Spider Silk: Ancient Ideas for New Biomaterials

Randolph V. Lewis†

Department of Molecular Biology, University of Wyoming, Laramie, Wyoming 82071-3944

Received August 12, 2005

Contents

Dr. Randy Lewis received his B.S. degree from CalTech in chemistry and his Ph.D. in biochemistry from the University of California, San Diego. Following postdoctoral work at the Roche Institute of Molecular Biology he joined the Molecular Biology Department at the University of Wyoming. In 2005 he was awarded the H. Duke Humphrey Distinguished Faculty Award and previously received the President's Outstanding Researcher Award.

1. Introduction

Spider silk has been noted for its extraordinary properties since ancient times. In many cultures there are stories based on the use of spider silk. Over the past 150 years numerous scientific and popular articles have been written about spider silk. However, it is only in the past few years that an understanding has emerged of the reasons for the unique mechanical properties that spider silk possesses. In particular, the proteins that comprise the silks and their sequences have provided key information that relates directly to these properties. This review is intended to present this information, put it into the context of silk fibers, and indicate where further studies are needed.

Spiders have been using protein-based nanomaterials with the ability to self-assemble into fibers and sheets for over 450 million years. Spiders are unique because of the use of silks throughout their life span and their nearly total dependence on silk for their evolutionary success.[1,2] Spiders also have evolved the ability to produce as many as six different silk fibers that have differing tensile strengths and elasticities. There were periods of fairly intense study of spider silk prior to World War II and in the late 1950s. However, progress was relatively slow, especially when compared to that on silkworm silk. Beginning in the 1970s Work, Gosline, and Tillinghast and their groups revived interest in spider silk with several papers describing the physical, mechanical, and chemical properties of spider silks. That information led to the current intense interest in these unique biomaterials.

1.1. Biological Aspects of Spider Silk Production

The typical orb-shaped spider web is constructed from different silks, each of which is produced in a separate gland (Figure 1). The orb-web-weaving spiders live off the ground and use the web to capture their prey. The web is designed to both stop the flying prey on the surface of the web and then immobilize it long enough for the spider to reach it. This design has been shown to be a nearly optimal combination of design and fiber properties.[3] The related cob web weavers have a similar lifestyle, but their webs typically trap the prey inside a maze of fibers instead of on the surface. The various silks produced by the orb-web-weaving spiders, the glands that produce them, and the uses of each silk are shown in Figure 1.

The non-orb-web-weaving spiders constitute a large majority of spiders and include those species that do not use a web to capture prey. Non-orb-web-weaving spiders produce fewer silks and use them differently in many cases. Their silks are used for eggcases, for lining underground burrows, and for above ground shelters among other uses.

† Address correspondence to Randolph V. Lewis, Department of Molecular Biology, University of Wyoming, 1000 E. University Ave., Laramie, WY 82071-3944 [telephone (307) 766-2147; fax (307) 766-5098; e-mail silk@uwyo.edu].

10.1021/cr010194g CCC: $59.00
Published on Web 06/15/2006

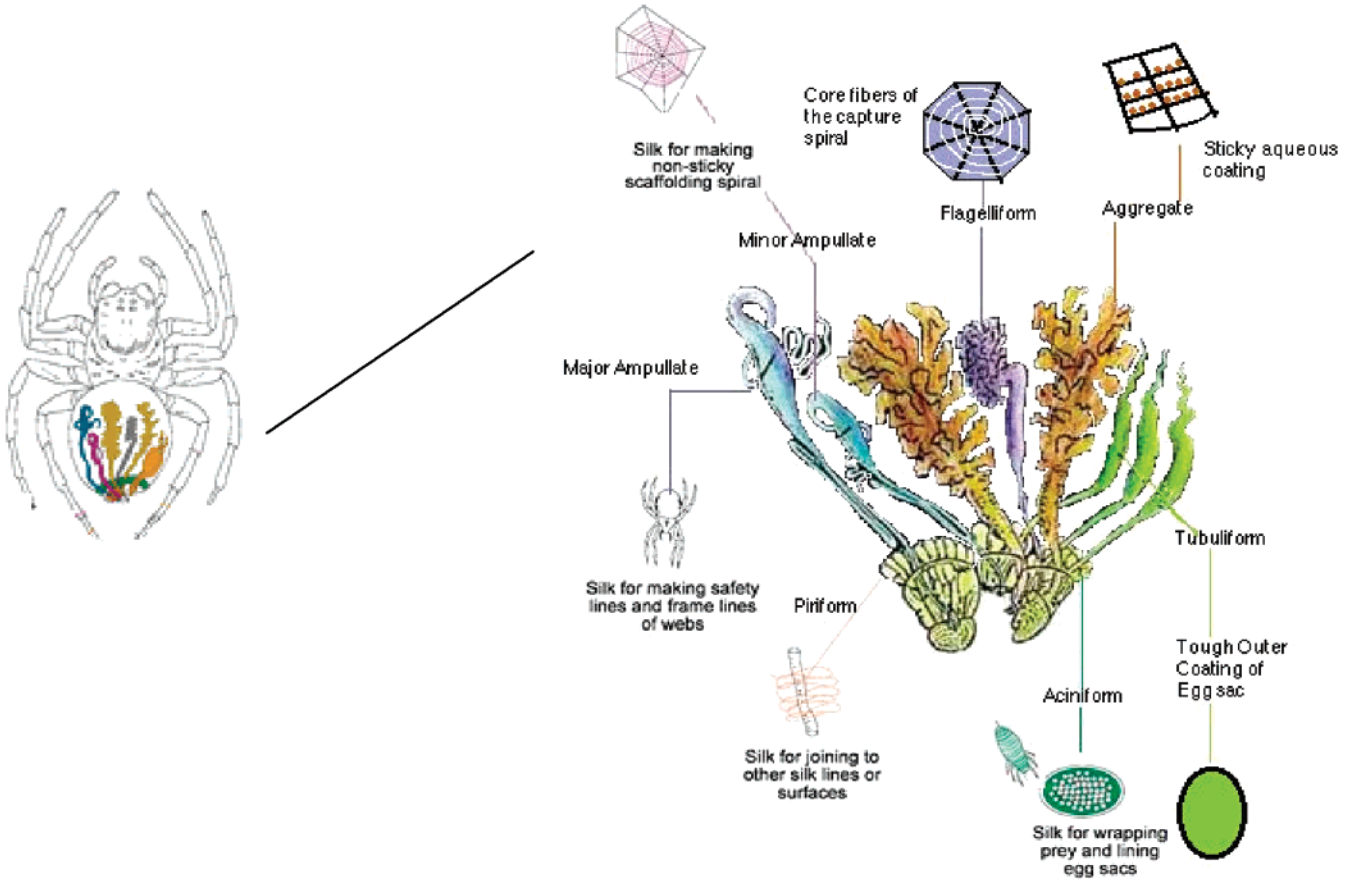

Figure 1. Spider silk glands and silk uses.

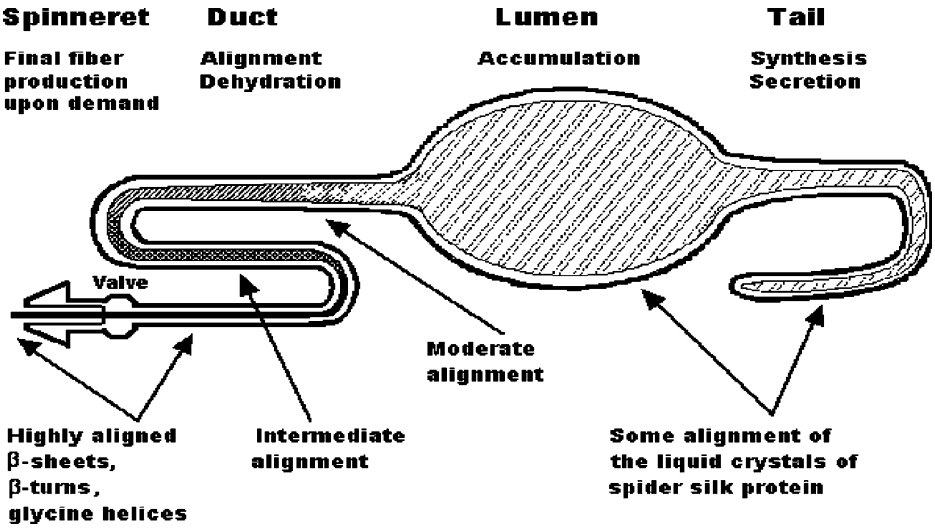

Figure 2. Major ampullate gland diagram. The various parts of the gland are labeled and their functions noted.

Although each of the silk glands has its own distinctive shape and size, they are functionally organized in a similar pattern (Figure 2). The majority of the gland is a reservoir of soluble silk protein synthesized in specialized cells at the distal end of the gland and secreted into the lumen of the gland. The soluble silk is then pulled down a narrow duct during which physical and chemical changes occur, producing the solid silk fiber. A muscular valve is present at the exit to the spinneret that can control the flow rate of the fiber and may control the fiber diameter. The silk exits through spigots on the spinnerets, of which there are three pairs: anterior, median, and posterior, each specific for a silk type.

Due to their large size and ease of study, the major ampullate glands have received the majority of attention, and most of what is known about the synthesis of silk proteins is based on this gland. However, morphological and histochemical studies of the other glands support the ideas developed from the major ampullate gland research. Synthesis of the silk protein(s) takes place in specialized columnar epithelial cells.[4] There appear to be at least two different types of cells producing protein in the major

Table 1. Comparisons of Mechanical Properties of Spider Silk[a,33]

material	strength ($N\ m^{-2}$)	elongation (%)	energy to break ($J\ kg^{-1}$)
dragline silk	4×10^9	35	4×10^5
minor ampullate silk	1×10^9	5	3×10^4
flagelliform silk	1×10^9	>200	4×10^5
tubuliform	1×10^9	20	1×10^5
aciniform	0.7×10^9	80	6×10^9
Kevlar	4×10^9	5	3×10^4
rubber	1×10^6	600	8×10^4
tendon	1×10^6	5	5×10^3

[a] Some data from Gosline, J. M.; Dennv, M. W.; DeMont, M. E. *Nature* **1984**, *309*, 551.

ampullate gland,[5] which is consistent with data showing two proteins in the silk fibers from these glands. The newly synthesized protein appears as droplets within the cell and are then secreted into the lumen of the gland.

The process of spinning the fiber is clearly one of nature's most amazing feats. The silk protein is spun at ambient temperature and pressure from an aqueous solution to form a fiber with the amazing mechanical properties described below. An excellent review of the entire spinning process has been published,[6] so only key factors will be presented here. The state of the protein in the lumen of the gland is believed to be a liquid crystal,[6,7] a state that prevents fiber formation until passage down the duct. This is probably accomplished by a combination of protein structure and concentration that prevents aggregation into large protein arrays. It has been shown that silk in the lumen is not birefringent, whereas silk becomes increasingly birefringent as it passes down the duct.[8] Thus, the ordered array of protein seen in the final fiber occurs in the duct. This appears to be due to the mechanical and frictional forces aligning the protein molecules and altering their secondary structure to the final fiber form. Iizuka[9] has proposed a similar mechanism for silkworm silk formation. Experimental evidence for this has been the ability to draw silk fibers directly from the lumen of isolated major, minor, and cylindrical glands (M. Hinman, personal communication), implying that the physical force of drawing the solution is sufficient for fiber formation.

1.2. Mechanical Properties

The unusual mechanical properties are the key features attracting researchers to spider's silks. Orb-web-weaving spiders appear to use the minimum amount of silk necessary in their webs to catch prey. The web has to stop a rapidly flying insect nearly instantly, so that the prey becomes entangled and trapped. To do this, the web must absorb the energy of the insect without breaking and yet not act as a trampoline and bounce the insect away from the web. Gosline et al.[3] reviewed several aspects of this property and concluded that spider silk and the web are nearly optimally designed for each other.

As with any polymer, especially those composed of protein, there are numerous factors that can affect the tensile strength and elasticity. These can include temperature, hydration state, and extension rate.[10] Another key factor is the diameter of the fiber, which can vary greatly along the length of a single fiber. Even with these caveats it is clear that dragline silk is a unique biomaterial. As seen in Table 1, dragline and flagelliform silks will absorb more energy prior to breaking than nearly any commonly used material. It is nearly as strong as several of the current synthetic fibers but can outperform them in many applications requiring total energy absorption. Also of note are the differences between the four different spider silks. There is a 5-fold range in tensile strength and a nearly 50-fold range in elongation.

Table 2. Amino Acid Composition of Silks from *A. diadematus* (Based on Andersen, 1970)[a,b]

amino acid	major ampullate	minor ampullate	flagelliform	aciniform	tubuliform
Asp	1.04	1.91	2.68	8.04	6.26
Thr	0.91	1.35	2.48	8.66	3.44
Ser	7.41	5.08	3.08	15.03	27.61
Glu	11.49	1.59	2.89	7.22	8.22
Pro	15.77	tr[c]	20.54	2.99	0.59
Gly	37.24	42.77	44.16	13.93	8.63
Ala	17.60	36.75	8.29	11.30	24.44
Val	1.15	1.73	6.68	7.37	5.97
Ile	0.63	0.67	1.01	4.27	1.69
Leu	1.27	0.96	1.40	10.10	5.73
Tyr	3.92	4.71	2.56	1.99	0.95
Phe	0.45	0.41	1.08	2.79	3.22
Lys	0.54	0.39	1.35	1.90	1.76
His	tr	tr	0.68	0.31	tr
Arg	0.57	1.69	1.13	4.09	1.49

[a] The amino acids are abbreviated with the three-letter code and the silks identified by their gland name. [b] Anderson, S. O. *Comp. Biochem. Physiol.* **1970**, *35*, 705. [c] Trace.

Another unique feature of major ampullate silks is their supercontraction when exposed to water. Depending on the spider species and other factors, these silks will contract to 50% or less of their original length in water.[11] This silk fiber supercontraction is the only known example of supercontraction in water. This supercontraction can occur repeatedly with virtually identical results.[11] Suggestions are that it provides an advantage to the spider by tightening the web whenever the humidity is very high by contraction of the attachment lines and the framework of the web.

2. Chemical Data

Spider silks have been known to be composed predominantly of protein since the 1907 studies of Fischer.[12] In fact, except for the sticky material deposited on the catching spiral thread, no significant amount (<0.1%) of any other compound has been show to be covalently linked to the silk proteins, including sugars, minerals, and lipids. In the major ampullate (Ma) silks the combination of Glu, Pro, Gly, and Ala comprise 80% of the silk from each species (Table 2). However, the proportion of Pro can differ significantly among species. As will be discussed below these differences can be accounted for by differing ratios of two proteins.

Minor ampullate (Mi) silk has an amino acid composition similar to that of major ampullate with two major differences. The first is that there is virtually no Pro present in Mi silk and, second, the Glu content is markedly reduced. Flagelliform silk, on the other hand, has a very high Pro content, a much reduced Ala level, and an elevated level of Val. Aciniform silk shows several divergent amino acids from the typical silk; nearly equal levels of Gly, Ala, and Ser, which comprise only ~40%, high levels (26%) of large hydrophobic amino acids, and nearly 5% positively charged amino acids. Tubuliform silk shows the greatest divergence, with Ser and Ala dominating and Gly much reduced and the other amino acid levels showing similarities to different silks.

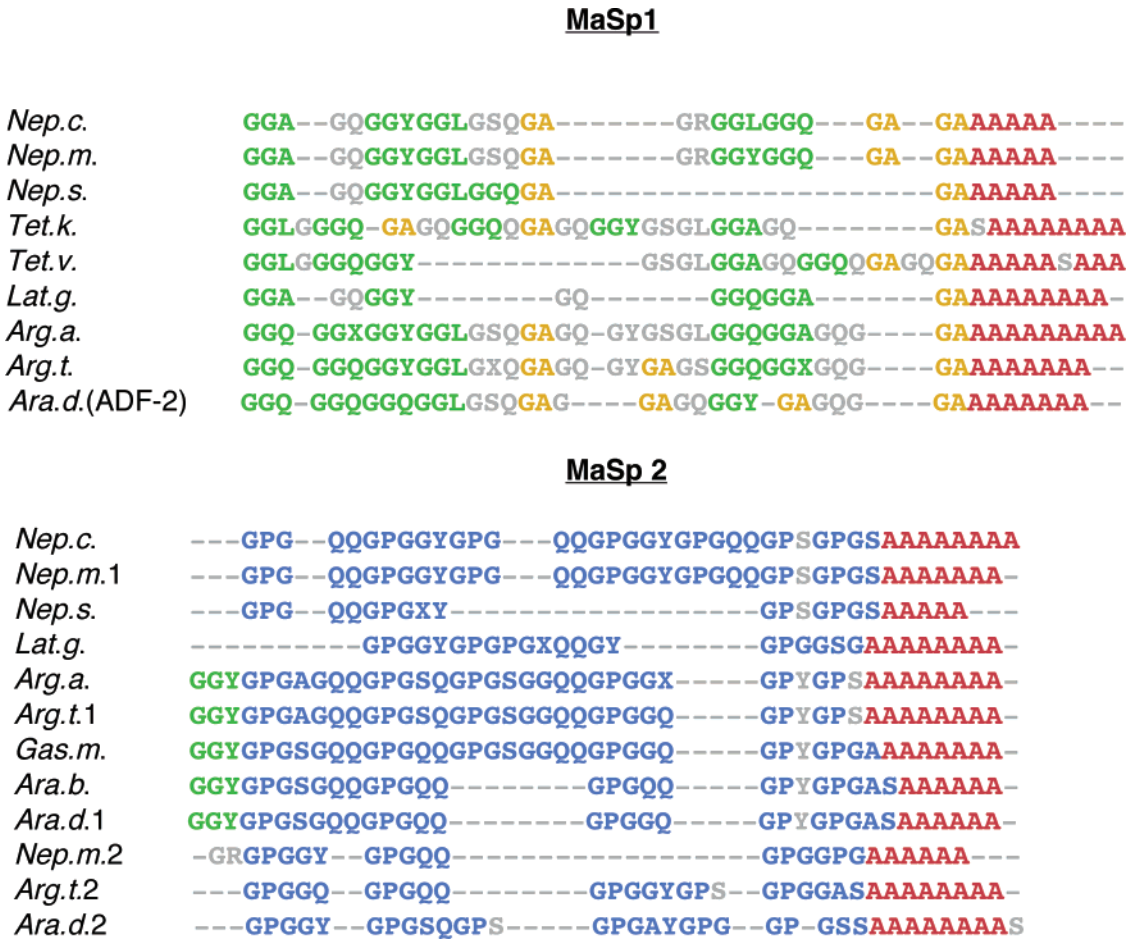

Figure 3. Amino acid sequences of the consensus repetitive sequence of the major ampullate silk proteins in various orb-weaving species. Amino acids are denoted in one-letter code. (−) indicates an amino acid not present when compared to the other sequences. Spiders: *Nep.c.*, *Nephila clavipes*; *Nep.m.*, *Nephila madagascariensis*; *Nep.s.*, *Nephila senegalensis*; *Lat.g.*, *Lactrodectus geometricus*; *Arg.t.*, *Argiope trifasciata*; *Arg.a.*, *Argiope aurantia*; *Gas.m.*, *Gasteracantha mammosa*; *Ara.b.*, *Araneus bicentenarius*; *Ara.d.*, *Araneus diadematus*; *Tet.k.*, *Tetragnatha kanaiensis*; *Tet.v.*, *Tetragnatha versicolor*.

2.1. Protein Sequences

2.1.1. Major Ampullate Silk

When detailed studies of spider silk protein were started, the amino acid compositions for a number of spider silks and mechanical test data were the only information available. The first cDNA sequence for a spider silk protein from the major ampullate silk of *Nephila clavipes*, termed MaSp1, was published in 1990 (Figure 3).[13] That partial sequence contained similar but nonidentical repeats with sequence motifs including stretches of polyAla of up to seven residues alternating with $(GGX)_n$ sequences, where the X residues are Y, L, and Q in that sequence order. There are differences in the number of GGX units in each repeat, but there are virtually no substitutions for the three X amino acids. Searches of protein sequence databases found no matches greater than five amino acids, indicating that they represent novel structures.

The second major ampullate silk protein was found on the basis of the identification of a proline-containing peptide in the silk, which was absent from protein encoded by the cDNA for MaSp1. This led to the cDNA for the second major ampullate silk protein, MaSp2.[14] The repetitive sequence and predicted structure of this protein forced a rethinking of the mechanism for elasticity and tensile strength and led to the models described below. The novel sequence features in the consensus repeats of this protein (Figure 3) include the polyAla regions as in MaSp1 alternating with sequences of GPGQQGPGGY. As with MaSp1, virtually no amino acid substitutions occur in the repetitive regions. The repeats shown in Figure 3 are the consensus of the hundreds of repeats in the protein, and most repeats have minor differences from the consensus either in the number of three or five amino acid motifs or the number of alanine residues in the polyAla motif.

Using Northern blotting, restriction enzyme digestion, and Southern blotting of genomic DNA, the sizes of the mRNA and genes for all of the above proteins were determined. The mRNA sizes for MaSp1 and MaSp2 were shown to be about 12.5 and 10.5 kb, respectively, generating proteins of >350 kDa. Analyses of genomic DNA indicated the absence of large introns in the coding regions and the lack of any detectable introns in the main repetitive portions of the genes.

Figure 3 shows published sequences for these two major ampullate silk proteins from other orb-weaving or derived orb-weaving spiders.[15−17] The data clearly show that high sequence conservation has been maintained over the 125 million years since these various species diverged from each other. In particular, the sequence motifs described above show almost no alterations in that period of time.

When the major ampullate silk cDNAs from more primitive spiders were examined, the results were very different (Figure 4).[17,18] In fact, except for *Dolomedes*, there is little similarity among the various species or to the orb weavers.

Figure 4. Major ampullate silk protein consensus sequences from non-orb-weaving spiders. The amino acids are listed by one letter code, and (−) indicates missing amino acids in comparison to other sequences. The spider family but not species name is given.

The only motifs that are conserved are polyAla and polyGA. However, several similar motifs are found such as polyGS and polyGV. Other sequences that may be similar but involve a very different structural amino acid include polyAQ, polyAY, polyS, and polyT. Particularly noticeable is the lack of proline-containing motifs that are so highly represented in the orb-weaving spiders' major ampullate silks. The amino acid compositions of the protein sequences of the repetitive regions correlate very well with the amino acid composition of the silks determined from proteins isolated from the same glands as the cDNA. Thus, it is unlikely that proteins similar to those of the orb-weaving spiders can represent a significant fraction of the fiber.

An interesting question that remains to be answered about these sequences is whether they are truly major ampullate silk proteins or whether the glands producing them are not directly related to the major ampullate glands of the orb weavers. The glands and species were all chosen because morphological data indicated the glands were major ampullate, but the incredible diversity of sequences indicates this may not be the case. If they are major ampullate glands, then a tremendous diversification in sequence must have occurred during the same time frame as high conservation was occurring during the species diversification of the orb weavers. In either case, it is clear that the constraints of web capture have substantially suppressed major ampullate silk protein divergence in the orb weavers, whereas silks with less specific requirements have diverged enormously.

2.1.2. Minor Ampullate Silk Proteins

Shortly after the initial sequences of the major ampullate silk proteins were published, cDNAs representing minor ampullate silk (Mi) protein transcripts from *N. clavipes* were sequenced.[19] The sequence obtained (Figure 5) shows similarity to the major ampullate silk proteins but conspicuous differences as well. GGX and short polyAla sequences are present, but the longer polyA motifs in the MaSps are replaced by $(GA)_n$ repeats. The consensus repeats show very similar organizations, but the number of GGX and GA repeats varies greatly.

A second cDNA from the same library shows considerable similarity to the first minor ampullate silk protein but clearly encodes a different protein (MiSp2). Multiple copies of a unique conserved 130 amino acid nonrepetitive sequence were found in both proteins; these were termed "spacer" regions as they break up the repetitive regions (Figure 5). These serine-rich spacer regions are reminiscent of the 30 amino acid nonrepetitive amorphous domains of silkworm silk, which also interrupt repetitive regions.[20] Interestingly, the MiSp spacer regions are highly conserved in sequence and thus differ from the repetitive regions, which show variation like the MaSp proteins. The function of this spacer

Figure 5. Consensus amino acid sequences of minor ampullate silk proteins from orb-weaving spiders. Abbreviations are the same as in Figure 3.

Figure 6. Flagelliform silk protein consensus sequences. Abbreviations are the same as in Figure 3.

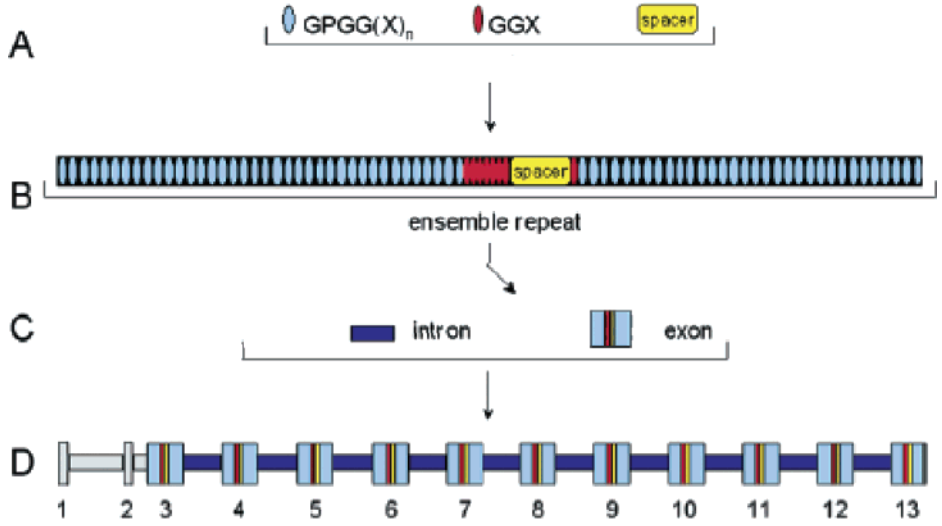

Figure 7. Schematic diagram of the flagelliform silk gene structure. The organization flows from the sequence of the individual protein modules (A) to the protein repeat unit (B) to the intron and exon units of the gene (C) to the complete gene structure (D), which is nearly 30 kb.

region is currently unknown, but it may serve to separate crystalline regions as well as participate in interchain protein associations through charged residues.

Northern blotting confirmed that minor ampullate silk is composed of at least these two repetitive proteins. The MiSp1 and MiSp2 transcripts are 9.5 and 7.5 kb, respectively, generating proteins of >250 kDa in size. Genomic DNA analysis by Southern blotting using probes to both the 3′ end of the clone and the repetitive region demonstrated the presence of only two MiSp genes. Restriction enzyme digestion and Southern blotting of the genomic DNA established that the entire genomic fragment, which corresponds closely in size to the transcript, is composed of the same sequences as found in the partial cDNA. This indicates the lack of any large introns or differing coding regions within this gene.[19]

2.1.3. Flagelliform Silk Proteins

The next silk protein cDNAs to be cloned were from the *N. clavipes* flagelliform gland encoding a protein for the catching spiral silk (Figure 6).[21] The cDNA clones were found to contain sequences encoding a 5′ untranslated region and a secretory signal peptide, numerous iterations of a five amino acid motif, and the C-terminal end. Northern blotting analysis indicated an mRNA transcript of ~15 kb, generating a protein of nearly 500 kDa. The amino acid sequence predicted from the gene sequence suggests a model of protein structure that explains the physical basis for the elasticity of spider silk and which is similar to that for MaSp2 (described below).

In comparison to the other known silk genes, the flagelliform protein has four very distinctive features: (1) It has the simplest repeat unit, a pentapeptide (GPGGX), that also appears as a motif within the MaSp2 repeat unit and the GGX motif found in MaSp1. (2) The pentapeptide repeat units have sequence variations not seen in the other silk proteins. When the repeats are aligned with each other, the fifth codon frequently results in an amino acid substitution. Although this variability suggests that the fifth amino acid is not critical to the protein structure, only a very limited set of amino acids (A, V, S, Y) are found in this position. Curiously, there are three predominant patterns to the strings of repeats: strings of As, alternating (V, S), and alternating (Y, S). (3) A highly conserved 34 amino acid spacer unit (sequence: TITEDLDITIDGADGPITISEELTISGA) occurs among the basic repeat units, which is reminiscent of MiSp1 and MiSp2. This region is also non-silk-like, with many charged and hydrophilic amino acids. This region has no similarity with

Acinform

Gly Ser Ala Gly Pro Gln Gly Gly Phe Gly Ala Thr Gly Gly Ala Ser Ala
GGA TCT GCT GGC CCT CAA GGT GGA TTC GGT GCC ACA GGT GGA GCG TCT GCT
Gly Leu Ile Ser Arg Val Ala Asn Ala Leu Ala Asn Thr Ser Thr Leu Arg
GGC CTT ATC TCC AGA GTA GCA AAC GCA CTT GCC AAT ACA TCA ACA TTG AGA
Thr Val Leu Arg Thr Gly Val Ser Gln Gln Ile Ala Ser Ser Val Val Gln
ACT GTC CTC AGA ACT GGT GTA TCC CAA CAG ATT GCC TCC AGC GTG GTA CAG
Arg Ala Ala Gln Ser Leu Ala Ser Thr Leu Gly Val Asp Gly Asn Asn Leu
AGA GCC GCT CAG TCG TTG GCC AGT ACT CTC GGA GTC GAC GGA AAT AAC TTG
Ala Arg Phe Ala Val Gln Ala Val Ser Arg Leu Pro Ala Gly Ser Asp Thr
GCC AGA TTC GCG GTA CAG GCC GTC TCT CGA CTG CCC GCC GGA TCA GAC ACT
Ser Ala Tyr Ala Gln Ala Phe Ser Ser Ala Leu Phe Asn Ala Gly Val Leu
TCT GCT TAC GCT CAA GCA TTC TCT AGT GCG CTC TTC AAT GCC GGA GTT CTC
Asn Ala Ser Asn Ile Asp Thr Leu Gly Ser Arg Val Leu Ser Ala Leu Leu
AAT GCA AGC AAC ATT GAC ACA TTG GGA TCC CGA GTT CTC TCA GCA CTT TTG
Asn Gly Val Ser Ser Ala Ala Gln Gly Leu Gly Ile Asn Val Asp Ser Gly
AAC GGA GTA TCA AGT GCG GCG CAA GGT CTT GGC ATC AAT GTA GAT AGC GGC
Ser Val Gln Ser Asp Ile Ser Ser Ser Ser Ser Phe Leu Ser Thr Ser Ser
AGT GTA CAA AGT GAC ATT AGT TCC AGT AGC AGC TTC CTC TCA ACA AGC TCG
Ser Ser Ala Ser Tyr Ser Gln Ala Ser Ala Ser Ser Thr Ser Gly Ala Gly
TCT TCG GCC AGT TAC TCT CAG GCA TCA GCT TCT TCG ACC AGC GGT GCC GGA
Tyr Thr Gly Pro Ser Gly Pro Ser Thr Gly Pro Ser Gly Tyr Pro Gly Pro
TAC ACA GGA CCT TCT GGA CCT TCC ACT GGA CCT TCT GGC TAC CCT GGG CCT
Leu Gly Gly Gly Ala Pro Phe Gly Gln Ser Gly Phe Gly
TTG GGT GGC GGA GCG CCG TTC GGT CAA TCA GGC TTT GGC

Figure 8. Consensus repeat sequence of *Arg. trifasciata* aciniform silk protein and cDNA. The protein sequence for the 14 repeats is shown above in the three-letter code, and the codon for that amino acid is shown below. The extremely limited number of bases that are not identical in the entire 14 repeats are underlined in the figure.

any known protein, and its significance to the structure of the protein is unknown. (4) The C-terminal nonrepetitive region of flagelliform silk also shows no homology to that of the MaSp and MiSp proteins. Because it appears that flagelliform silk arose near the time of the divergence of the orb-weaving spiders from the others, it is interesting that it shows such differences from the silks it presumably arose from.

The first genomic sequence data for any spider silk were from the flagelliform silk.[22] This gene with a size of nearly 30 kb was of particular interest as it showed a pattern of repeating introns and exons, with the introns more highly conserved than the exons (Figure 7). When two related species (*N. clavipes* and *Nephila madagascariensis*) were compared, the introns were more similar within each species than the correspondingly located introns in the other species. Thus, a significant homogenization must be occurring in these genes, leading to the intron sequence conservation in each gene. The reason for this intron−exon structure as well as the high sequence conservation of the introns is unknown. It is likely necessary for either gene or initial transcript stability.

2.1.4. Aciniform Silk Proteins

There are numerous aciniform glands found below two different spinnerets (see Figure 1). The two separate sets of glands were isolated to make cDNA libraries from *Argiope trifasciata*.[23] From the two libraries 59 silk clones were isolated, which all contained similar sequences. The longest cDNA clone was 8618 bp in length and was completely sequenced. The transcript could be translated in just one reading frame and was named AcSp1, an abbreviation of "aciniform spidroin 1."

The predicted amino acid composition of AcSp1 generally agreed with the composition of the protein from the aciniform silk glands from *Argiope*. This correspondence was consistent

```
Ar.g 1 KTTSTSTSGSQADSRSASSSASQASASAFAQQSSASLSSSSSFSS
Ar.g 2 KTTSTSTSGSQADSRSASSSASQASASAFAQQSSASLSSSSSFSS
Ar.g 3 KTTSTSTSGSQADSRSASSSASQASASAFAQQSSASLSSSSSFSS

Ar.g 1 AFSSATSISAVGNVGYQLGLKVANSLGLGNAQALASSLSQAVSAVG
Ar.g 2 AFSSATSISAVGNVGYQLGLKVANSLGLGNAQALASSLSQAVSAVG
Ar.g 3 AFSSATSISAVGNVGYQLGLKVANSLGLGNAQALASSLSQAVSAVG
```

Figure 9. Amino acid sequence alignment among the three repeat units of *Arg. gemmoides* (*Ar.g*) TuSp1. The alignments were created using Mac vector default settings. Darkly shaded regions indicate identical amino acids. Lightly shaded region with bold letters indicate similar amino acids.

with AcSp1 being a major constituent of the protein stored in the aciniform silk glands. AcSp1 consists of a series of 200 amino acid long repeats ending with a 99 amino acid long, nonrepetitive C terminus. In BLAST searches with the AcSp1 C terminus, the top matches were to published spidroins. This similarity with other spidroins and the repetitive genetic architecture of the AcSp1 cDNA clearly show AcSp1 as one of the spider silk fibroin gene family.

The individual repeat units within the protein are remarkably conserved at both the DNA and amino acid sequence levels (Figure 8). Several of the repeats show an amazing 100% identity to each other. Although the repeat unit is 600 bp long, there were only 12 variable sites in the alignment of 14 repeats, with most of the variation in the terminal repeat. Homogeneity of repeat units within a gene has been seen in all of the characterized spider fibroins (summarized in Gatesy et al.[17]) and in fibroins from lepidopterans,[20,24,25] but this low level of variation among AcSp1 intragenic repeats is exceptional. On average, the 600 bp long units were ~99.9% similar at the DNA level.

Whereas the repeat units are highly conserved within AcSp1, sequence searches with the repetitive region found no matches in the nucleotide and protein databases. In contrast to the C-terminal region, the AcSp1 repetitive region has no real similarity to previously characterized genes and proteins. Aciniform fibroin has few of the subrepeats that characterize fibroins from the major ampullate, minor ampullate, flagelliform, and tubuliform glands of araneoid spiders.[13,14,17,19] PolyAla, a motif that has been hypothesized to account for the high tensile strength of major ampullate silk,[26−28] is notably lacking in AcSp1. Recently, a cDNA for AcSp1 from *Nephila* has been partially sequenced (David Perry and Randy Lewis, unpublished data) which shows a sequence similar to but distinct from that of *Argiope*.

2.1.5. Tubuliform (Eggcase) Silk

Tubuliform silk is unique among the orb-weaving spider silks due to its distinct amino acid composition, specific time of production, and atypical mechanical properties. A recent publication[29] showed the tubuliform gland cDNA and protein sequences from three orb-weaving spiders, *Argiope aurantia*, *Araneus gemmoides*, and *N. clavipes*. Amino acid composition comparison between the predicted tubuliform silk protein sequence (TuSp1) and the corresponding gland protein confirmed that TuSp1 is the major component in tubuliform gland in the three spiders. The majority of previously studied spider silk proteins can be explained by different combinations and arrangements of four amino acid motifs, A_n, $(GA)_n$, $(GGX)_n$, and $(GPGXX)_n$ (as described above). However, these amino acid motifs are rarely represented in tubuliform silk proteins (Figure 9). Instead, TuSp1 has a more complex molecular architecture with new amino acid motifs such as S_n, $(SA)_n$, $(SQ)_n$, and GX (X represents Q, N, I, L, A, V, Y, F, or D). In addition, some of these motifs are also found in silk proteins of the basal taxa spiders such as *Plectreurys tristis* and the mygalomorph *Euagrus chisoseus*.[17]

The relatively high levels of the large amino acids are consistent with the results of fiber X-ray diffraction.[30] These data indicate that eggcase silk has a larger *b* dimensional value in the β-sheet region than either major or minor ampullate silk, an indication of the presence of large side-chain amino acids.[30] Secondary structure predictions of TuSp1 sequence are consistent with this as large stretches of β-sheet are likely to occur with this protein. Furthermore, the presence of large side-chain amino acids in tubuliform silk proteins also agrees with the results of transmission electron microscopy (TEM), which show the presence of streaks in tubuliform silk indicating variations in β-sheet spacing.[31] Mechanical testing has shown tubuliform silk to have a relatively high tensile strength with a fairly low elasticity, properties that are similar to the mechanical properties of minor ampullate silk.[32] However, its low ability to withstand bending before breaking, a glasslike or crystalline behavior, makes it different from all of the other silks.[33] This glasslike or crystalline feature of tubuliform silk can be explained by the large side-chain amino acids forming the highly frustrated β-sheet crystalline structure, which would not allow sufficient protein movement upon bending.

Repeat unit comparison within species as well as among three spiders exhibits high sequence conservation. The only divergence among repeat units within species is due to single-base substitutions. The high homogeneity among TuSp1 repeat units within species in all three spiders is an indication of within-gene concerted evolution, probably through gene conversion and unequal crossing-over events. Similar results have been reported in other spider silk gene family members (as described above). Parsimony analysis based on C-terminal sequence shows that *Argiope* and *Araneus* are more closely related than either is to *Nephila*, which is consistent with phylogenetic analysis based on morphological evidence.

2.1.6. C-Terminal Region

Interestingly, the only sequence that appears to be highly conserved among the spider silk proteins is the nonrepetitive sequence located at the C terminus. There is at least 45% amino acid identity between the most divergent pair in this group (Figure 10). This conservation in the C-terminal region of major and minor ampullate silks was noted previously for the orb weavers[21] and was used by others as part of their cloning strategy.[16] It remains surprising in view of the major sequence changes occurring in these species that this sequence was conserved, but it confirms the identity of these proteins as silk proteins.

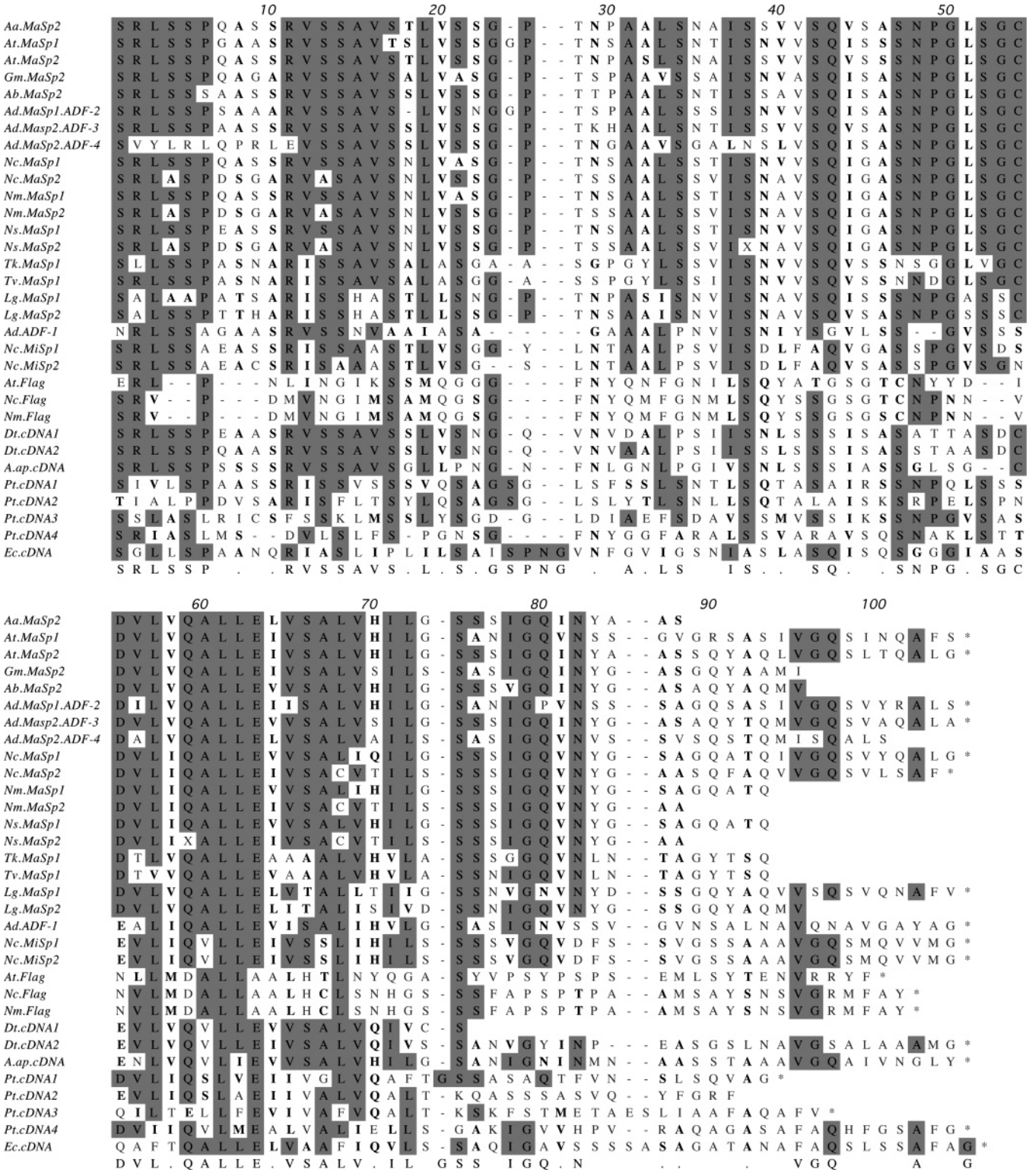

Figure 10. Spider silk protein C-terminal amino acid alignment [ranging from 71 (Dt.cDNA1) to 104 (Ec.cDNA)]. Shaded regions indicate identical amino acids: bold, similar; gray, mismatch. "−" in the sequence represents missing amino acid, and "∗" represents a stop codon. Consensus sequence shows identities in upper case, similarities are marked as ".", and mismatches are blank. Abbreviations of spider species used in this figure (from top to bottom): Aa, *Argiope aurantia*; At, *Argiope trifasciata*; Gm, *Gasteracantha mammosa*; Ab, *Araneus bicentenarius*; Ad, *Araneus diadematus*; Nc, *Nephila clavipes*; Nm, *Nephila madagascariensis*; Ns, *Nephila senegalensis*; Tk, *Tetragnatha kauaiensis*; Tv, *Tetragnatha versicolor*; Lg, *Latrodectus geometricus*; Dt, *Dolomedes tenebrosus*; Aap, *Agelenopsis aperta*; Pt, *Plectreurys tristis*; Ec, *Euagrus chisoseus.* Abbreviations used for the silk fibroins: MaSp1, major ampullate spidroin 1; MaSp2, major ampullate spidroin 2; MiSp1, minor ampullate spidroin 1; MiSp2, minor ampullate spidroin 2; Flag, flagelliform silk protein; cDNA, fibroins from unspecified glands.

It is of particular interest that the conservation of sequence is focused on certain stretches of amino acids, suggesting strongly that these regions have important functions that have been preserved for nearly 250 million years. It has now been shown that this region of the protein is present in the fiber.[33] This may be a result of design and a role the sequence plays

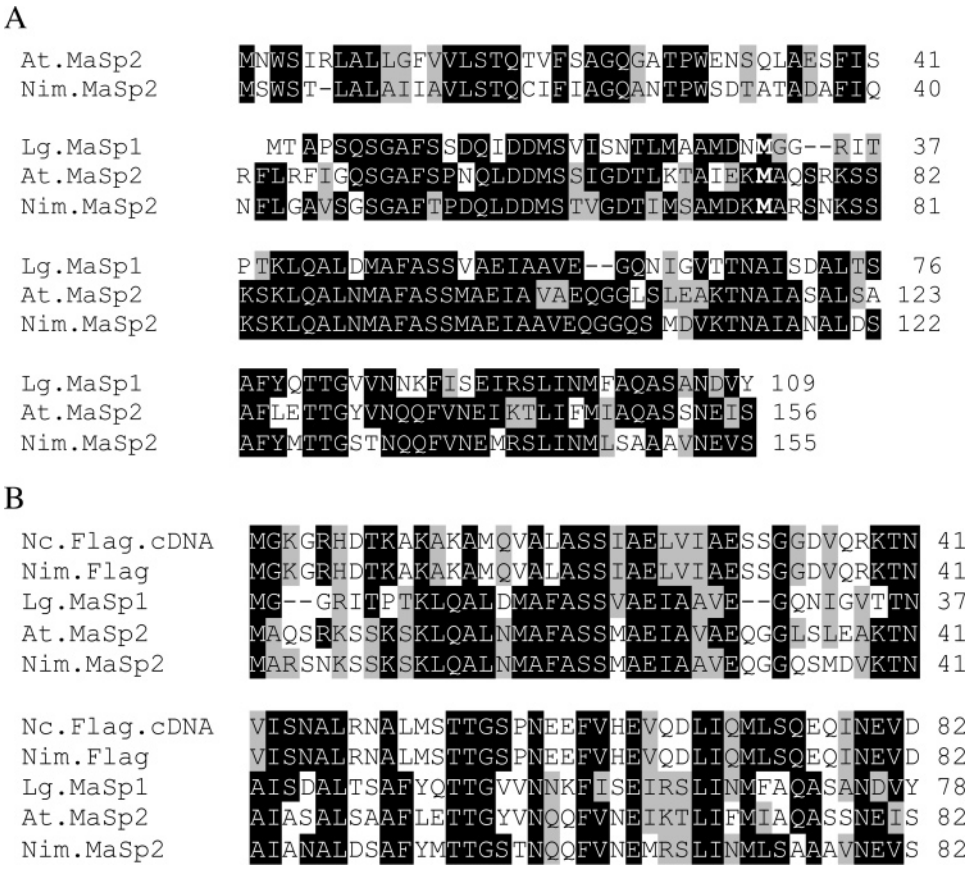

Figure 11. N-Terminal sequence alignments: (A) long isoform; (B) short isoform. Protein sequences start with the first Met in frame. Protein sequences in (A) reported in this study are derived from genomic DNA and correspond to the following GenBank accession numbers: Lg.MaSp1 (DQ059133), At.MaSp2 (DQ059136), and Nim.MaSp2 (DQ059135). Met downstream of the first, which is the first amino acid of the short isoform, is bold. Short isoform alignment (B) represents the five known N-terminal sequences. Two additional sequences were used here: Nc.Flag, flagelliform silk from *N. clavipes* cDNA (AF027972), and Nim.Flag, flagelliform silk from *N. madagascariensis* genomic DNA (AF218623). Identical amino acids have black background, similarities are gray, and mismatches are not shaded.

in fiber formation. It may also be the result of the lack of need to remove this protein segment for fiber formation, and its function could be preventing premature fiber formation while the protein is in the gland.

2.1.7. N-Terminal Region

Little is known about the N-terminal region of these spider silk proteins as only two sequences are known from flagelliform silk. The sequences appear typical of secreted proteins (Figure 11) with the usual amino acid composition and a likely enzymatic cleavage site. Three additional N-terminal sequences have recently been determined[34] from MaSp protein genomic clones, which show distinct homology to the published sequences. Additional studies are needed to determine how well conserved this part of the protein is during evolution of different silks and different species.

3. Biophysical Studies

The majority of biophysical data on spider silk have been obtained from major ampullate (dragline) silk. There are two reasons for this: (1) it is easy to obtain, and (2) it exhibits a unique combination of elasticity and high tensile strength. X-ray fiber diffraction measurements[35,36] and NMR studies[27,30,37,38] have established that spider silks contain extended, ribbonlike β-sheets, oriented such that the β-sheets run approximately parallel to the fiber axis. Solid-state NMR experiments on spiders' silks[38,39] and silkworms[40−42] characterized the composition and orientation of ribbonlike β-sheet structures in silk fibers. Silk samples labeled at both Ala and Gly residues lead to ψ and φ dihedral angles that indicate β-strand conformations in silk fibroins (Figure 12). This was confirmed by (i) DOQSY spectroscopy,[43] (ii) two-dimensional exchange spectroscopy,[44] and (iii) ^{13}C CSA and ^{2}H quadrupolar NMR experiments.[45] In addition to the β-sheet motif, solid-state NMR has produced preliminary evidence for β-turns[40] and 3_1-helices[43] in silk materials. Several early studies of silk fibers using X-ray diffraction provided some information, much of which was interpreted on the basis of the structure of silkworm silk. These studies led to the classification of dragline silk as β-sheet proteins but also showed that much of the structure was not β-sheet, appearing to be unordered.

Two recent papers[46,47] provide an interesting perspective on these spider silk sequences. The first shows that polyAla in the unfolded state tends to form a polyproline II helix, which would explain the lack of intermolecular interactions while the silk protein is in the gland. In contrast, the second paper shows that polyAla forms aggregates from intermolecular β-sheets using molecular dynamics simulations. Although this paper focused on the prion proteins, it is believed the same sequences found in the insect and spider silk proteins will behave exactly as predicted by their simulations. Thus, the protein prior to fiber formation could be in the proline helix, and during the spinning process the

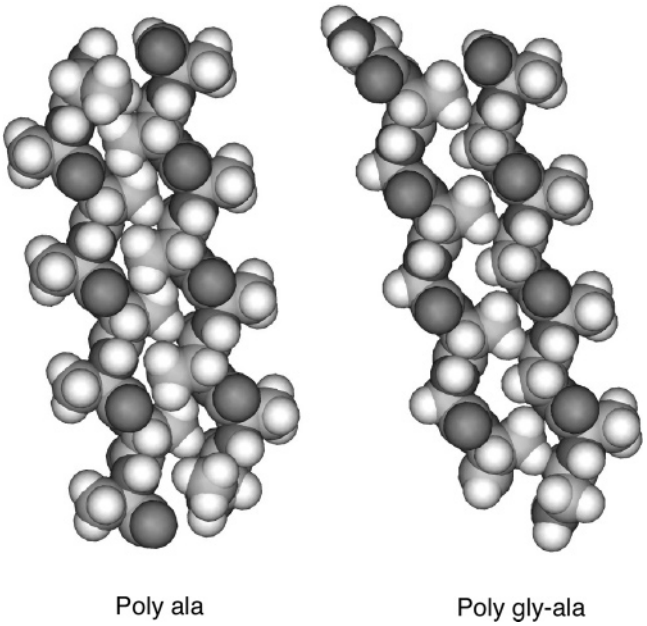

Figure 12. Computer models of the polyAla and polyGly-Ala segment shown with two chains oriented in the vertical direction.

conversion to β-sheet drives the aggregation that forms the fiber.

The GGX-rich regions of MaSp1 were originally thought to be responsible for elasticity. However, later studies with molecular modeling suggested that the GPGXX structures in MaSp2 were more likely candidates, and the larger number of these sequences in flagelliform silk further confirmed this hypothesis. The GGX structure was proposed by us to be poly-glycine II helix on the basis of molecular modeling.[30] Recent NMR data from Kummerlen et al. have supported the helical structure with a glycine II helix best fitting their data.[44] Figure 13 is a computer model showing the interdigitation possible with the GGYGGLGGQGGA repeat sequence in antiparallel strands. The Tyr and Gln residues form hydrogen bonds with the opposing main chain, and the methylene groups of all three large side chains form a hydrophobic core. Thus, these residues can provide a stabilizing energy to the fiber formation as well. This same type of effect may be the major driving force for the formation of flagelliform silk, which lacks the polyAla regions and has extended GGX regions instead.

The Pro-rich regions of MaSp2 have been proposed from computer modeling studies to generate silk elasticity, with the GPGXX pentapeptide segments thought to form β-turns that form a β-spiral.[26] The properties of flagelliform silk are consistent with this model as it is highly elastic. These protein structures would be similar to those thought to exist in other proteins with extensive turn structures (review by Tatham and Shewry[48]). The linked β-turns can easily form β-spirals, as has been suggested for elastin, and provide elasticity by simple extension and contraction of that spiral.[49] A computer model for the β-spiral is shown below (Figure 14).

The biophysical studies further demonstrate the importance of conserved protein sequence motifs found from sequence studies. These also provide the structural basis for the unique materials properties of the different silks. The model shows four GPGXX motifs forming the β-spiral. The long axis of the fiber would be at 45° from left to right. The dashed lines show some of the likely hydrogen bonds, both across the turns and between layers of the spiral.

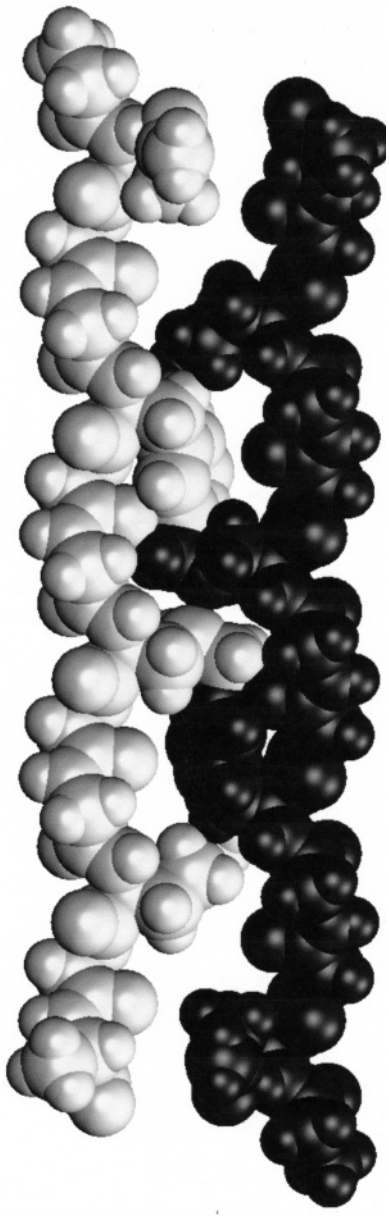

Figure 13. Computer model of the GGX repeat region. The model is a space-filling energy-minimized antiparallel two-strand GGS region. The starting configuration was a Gly II helix for both strands.

In natural spider silks, proteins with no elastic segments show $<5\%$ elasticity, whereas proteins with 9 motifs have $20-30\%$ elasticity and those containing up to 60 elastic motifs have $>200\%$ elasticity. Thus, there appears to be a relationship of increasing motif number to increased elasticity, as we predicted. The length and number of the crystalline strength motifs are less variable in the spider silk proteins, but polyAla-containing proteins show tensile strengths 4-fold higher than those with polyGly-Ala motifs.

As described previously, the other extraordinary property of the dragline silks from most orb weavers is supercontraction. This results in more rubberlike mechanical properties with increased elasticity and decreased tensile strength and stiffness. The structural changes in the proteins within the fiber that allow this to occur are still not fully understood. Raman spectra of supercontracted silk suggest an increase in random coil structure with a decrease in β-sheets. However, NMR studies by Jelinski et al.[50] and van Beek et

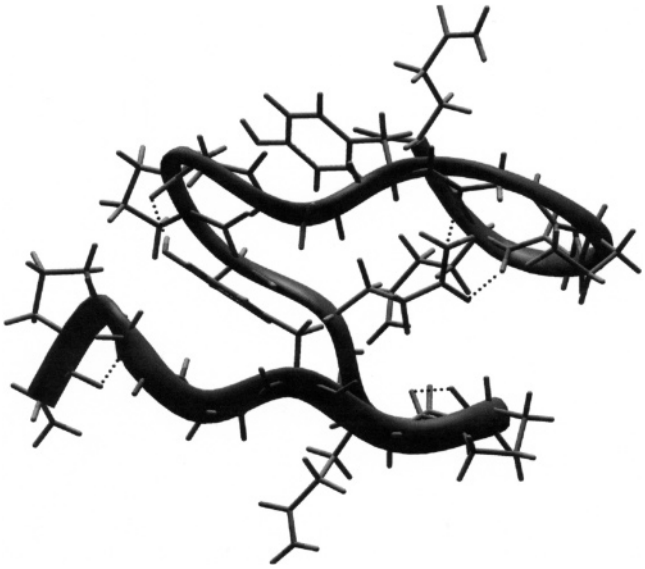

Figure 14. Computer model of a β-spiral. The model is an energy-minimized sequence of $(GPGGQGPGGY)_2$. The starting configuration was type II β-turns at each pentamer sequence.

al.[51] showed little change in the β-sheet regions as did fiber X-ray diffraction data from Parkhe et al.[30]

A recent NMR study[52] provides more information concerning the molecular details of the effects of water on the fiber proteins. The broad 1H line widths observed in the NMR spectrum of the dry *N. clavipes* dragline silk showed that the protein chain is completely rigid prior to hydration. In wet dragline the β-sheets remain rigid, but significant chain motion occurs in the glycine-rich region and a newly observed alanine helical environment. The mobile alanine region is assigned to a loose helical environment that probably links the crystalline and amorphous domains and could facilitate the crystallite reorientation observed by fiber X-ray diffraction. The appearance of this mobile alanine region in the NMR spectrum of the wet silk could explain the increased elasticity and decreased stiffness observed for wet, supercontracted dragline silk. It was shown that this process is reversible as the new resonance disappears when the silk is allowed to dry and all of the 1H line widths broaden. This changeable alanine environment is likely the GAG repeat that terminates the polyAla runs. The reversible conversion of this protein segment between β-sheet and helical conformation could provide the driving force for the supercontraction process.

The morphology and size of the β-sheet crystallites presumably influence the mechanical properties of the silk. There is a large discrepancy in the literature regarding the actual crystallite domain size in both the native and supercontracted states. Transmission electron microscopy (TEM) methods measure crystalline domain sizes on the order of 70−500 nm, whereas wide-angle X-ray diffraction (WAXD) techniques estimate significantly smaller crystallites, having dimensions that are $2 \times 5 \times 7$ nm. This discrepancy between WAXD and TEM results has been explained as due to a nonperiodic lattice (NPL) where the larger domains observed with TEM are not completely ordered but are made up of the smaller, well-ordered polyAla crystallites that are prevalent from WAXD. Molecular models favor the smaller crystallites. Spin−diffusion NMR experiments done as part of the supercontraction study above estimate average polyAla crystalline domains that are 6 ± 2 nm, in excellent agreement with the WAXD measurements on dry silk, which shows that the crystallites remain intact when the silk is in contact with water. The small crystallites measured indicate that the water enters the larger domains observed by TEM, proving that these domains are in fact polycrystalline.

4. Synthetic Gene and Bacterial Expression Studies

Several groups have used synthetic genes based on spider silk to produce proteins in bacteria. Kaplan's group[53] expressed several different gene constructs based on our major ampullate silk protein sequences. The DuPont group has also published a method for the production of synthetic spider silk proteins similar to ours.[54] Other groups have shown that the spider silk protein can be produced in plants, although the protein levels are not very high.[55]

The sequences above are the basis for designing elastic proteins and biomaterials from them. We have constructed synthetic genes encoding 4, 8, 16, and 32 units of the consensus repeat sequence of MaSp2.[56] This method starts with a synthetic DNA representing the consensus sequence for the silk protein. This DNA sequence is then repeatedly doubled using compatible but nonregenerable restriction enzymes until the desired number of repeats is reached. Each of these constructs has been shown to produce a protein of the expected size in *Escherichia coli*. We have overexpressed the 16 repeat protein in *E. coli* to a level that has allowed us to generate >25 g of purified protein from a series of 10 L cultures. We have also constructed synthetic genes encoding the MaSp1 protein consensus repeats and have expressed that protein.

The recent paper from Nexia Biotechnologies in collaboration with the Natick Army Laboratory[57] demonstrates that genes were expressed at high levels in their mammalian cell culture system. In addition, they have developed methodologies to make the fibers. This system is water-based, gives the best materials properties available, and can be used on small quantities of silk. They were able to achieve reasonable tensile strengths that, combined with the higher than native silk elongation, gave breaking properties similar to those of the natural dragline silk. However, these fibers were composed only of MaSp2.

5. Biological Testing

Very few studies of biological testing of spider silk have been done in a rigorous manner. There is a large body of folklore concerning the antibiotic, wound-healing, and clot-inducing activity of spider webs. However, much of that lore has not been seriously tested. There are two recent publications that deal with the response of animals to implantation of natural spider silks.[58,59] Both of these publications show that the natural spider silks do not induce an immune response whether implanted subcutaneously or intramuscularly in rats, mice, or pigs. In addition, the Vadlamudi paper shows that the spider silks tested show no antibiotic activity toward several bacteria and that the tensile strength shows no changes after 90 days of incubation in rat plasma.

In unpublished studies of ours conducted with U.S. Surgical the findings of these two papers were reinforced. Their studies showed no tissue reaction greater than the bare polyethylene rod control that the silk was wrapped around for implantation. In a curious finding, with the 10 male rats there was a lower tissue response to the wrapped rod than

the control rod that was not seen in the female rats. This is likely a statistical fluke but further emphasizes the benign nature of the implanted spider silks.

6. Summary

The past 15 years has seen a tremendous increase in information about many aspects of spider silk. This includes protein sequences, biophysical studies of the proteins in the fiber, the fiber-spinning process, and evolutionary studies. Despite this progress there are still several key connections needed for a full understanding of spider silk. These include (1) the relationships between protein structure and fiber mechanical properties, (2) the role that the fiber spinning plays in determining the mechanical properties of the fiber, (3) the ability of biotechnology approaches to protein production and fiber spinning to mimic the properties of the natural fibers, and (4) whether commercial applications of these unique fibers will occur.

7. Acknowledgment

I am indebted to the students and postdoctoral fellows who have generated much of the data described here over the past 15 years. The research has been supported by the ARO, NIH, NSF, and AFOSR. I thank the reviewers for their insightful and useful comments and corrections.

8. References

(1) Lucas, F. *Discovery* **1964**, *25*, 20.
(2) Vollrath, F. *Sci. Am.* **1992**, *266*, 70.
(3) Gosline, J. M.; DeMont, M.; Denny, M. *Endeavour* **1986**, *10*, 37.
(4) Bell, A. L.; Peakall, D. B. *J. Cell Biol.* **1969**, *42*, 284.
(5) Kovoor, J. *Ann. Sci. Nat. Zool. Biol. Anim.* **1972**, *14*, 1.
(6) Vollrath, F.; Knight, D. *Nature* **2001**, *410*, 541.
(7) Willcox, P. J.; Gido, S. P.; Muller, W.; Kaplan, D. L. *Macromolecules* **1996**, *29*, 5109.
(8) Work, R. W. *Textile Res. J.* **1977**, *47*, 650.
(9) Iizuka, E. *Experientia* **1983**, *39*, 449.
(10) Madsen, B.; Shao, Z. Z.; Vollrath, F. *Int. J. Biol. Macromol.* **1999**, *24*, 301.
(11) Work, R. W. *J. Exp. Biol.* **1985**, *118*, 379.
(12) Fischer, E. *Hoppe-Seyler's Z. Physiol. Chem.* **1907**, *53*, 440.
(13) Xu, M.; Lewis, R. V. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7120.
(14) Hinman, M. B.; Lewis, R. V. *J. Biol. Chem.* **1992**, *267*, 19320.
(15) Beckwitt, R.; Arcidiacono, S. *J. Biol. Chem.* **1994**, *269*, 6661.
(16) Guerette, P. A.; Ginzinger, D. G.; Weber, B. H. F.; Gosline, J. M. *Science* **1996**, *272*, 112.
(17) Gatesy, J.; Hayashi, C. Y.; Motriuk, D.; Woods, J.; Lewis, R. *Science* **2001**, *291*, 2603.
(18) Tian, M.-Z.; Liu, C.-Z.; Lewis, R. *Biomacromolecules* **2004**, *5*, 657.
(19) Colgin, M. A.; Lewis, R. V. *Protein Sci.* **1998**, *7*, 667.
(20) Mita, K.; Ichimura, S.; James, T. C. *J. Mol. Evol.* **1994**, *38*, 583.
(21) Hayashi, C. Y.; Lewis, R. V. *J. Mol. Biol.* **1998**, *275*, 773.
(22) Hayashi, C. Y.; Lewis, R. V. *Science* **2000**, *287*, 1477.
(23) Hayashi, C. Y.; Blackledge, T. A.; Lewis, R. V. *Mol. Biol. Evol.* **2004**, *21*, 1950.
(24) Sezutsu, H.; Yukohiro, K. *J. Mol. Evol.* **2000**, *51*, 329.
(25) Fedic, R.; Zurovec, M.; Sehnal, F. *Sericology* **2002**, *71*, 1.
(26) Hayashi, C. Y.; Shipley, N. H.; Lewis, R. V. *Int. J. Biol. Macromol.* **1999**, *24*, 271.
(27) Simmons, A.; Michal, C. A.; Jelinski, L. W. *Science* **1996**, *271*, 84.
(28) Gosline, J. M.; Guerrette, P. A.; Ortlepp, C. S.; Savage, K. N. *J. Exp. Biol.* **1999**, *203*, 3295.
(29) Tian, M.-Z.; Lewis, R. V. *Biochemistry* **2005**, *44*, 8006.
(30) Parkhe, A. D.; Seeley, S. K.; Gardner, K.; Thompson, L.; Lewis, R. V. *J. Mol. Recognit.* **1997**, *10*, 1.
(31) Barghout, J. Y. J.; Thiel, B. L.; Viney, C. *Int. J. Biol. Macromol.* **1999**, *24*, 211.
(32) Stauffer, S. L.; Coguill, S. L.; Lewis, R. V. *Journal of Arachnology* **1994**, *22*, 5.
(33) Sponner, A.; Unger, E.; Grosse, F.; Weisshart, K. *Biomacromolecules* **2004**, *5*, 840.
(34) Motriuk-Smith, D.; Smith, A.; Hayashi, C. Y.; Lewis, R. V. *Biomacromolecules* **2005**, *6*, 3152.
(35) Zhang, S.-G.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 23.
(36) Riekel, C.; Branden, C.; Craig, C.; Ferrero, C.; Heidelbach, F.; Muller, M. *Int. J. Biol. Macromol.* **1999**, *24*, 179.
(37) van Beek, J. D.; Hess, S.; Vollrath, F.; Meier, B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10266.
(38) Hijirida, D. H.; Do, K. G.; Mical, C.; Wong, S.; Zax, D.; Jelinski, L. W. *Biophys. J.* **1996**, *71*, 3442.
(39) Liivak, O.; Flores, A.; Lewis, R. V.; Jelinski, L. W. *Macromolecules* **1997**, *30*, 7127.
(40) Lazo, N. D.; Downing, D. T. *Macromolecules* **1999**, *32*, 4700.
(41) Asakura, T.; Ito, T.; Okudaira, M.; Kameda, T. *Macromolecules* **1999**, *32*, 4940.
(42) Demura, M.; Minami, M.; Asakura, T.; Cross, T. A. *J. Am. Chem. Soc.* **1998**, *120*, 1300.
(43) van Beek, J. D.; Kummerlen, J.; Vollrath, F.; Meier, B. H. *Int. J. Biol. Macromol.* **1999**, *24*, 173.
(44) Kummerlen, J.; van Beek, J. D.; Vollrath, F.; Meier, B. H. *Macromolecules* **1996**, *29*, 2920.
(45) Asakura, T.; Minami, M.; Shimada, R.; Denura, M.; Osanai, M.; Fujito, T.; Imanari, M.; Ulrich, A. S. *Macromolecules* **1997**, *30*, 2429.
(46) Pappu, R. V.; Rose, G. D. *Protein Sci.* **2002**, *11*, 2437.
(47) Ma, B.; Nussinov, R. *Protein Sci.* **2002**, *11*, 2335.
(48) Tatham, A. S.; Shewry, P. R. *Philos. Trans. R. Soc. London Ser. B, Biol. Sci.* **2002**, *357*, 229.
(49) Urry, D. W.; Hugel, T.; Seitz, M.; Gaub, H. E.; Sheiba, L.; Dea, J.; Xu, J.; Parker, T. *Philos. Trans. R. Soc. London, Ser. B* **2002**, *357*, 169.
(50) Jelinski, L. W.; Blye, A.; Liivak, O.; Michal, C.; LaVerde, G.; Seidel, A.; Shah, N.; Yang, Z. *Int. J. Biol. Macromol.* **1999**, *24*, 197.
(51) van Beek, J. D.; Hess, S.; Vollrath, F.; Meier, H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10266.
(52) Holland, G. P.; Lewis, R. V.; Yarger, I. L. *J. Am. Chem. Soc.* **2004**, *126*, 5867.
(53) Prince, J. T.; McGrath, K. P.; DiGirolamo, C. M.; Kaplan, D. L. *Biochemistry* **1995**, *34*, 10879.
(54) Fahnestock, S. R.; Irwin, S. L. *Appl. Microbiol. Biotechnol.* **1997**, *47*, 23.
(55) Scheller, J.; Guhrs, K.-H.; Grosse, F.; Conrad, U. *Nat. Biotechnol.* **2001**, *19*, 573.
(56) Lewis, R. V.; Hinman, M.; Kothakota, S.; Fournier, M. *Protein Express. Purif.* **1996**, *7*, 400.
(57) Lazaris, A.; Arcidiacono, S.; Huang, Y.; Zhou, J.-F.; Duguay, F.; Chretien, N.; Welsh, E. A.; Soares, J. W.; Karatzas, C. N. *Science* **2002**, *295*, 472.
(58) Vollrath, F.; Barth, P.; Basedow, A.; Engstronm, W.; List, H. *In Vivo* **2002**, *16*, 229.
(59) Vadlamudi, S.; Hussein, M.; Lewis, R. V.; Ferris, C. *Pharm. Res.* **1995**, *12*, 235.

CR010194G