaks and statistical coil peaks makes it impossible to use  $^{13}\text{C}$ -NMR to define the fraction of ordered silk I within the silk gland. However, it does imply that all the protein which would eventually adopt the silk II crystalline form may already be in a silk I form but the degree of organization has not yet been determined. NMR data suggest chains have the same conformation over a concentration range of 0.3 to 14.5 g/l, which is the same conformation as silk I [24]. Domains responsible for this conformation extend over about 300 residues along the chain.

Epithelial cells lining the posterior region of the silk gland have a radial microtubule system and a circular microtubule-microfilament system which provide for intracellular transport and secretion of fibroin in the form of secretory granules [26]. The secretory granules of fibroin originate in the golgi apparatus in association with the endoplasmic reticulum. Electron microscopy indicates that the microtubules run from the basal to luminal membranes in the cytoplasm and fibroin "globules" have been observed associated with the microtubules. The polypeptide chain associations, concentrations and conformations in these granules are unknown. Fibroin globules also store calcium which is released upon the exocytosis of secretory granules of fibroin in the luminal plasma membrane. Different size pores on the cuticular membrane have been observed along with spherical fibroin masses 1  $\mu\text{m}$  to 3  $\mu\text{m}$  in diameter in the lumen of the gland [27]. Once in the lumen, additional interactions between polymer chains initiate the formation of the characteristic beta sheet structure. Silks can adopt a variety of secondary structures, including alpha helices, beta sheets and cross-beta sheets. The beta sheet structure is characteristic of the silkworm cocoon silk and spider dragline silk discussed here. The polymer chains run parallel to the fiber axis and run antiparallel with respect to each other, with hydrogen bonds between chains and hydrophobic interactions between overlying sheets due to the predominance of short side chain amino acids [28].

Details of the conformation of silk at this stage (silk I), have been difficult to assess experimentally because a variety of factors cause the premature conversion of the metastable silk I to the stable silk II conformation. Energy minimization studies on Gly-Ala repeats have been carried out to determine the relative stabilities of the different conformations in these stacked sheets (Table III) and the relative contribution of serine in the Gly-Ala repeat [29,30]. The serine contributes solubility to the uncrystallized silk and may contribute stability in the crystal by forming a hydrogen bond between the serine side chain and the carbonyl oxygen of the adjacent chain in both silk I and silk II conformations (unpublished observations). Hydrogen bond formation is supported by NMR experiments which show that on the NMR time scale the serine side chain is not rotating freely, while the methyl hydrogens of the alanine residues show rotational averaging [25].

Table III. Relative stabilities of crystals of model polypeptides based on empirical conformational energy program for peptides (ECEPP).

Peptide Sequence	Conformation	Total ECEPP Energy	
		IR <sup>1</sup>	OR <sup>2</sup>
Ala/Gly	Beta Sheet (extended)	-481	-299
Ala/Gly	Silk I form (alternating L/R 3-fold helix)	-368	-394

<sup>1</sup>IR = In register sheets; <sup>2</sup>OR = Out of register sheets

Conformational energy calculations were carried out on representative model polypeptides for the crystalline domains of *B. mori* silk. The  $\text{CH}_3\text{CO-AGAGAG-NHCH}_3$  repeat was used to construct five stranded hydrogen bonded sheets which were stacked three high to model three dimensional crystals. Since the strands in these sheets are two-fold helices the side chains of every other residue along the chain point to the same side of the sheet. Chains in register have all alanine side chains on the same side of the plane, while chains out of register have identical faces leaving only one orientation for stacking the sheets. Based on these conformational energy calculations Fossey *et al.* [29,30] have proposed an orthorhombic unit cell containing out of register chains as the silk I structure. Recently, it has been suggested that based on an analysis of X-ray intensities that the silk II form is actually composed of out of register chains (antipolar in their nomenclature) [31]. Fossey *et al.* [29,30] have found based on conformational energy calculations that this is indeed the lowest energy conformation for single sheets; however for stacked sheets the additional stabilization which is obtained from the close methyl-methyl interactions makes the in register [polar in the nomenclature of 31] a more stable conformation.

The orthorhombic unit cell for silk I has dimensions  $a = 8.9 \text{ \AA}$  (perpendicular to the chain axis in the hydrogen bonded direction),  $b = 11.3 \text{ \AA}$  (two-times the intersheet spacing), and  $c = 6.46 \text{ \AA}$  (chain direction). The proposed model agrees with X-ray diffraction data and density calculations. An alternating left and right handed 3 by 1 helix was found with hydrogen bonds  $= 1.99 \text{ \AA}$  in length. This compares with silk II, where the unit cell dimensions are  $a = 9.40 \text{ \AA}$  (interchain distance),  $b = 6.97 \text{ \AA}$  (fiber axis distance), and  $c = 9.20 \text{ \AA}$  (intersheet distance) [28]. Silk I may be a hydration stabilized conformation, since it is found when the contents of the silk gland are allowed to dry undisturbed.

### PROCESSING INTO FIBRILS

Little evidence is available to substantiate a distinct fibrillar microstructure in silk. Kratky [32] reported fibroin rod-like particles of approximately  $60 \text{ \AA} \times 90 \text{ \AA}$  in gels over a 10% to 20% range. The  $60 \text{ \AA}$  dimension is probably along the hydrogen bond axis and  $90 \text{ \AA}$  along the chain axis. These dimensions would accommodate 13 chains along the hydrogen bond direction and 28 residues along the chain axis based on the silk I model of Fossey *et al.* [29].

Fraser and MacRae [18] calculated, based on the half width of equatorial X-ray reflections for silk II, that the crystallites were  $59 \text{ \AA}$ , parallel to the hydrogen bond direction, and  $22 \text{ \AA}$  across some small number of sheets (a small number of sheets is consistent with the broad 001 reflections seen). Assuming that these two particles each represent a crystalline domain, the crystalline domains in silk II are  $59 \text{ \AA} \times 97 \text{ \AA} \times 22 \text{ \AA}$  or approximately  $125.9 \times 10^3 \text{ \AA}^3$  (1679 residues), implying that five or six of the crystalline chain segments of one or more likely several fibroin chains of Gage and Manning [8] are associated into a crystalline domain. It should be noted that Kratky [32] found the particle size changed from  $60 \text{ \AA} \times 90 \text{ \AA}$  to  $60 \text{ \AA} \times 180 \text{ \AA}$  when the gel was treated with solvent which might be expected to produce silk II. The reason for this change is not clear. Kratky [32] reports that aggregates in the gland of a similar size to the renatured gels are found as well as larger cylindrical particles of radius  $103 \text{ \AA}$  by a length of  $618 \text{ \AA}$ . It should be noted that these data only involve a single experiment. Flow conditions will require that the long axis of the crystal be in the fiber axis direction if these crystallites behave as a solid in the flow field.

Once the silk polypeptides have been synthesized and exported from the epithelial cells lining the lumen of the anterior silk gland, and subsequent to initial chain recognition and interaction, fibril formation begins during processing of the soluble silk (silk I) through the gland up to the spinning step (Figure 1). During this process at least four key events occur: (1) conformational change from silk I to silk II, (2) formation of a nematic liquid crystalline phase, (3) loss of water, and (4) shear.

First, the conformational shift has recently been described in detail [29]. In the silk II conformation, the glycine side chains project out from the same side of the sheet and the alanine and serine side chains from the other side of the same sheet. The sheets stack with the

glycine faces together and the alanine/serine faces together. The processing changes associated with the conversion from the silk I conformation in the gland to the silk II form involve only a small interchain energy barrier [29] but requires an 180 degree rotation of half the chains. The silk I to silk II conversion has been reported by a number of other investigators and occurs quite readily, suggesting that the overall energy barrier is not excessively large. The silk I-silk II transition can also be induced with heat [32]. The crystal associated with the silk II conformation is more tightly packed and excludes water molecules. The resulting fiber is no longer water soluble and is also very resistant to proteolytic digestion. Apart from the natural process, the conversion of the metastable silk I to the silk II conformation has been demonstrated experimentally through mechanical agitation, exposure to hydrophilic organic solvents such as methanol or acetone, exposure to electrical fields, and changes in temperature.

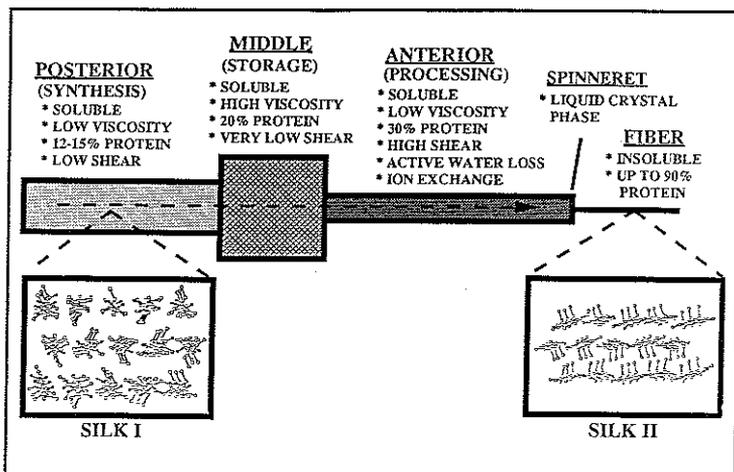


Figure 1. Diagrammatic representation of processing steps within a generalized silk gland.

The collagen triple helix, with its characteristic Gly-X-Y repeat, is perhaps the most well characterized structural protein in terms of fibril and fiber assembly [33]. Of interest is that reconstituted collagen fibrils do not provide the same packing as native material [34]. This indicates that the lateral packing process may be controlled by other factors than just self-assembly which controls axial packing and registry. For silks, the lateral control may derive from the spinneret in a physical process, while the axial control is also through self-assembly but based on associations and registry derived from the crystalline domains. Control of fiber shape and length are not clearly determined by the physical spinning process, since different geometries are reported for different silks and correlation to spinneret morphology has not been documented.

Wool filaments, formed from keratin with a high percentage of disulfide bonds, derive from the interaction of type I and type II intermediate filament proteins which exhibit high specificity [35]. Heterodimers between the intermediate filament proteins first form a coiled coil assembly followed by tetramer formation to produce a protofibril and then the final wool fiber.

The relative understanding of silk fibril assembly is poor when compared with collagen and keratin. The registry between chains and sheets presumably derives from recognition between crystalline domains. The recognition process may be influenced by the Ala-Ala repeat which

would allow the chains to remain in register and the tyrosine residue which it has been suggested cannot be accommodated within the crystalline domains [36]. This process is initiated in the posterior region of the silk gland and relates to the size and distribution of these domains. This interaction results in the optimization of chain orientation (shear alignment during processing, antiparallel vs. parallel, global alignment), control of crystallization, control of solubility (to avoid premature precipitation in the gland by maintenance of the silk I conformation up to the spinning step), and reduces chain entanglement (because initial associations occur early, close coupling of biosynthesis and processing occur at this stage).

It has been suggested that the disulfide bond between the light and heavy fibroin chains in silkworm cocoon silk of *B. mori*, along with hydrophobic interactions, may serve to prevent premature precipitation of the protein while in the gland by interfering with secondary structure formation [36]. This factor, coupled with the control of secondary structure (e.g., maintenance of a silk I conformation) are possible keys. Thus, these interchain interactions play a critical role in solubilization of the protein during synthesis, export, transport and secretion until the final spinning process.

Second, polarized light microscopy was used to identify characteristic microstructures (specifically, disclinations or local discontinuities in molecular orientation) in silk exudates from the silkworm and spider, establishing a role for a nematic liquid crystalline phase during processing [2,38]. In addition, hand-drawn silk fibers exhibited a banded microstructure perpendicular to the fiber axis, while native silk fibers show global molecular alignment. Of interest is that the band size was ten-fold larger in the hand-drawn silkworm silk vs. the hand-drawn spider silk. A liquid crystalline phase had been postulated because of the tensile properties of the fibers (high strength, stiffness, and toughness), optical birefringence, low viscosity for a high protein concentration in the gland (about 30%) which is still spinnable, molecular order exhibited for some polypeptides in solution, and minimal draw imparted by the organism. The nature of the rodlike entities responsible for the nematic liquid crystalline behavior of silk secretions is uncertain, although alpha-helices do not appear to be the structural unit involved [22,23]. Extension rates for glandular silk of 500 mm/min (at 35°C) or greater result in a conformational shift to the beta sheet and the appearance of birefringence [39].

Third, the loss of water during processing has been reported, since the concentration of protein changes from 12% to 14% in the posterior region of the silkworm silk gland where synthesis and transport occur, to 30% in the middle gland, and nearly dry in the final fiber product. In addition, from modeling studies, a volume reduction of 7% can be predicted upon conversion of silk I to silk II, which would exclude water from the crystalline domains [unpublished observations]. Absorption of water from the soluble silk during transit down the major ampullate gland duct was demonstrated, along with an exchange of sodium and potassium ions [40].

Fourth, high shear rates have been reported at 2 to 400 sec<sup>-1</sup> at 1.0 cm/sec spinning rates in the anterior narrow region of the silkworm gland which is 0.05 mm to 0.3 mm in diameter [39]. In the silkworm, crystallinity of silk correlates positively with shear rate and rate of draw, and negatively with the diameter of spinneret [39].

The result of these processes is the conversion of a soluble protein at low concentration to an insoluble protein fiber exhibiting global molecular alignment. These changes occur under ambient conditions, without external heating, involve the active transport of water and spinning from an aqueous solution, occur with minimal draw by the organism on the spun fiber, and take advantage of physiological control of the spinneret to effect fiber diameter.

## PROCESSING INTO FIBERS

Key elements in processing silk polypeptides and fibrils into silk fiber include: (1) control at the spinneret, (2) solution (water) spinning process with active transport and loss of water (addressed above), and (3) minimal fiber draw. Spider silk fibers generally vary in diameter

from 1  $\mu\text{m}$  to 5  $\mu\text{m}$  or larger depending on the species of spider and the glandular source of silk, although some silks are considerably finer in diameter (0.01  $\mu\text{m}$ ) [41]. Silkworm cocoon silk fibers range from 10  $\mu\text{m}$  to 25  $\mu\text{m}$  in diameter. The spinneret can be controlled to change fiber diameter. Some fiber draw occurs due to the characteristic figure-8 head movement of the silkworm and the action of the spider's legs during web formation. The spinneret serves as a press and a fulcrum for drawing silk by the silkworm [39].

At this stage of assembly and processing, key conformational changes associated with the crystalline domains are used to effect major shifts in physical and mechanical properties of the silk fibers formed. This shift is coupled with the nematic liquid crystalline phase used to organize the polymer chains during the spinning step to gain suitable alignment without the benefit of large external draw. During these changes the appropriate solubility for the fibrils must be maintained (silk I), since premature conversion to the insoluble silk II form would be catastrophic, essentially clogging the spinning apparatus. In the silkworm a role for a single disulfide bond in the process has been suggested [36]. The principal benefit of liquid crystallinity appears to be minimization of the energy required to expel the protein secretion through the duct and spinneret, due to the fact that the viscosity of a liquid crystalline polymer solution can be several orders of magnitude below that of a conventional polymer solution at the same concentration [42]. Since liquid crystalline phases are also formed by *N. clavipes* cocoon and capture silks, which form fibers significantly less strong and stiff than dragline, molecular alignment for improved mechanical properties does not appear to be the major driving force in this process. The improved ease of processing appears to be the principal benefit of the liquid crystalline phase.

Properties of the fibers formed by these processes are impressive (Table I). High strength, stiffness, toughness, resistance to water, resistance to proteolytic enzymes, and birefringence characterize the fibers. Silk decomposes at around 149°C.

#### COMPLEX ENGINEERING INTO SUPRA-ASSEMBLES

The single silk fibers spun from pairs of glands are often organized into more complex hierarchical structures such as the cocoon or the orb web. In terms of scales, this moves into the centimeter to meter range when dealing with these supra-assemblies (Table IV). This process involves coordination and regulation of multiple functions and activities which are outside of the scope of this paper. The complexity of engineering such composites and highly evolved structures that can provide a protective barrier against environmental hazards during molting (silkworm), or the alignment, fastening, and interaction of many different fibers and materials to form a highly engineered orb web structure, represent orders of magnitude increases in complexity that are not well understood. This final level of hierarchy is primarily behavioral/neurological, once all the requisite materials have been synthesized and processed. Therefore, although falling under the concept of hierarchies it addresses a completely different scientific question. Additional issues of silk recycling and degradation are interesting questions, since common proteolytic enzymes do not readily degrade the silk structure, however, some spiders are able to recycle webs on a daily basis [43] as demonstrated with radioactive tracers [44].

The assembly processes for silk fibers reviewed here provide preliminary support to satisfy the definition of hierarchical structures, since silks exhibit different levels of structural organization (molecular to macroscopic), specific interactions (crystalline domains, hydrogen bonding, hydrophobic interactions), and complex architecture, to achieve complex multifunctional performance. Clearly there is still much more that we do not understand, particularly at the mesoscopic level, but with the more recent application of the tools of genetic engineering to these polypeptides the knowledge base should expand rapidly.

Table IV. Generalized structural hierarchy in silk fibers.

Scale	Material	Size, Volume <sup>1</sup> , Length <sup>2</sup>	Location
<u>Molecular</u> (polymer chain)	protein	~350kd; ~5Kaa; 375,000Å <sup>3</sup> ; 72.1Å	Epithelial Cells
	crystalline domain [derived from 17]	59aa, 4,425Å <sup>3</sup> , 16.4Å	Epithelial Cells
<u>Micro</u> (beta sheet & crystallites)	X-ray crystallite [derived from 18,32]	59Å (12 chains) X 22Å (5 sheets) X 90 or 180Å (116,820Å <sup>3</sup> or 233,640Å <sup>3</sup> ; 48.9Å or 61.6Å)	Posterior Silk Gland
	Genetic crystalline domain, per gene copy [derived from 8]	~295aa; 22,125Å <sup>3</sup> ; 28.1Å	Posterior Silk Gland
	Genetic crystalline segment to fit X-ray data	10 crystalline blocks of 295aa, 221,250Å <sup>3</sup>	Posterior Silk Gland
<u>Meso</u> (fibrils, fibers)	Fibers	silkworm: 9X10 <sup>3</sup> m long, 10 - 30 μm diam. spider: variable length, ~ 1- 5 μm diam.	Spinneret
<u>Macro</u> (composites)	Cocoon, Orb Web	centimeter to meter	External Environment

<sup>1</sup>based on 75 Å<sup>3</sup> per residue; <sup>2</sup>length = cube root; aa = amino acid.

#### ACKNOWLEDGEMENTS

We thank S. Lombardi, R. Beckwitt, C. Mello, S. Arcidiacono, K. Senecal, S. Stockwell, K. Kerkam, T. Thannhauser, G. Nemethy, and H. Scheraga for various contributions to this work.

#### REFERENCES

1. E. Baer, A. Hiltner and H. D. Keith, *Science* **235**, 1015 (1987).
2. K. Kerkam, D. Kaplan, S. Lombardi and C. Viney, *Nature* **349**, 596 (1991).
3. J. Kovoov, in *Ecophysiology of Spiders*, edited by W. Nentwig (Springer-Verlag, Heidelberg, 1987), p. 160.
4. J. M. Gosline, M. E. DeMont and M. W. Denny, *Endeavour* **10**, 37 (1986).
5. D. L. Kaplan, S. J. Lombardi, W. Muller and S. Fossey, in *Biomaterials: Novel Materials from Biological Sources*, edited by D. Byrom (Stockton Press, New York, 1991).
6. S. O. Andersen, *Comp. Biochem. Physiol.* **35**, 705 (1970).
7. Y. Tsujimoto and Y. Suzuki, *Cell* **18**, 591 (1979).
8. L. P. Gage and R. F. Manning, *J. Biol. Chem.* **255**, 9451 (1980).
9. M. Xu and R. V. Lewis, *Proc. Natl. Academy Sci.* **87**, 7120 (1990).
10. S. J. Lombardi and D. L. Kaplan, *Polym. Preprints, Div. Polym. Chem., Am. Chem. Soc.* **31**, 195 (1990).
11. S. J. Lombardi and D. L. Kaplan, *Acta Zool. Fennica* **190**, 243 (1990).
12. I. W. Prosser and R. P. Mecham, in *Self-Assembling Architecture*, edited by J. E. Varner (Alan Liss, Inc., New York, 1988) p. 1.
13. G. C. Candelas and F. Lopez, *Comp. Biochem. Physiol.* **74**, 637 (1983).

14. F. Lucas, J. T. B. Shaw and S. G. Smith, in Advances in Protein Chemistry, edited by C. B. Anfinsen, M. L. Anson, K. Bailey and J. T. Edsall (Academic Press, New York, 1958) p. 107.
15. K. Shimura, A. Kikuchi, K. Ohtomo, Y. Katagata and A. Hyodo, *J. Biochem.* **80**, 693 (1976).
16. S. J. Lombardi and D. L. Kaplan, *J. Arachnol.* **18**, 297 (1990).
17. D. J. Strydom, T. Haylett and R. H. Stead, *Biochem. Res. Commun.* **79**, 932 (1977).
18. R. D. B. Fraser and T. P. MacRae, Conformation of Fibrous Proteins (Academic Press, New York, 1973).
19. F. Lucas, J. T. B. Shaw and S. G. Smith, *Biochem. J.* **83**, 164 (1962).
20. R. M. Robson, in Fiber Chemistry Handbook of Fiber Science and Technology, edited by M. Lewing, E. Pearle (Marcel Dekker, NY, 1985).
21. T. Asakura, *Makromol. Chem.* **7**, 755 (1986).
22. C. Viney, in Structure, Cellular Synthesis and Assembly of Biopolymers, edited by S. T. Case (Springer Verlag, Heidelberg, 1992), in press.
23. C. Viney, K. Kerkam, L. Gilliland, D. Kaplan and S. Fossey, in Complex Fluids, edited by E. Sirota (Mater. Res. Soc. Proc., Pittsburg, PA 1992) in press.
24. T. Asakura, *Makromol. Chem. Rapid Commun.* **7**, 755 (1986)
25. M. Ishida, T. Asakura, M. Yokoi, H. Saito, *Macromol.* **23**, 88 (1990)
26. A. L. Bell and D. B. Peakall, *J. Cell Biol.* **42**, 284 (1969).
27. H. Akai, *J. Sericulture Sci. Japan* **55**, 163 (1986).
28. R. E. Marsh, R. B. Corey and L. Pauling, *Biochem. Biophys. Acta* **16**, 1 (1955).
29. S. A. Fossey, G. Nemethy, K. D. Gibson and H. A. Scheraga, *Biopolym.* in press (1991).
30. S. A. Fossey, G. Nemethy, K. D. Gibson and H. A. Scheraga, in Materials Synthesis Based on Biological Processes, edited by M. Alper, P. Calvert, R. Frankel, P. Rieke, D. Tirrell (Mater. Res. Soc. Proc., Pittsburgh, PA 1991) pp. 239-244.
31. Y. Takahashi, M. Gehoh, K. Yuzuriha, *J. Polym. Phys.* **29**, 889 (1990).
32. O. Kratky, *Farraday Soc.* **52**, 558 (1956).
33. A. Veis, in Self-Assembling Architecture, edited by J. E. Varner (Alan Liss, Inc., New York, 1988) p. 129.
34. D. E. Birk, F. H. Silver and R. C. Trelstad, in Cell Biology of Extracellular Matrix, 2nd ed, edited by E. D. Hay (Plenum Press, New York, 1991) p. 221.
35. J. Herrling and L. G. Sparrow, *J. Biol. Macromol.* **13**, 115 (1991).
36. B. Lotz and F. Colonna-Cesari, *Biochimie* **61**, 205 (1979).
37. K. Yamaguchi, Y. Kikuchi, T. Takagi, A. Kikuchi, F. Oyama, K. Shimura and S. Mizuno, *J. Mol. Biol.* **210**, 127 (1989).
38. K. Kerkam, D. L. Kaplan, S. J. Lombardi and C. Viney, in Materials Synthesis Based on Biological Processes, edited by M. Alper, P. Calvert, R. Frankel, P. Rieke, D. Tirrell (Mater. Res. Soc. Proc., Pittsburgh, PA 1991) pp. 239-244.
39. J. Magoshi, Y. Magoshi and S. Nakamura, *J. Appl. Polym. Sci.* **41**, 187 (1985).
40. E. K. Tillinghast, S. F. Chase and M. A. Townley, *J. Insect Physiol.* **30**, 591 (1984).
41. R. F. Foelix, in Biology of Spiders (Harvard University Press, Cambridge, MA 1982) p.121.
42. M. G. Dobb and J. E. McIntyre, *Adv. Polym. Sci.* **60/61**, 61 (1984).
43. D. B. Peakall, *J. Experiment. Zool.* **176**, 257 (1971).
44. M. A. Townley and E. K. Tillinghast, *J. Arachnol.* **16**, 303 (1988).