Molecular and Mechanical Characterization of Aciniform Silk: Uniformity of Iterated Sequence Modules in a Novel Member of the Spider Silk Fibroin Gene Family

Cheryl Y. Hayashi, Todd A. Blackledge,* and Randolph V. Lewis†*

*Department of Biology, University of California, Riverside; and †Department of Molecular Biology, University of Wyoming

Araneoid spiders use specialized abdominal glands to produce up to seven different protein-based silks/glues that have diverse physical properties. The fibroin sequences that encode aciniform fibers (wrapping silk) and the mechanical properties of these fibers have not been characterized previously. To gain a better understanding of the molecular radiation of spider silk fibroin genes, cDNA libraries derived from aciniform glands of the banded garden spider, *Argiope trifasciata*, were constructed, and unique silk transcripts were sequenced. There was evidence for a single silk fibroin gene that was expressed in the aciniform glands, and the inferred amino acid composition of the novel fibroin closely matched the amino acid contents of these glands. The inferred protein, aciniform spidroin 1 (AcSp1), is composed of highly homogenized repeats that are 200 amino acids in length. The long stretches of poly-alanine and glycine-alanine subrepeats, which are thought to account for the crystalline regions of minor ampullate and major ampullate fibers, are very poorly represented in AcSp1. The AcSp1 repeat unit is iterated minimally 14 times and does not display substantial sequence similarity to any previously described genes or proteins. Database searches, however, showed that the nonrepetitive carboxy-terminus contains stretches of matches to known spider fibroin sequences, suggesting that the AcSp1 gene is a highly divergent member of the spider silk gene family. In phylogenetic analyses of carboxy-terminal sequences from araneid spiders, the aciniform sequence did not group strongly with clusters of fibroins from the flagelliform, minor ampullate, or major ampullate silk glands. Comparisons of stress/strain curves for major ampullate, minor ampullate, and aciniform silks from *Argiope trifasciata* showed significant differences in ultimate strength, extensibility, and toughness. Remarkably, the toughness of aciniform silk was 50% greater than the highest values typically recorded for major ampullate silk. These differences in performance, in combination with the radical divergence at the sequence level among fibroin paralogs, suggest a possible linkage between silk fibroin sequences and performance that should be explored in future structural/functional studies of aciniform silk.

Introduction

Derived orb-web weaving spiders (Araneoidea) use specialized sets of abdominal silk glands to manufacture up to seven types of fibers and glues (Foelix 1996). Empirical studies have shown that these secretions are encoded, at least in part, by different members of the spider silk gene family (Guerette et al. 1996; Gatesy et al. 2001; Hayashi 2002). Of the seven different sets of silk glands in a typical araneoid, fibroin complementary DNAs (cDNAs) have been characterized from only four glandular types: major ampullate (produces dragline and frame silk; Xu and Lewis 1990; Hinman and Lewis 1992; Beckwitt and Arcidiacono 1994; Guerette et al. 1996; Beckwitt, Arcidiacono, and Stote 1998; Gatesy et al. 2001), minor ampullate (makes temporary capture spiral silk; Guerette et al. 1996; Colgin and Lewis 1998), flagelliform (synthesizes core fiber of the capture spiral; Hayashi and Lewis 1998; Gatesy et al. 2001; Becker et al. 2003), and tubuliform (generates egg case silk; Guerette et al. 1996). Flagelliform and minor ampullate fibroin genes are not known to be expressed in other glands, but Northern blot experiments suggest that some gene family members are present in more than one gland type (Guerette et al. 1996). No cDNA, gene, or protein sequences have been described for aciniform (wrapping silk), aggregate (sticky glue), and piriform (attachment disc) silks.

The cDNAs of all the published araneoid silks have a similar structural organization. The transcripts are long (~4–16 kilobases; Hayashi, Shipley, and Lewis 1999) and highly internally repetitive with a relatively conserved, nonrepetitive, carboxy-terminal region. Because of extensive divergence in the repetitive regions of different silk gene paralogs, sequence similarities in the 3′ carboxy-terminus have been used to identify members of the spider silk gene family (Beckwitt and Arcidiacono 1994; Guerette et al. 1996; Gatesy et al. 2001; Hayashi 2002). Comparative analyses of the described spider fibroin cDNAs suggested sequence conservation within particular ortholog groups, with extensive divergence among paralogous gene copies (Gatesy et al. 2001; Hayashi 2002). This pattern is consistent with ancient functional diversification of gene duplicates and subsequent conservation of critical sequence motifs. Not only are the repetitive sequences that characterize each fibroin highly diverged in comparisons among paralogs, but the mechanical properties of the spun silk fibers from different glands also are notably dissimilar from each other (Stauffer, Coguill, and Lewis 1994; Köhler and Vollrath 1995; Gosline et al. 1999).

Four simple amino acid motifs, in various combinations and frequencies, explain most of the diversity in known araneoid silk fibroin sequences (Hayashi, Shipley, and Lewis 1999; Gatesy et al. 2001). These are: (1) poly-alanine (A_n), (2) alternating glycine and alanine couplets $(GA)_n$, (3) triplets composed of two glycines followed by a variable amino acid $(GGX)_n$, and (4) glycine-proline-glycine containing units ($GPGX_n$). The four small subrepeats, arranged in combinations that are characteristic of particular fibroins, form larger ensemble repeat units that are themselves iterated many times.

The amino acid compositions of aciniform, aggregate, piriform, and tubuliform silks have substantially less

Key words: aciniform silk, *Argiope trifasciata*, cDNA, fibroin, gene family, spider.

E-mail: cheryl.hayashi@ucr.edu.

Mol. Biol. Evol. 21(10):1950–1959. 2004
doi:10.1093/molbev/msh204
Advance Access publication July 7, 2004

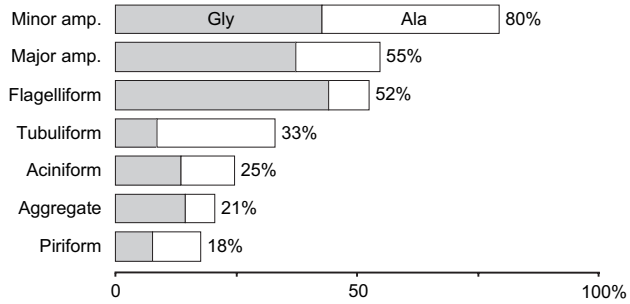


FIG. 1.—The minor ampullate (amp.), major ampullate, flagelliform, tubuliform, aciniform, aggregate, and piriform silk glands of *Araneus diadematus* differ in their amino acid compositions. The sum of the percentages of glycine (gray rectangle) and alanine (white rectangle) within a gland type are indicated to the right of each bar (data from Andersen 1970). One cDNA, the glycine-rich ADF-2 (Guerrette et al. 1996), has been associated with the tubuliform glands of *A. diadematus*. However, the predicted amino acid composition of ADF-2 does not correspond well with the published amino acid composition of tubuliform silk (Guerrette et al. 1996; also see supplementary material for Gatesy et al. 2001). Thus, tubuliform silk likely includes at least one additional fibroin protein that has not been described yet.

glycine and alanine relative to silks from the major, minor, and flagelliform glands. This suggests that the four simple glycine- and alanine-rich amino acid motifs (above) that explain most of the variation in major ampullate, minor ampullate, and flagelliform silks cannot account for much of the protein sequence in the other araneoid silks (fig. 1). Three types of araneoid silks have not been described at the sequence level. Of these, aggregate silk is an adhesive, and piriform silk is used to affix dragline silk to substrates. In contrast, aciniform silk is fibrous and emerges from multiple spigots on the posterior median spinnerets (PMS) and posterior lateral spinnerets (PLS). Araneoid spiders use aciniform silk to wrap and immobilize prey, construct web decorations, build sperm webs, and encase eggs (Foelix 1996).

Given the varied functions of this silk and its amino acid composition with a relatively low percentage of glycine and alanine (fig. 1), we hypothesized that unique repeat modules should compose aciniform fibroins. Here, we characterized silk cDNAs derived from the aciniform glands of *Argiope trifasciata* (Araneidae), the banded garden spider. To better understand the diversity of sequences that encode fibrous silks, we compared the aciniform cDNA sequences to other known araneid silk genes and performed mechanical tests to describe some of the physical differences between major ampullate, minor ampullate, and aciniform silk fibers spun by *A. trifasciata*. Characterization of divergent fibroin sequences at the molecular and mechanical levels is the first step toward a better understanding of how gene duplication, repetitive silk protein sequence, and the spinning process (Vollrath and Knight 2001) determine the diverse, and sometimes extraordinary, mechanical properties of spider silk.

Materials and Methods

cDNA Library Construction and Screening

A. trifsaciata were collected from Wheatland, Wyoming (Platte County). Aciniform silk glands attached to the PMS and PLS were dissected from euthanized spiders and flash frozen in liquid nitrogen. Messenger RNA (mRNA) was extracted from the glands using oligo-$(dT)_{25}$ coupled to magnetic beads (Biotech Inc., Brown Deer, Wisc.). cDNA was synthesized with the SuperScript II Choice protocol (Invitrogen, San Diego, Calif.) using oligo-$(dT)_{18}$ as the primer for first strand synthesis. The cDNA was passed through a ChromaSpin 1,000 column (Clontech, Palo Alto, Calif.) to select for large fragments. The cDNA was then blunt-end ligated into pZErO-2 (Invitrogen) that had been cut with *Eco*RV (New England Biolabs, Beverly, Mass.). TOP10 electrocompetent cells (Invitrogen) were transformed with the ligation, and separate libraries were made for the PMS and PLS aciniform cDNA. From each library, approximately 1,200 recombinant colonies were replicated onto nylon membranes for screening.

The libraries were sequentially screened with a set of γ^{32}P-labeled oligonucleotide probes (5′-CCWAYWCC-NCCATATCCWCC-3′, 5′-CCWCCWGGWCCNNNW-CCWCCWGGWCC-3′, 5′-CCWGGWCCTTGTTGW-CCWGGWCC-3′, 5′-GCDGCDGCDGCDGCDGC-3′, 5′-CCWGCWCCWGCWCCWGCWCC-3′, and 5′-CCA-GADAGACCAGGATTACT-3′) that correspond to short amino acid sequence motifs that have been found to be conserved among spider silks (Hayashi, Shipley, and Lewis 1999; Gatesy et al. 2001). Additionally, a sampling of colonies from each library were screened for insert size. Because the known spider silk transcripts are all exceptionally long (Hayashi Shipley, and Lewis 1999), inserts greater than 3,000 base pairs (bp) were sequenced. Two-hundred forty clones with shorter inserts also were sequenced to verify that the short clones did not contain silk transcripts (C. J. Vink, personal communication). Based on the findings of this preliminary sequencing, two aciniform fibroin specific oligonucleotides, 5′-CGAAGAAGCTGATGCCT-GAGAGTAAC-3′ and 5′-TCCAGTGGAAGGTCCAG-AAGGTCCTG-3′, were synthesized and used to probe both libraries. A subset of the identified silk cDNA inserts were sequenced in their entirety using the transposon-based GPS-1 Genome Priming System (New England Biolabs). It was necessary to use transposons because the silk cDNAs were too long to characterize using only the T7 and SP6 universal primer sites on the plasmid vector. Furthermore, the repetitive nature of the inserts precluded the possibility of using internal primers for sequencing.

Comparative Analyses

To identify potentially homologous proteins and genes, Blast (www.ncbi.nlm.nih.gov/) searches were performed with the aciniform cDNA and with translated cDNA sequences. The default query parameters were used for nucleotide and protein Blast searches against the nr databases. Separate searches were done with the repetitive region and the nonrepetitive carboxy-terminus.

The final 99 amino acids of the translated aciniform cDNA sequence were compared to the corresponding carboxy-terminal regions of published silk fibroins from araneid spiders. The amino acid sequences were aligned with ClustalW (MacVector 7.2, Accelrys Inc., San Diego, Calif.) using the identity matrix and a gap penalty of one.

One published sequence, MaSp1 from *Argiope aurantia* (GenBank accession number AF350262) was not used because it did not include over half of the conserved carboxy-terminal region. The DNA sequences that encode the c-termini were aligned according to the amino acid alignment.

Phylogenetic analyses of the data sets were performed with PAUP* (Swofford 2002). Branch-and-bound parsimony searches were done with gaps treated as missing data and also with gaps treated as characters. For maximum-likelihood analysis, the best fitting model of sequence evolution, among 56 possibilities, was chosen using likelihood ratio tests implemented in ModelTest (Posada and Crandall 1998), and a heuristic likelihood analysis was executed using PAUP* (100 random taxon addition replicates with tree-bisection-reconnection (TBR) branch swapping). Bootstrap resampling was used to assess the robustness of nodes (Felsenstein 1985). For parsimony analyses of the amino acid and nucleotide data sets, uninformative characters were excluded and 1,000 bootstrap replicates were performed (branch and bound searches); one hundred replicates were done in the maximum-likelihood bootstrap analysis (heuristic searches with simple taxon addition and TBR branch swapping).

Amino Acid Composition Analyses

Silk protein was extracted from the aciniform glands by grinding the tissue in 2% sodium dodecyl sulfate buffer, followed by two rinses with acetone. The silk protein was then hydrolyzed in 6 N hydrochloric acid, and the relative amount of each amino acid was determined with the AccQ-Tag chemistry and method (Waters Corp., Milford, Mass.). Six compositions were determined: three extracts each from the PLS and PMS aciniform glandular fibroins.

Mechanical Tests

Silk samples were obtained from two adult female *A. trifasciata* collected as juveniles and maintained in the laboratory until they matured after 2–3 instars. The spiders were maintained in 30×30×5 cm cages and allowed to build webs freely. They were fed a diet of crickets (*Gryllus* sp.) and honeybees (*Apis mellifera*) and were misted with water daily. Samples of aciniform silk were obtained by first inducing the spiders to wrap prey that were placed into webs. Then "c"-shaped cards covered with double-sided tape were inserted between the spiders and prey during wrapping such that the spiders wrapped several swaths of silk onto the cards. Under a dissecting microscope, insect minuten pins were then used to tease one to two fibers of silk from these sheet-like swaths. The fibers were then affixed to "c"-shaped cardboard sample mounts, across 10 mm gaps, using fast-drying cyanoacrylate glue (Super Glue™). A total of eight samples were collected from one spider and three samples were collected from the other spider. An additional three samples of wrapping silk from the closely related species *Argiope argentata* were collected using the same protocol.

Seven single-stranded samples of major ampullate silk were collected from each of the two *A. trifasciata* by forcibly silking them (prior to silking, spiders were briefly anesthetized with CO_2). Major ampullate silk samples were secured to cardboard mounts, across 21 mm gaps, using cyanoacrylate glue. Eight samples of double-stranded minor ampullate silk were collected from one of the *A. trifasciata* during the same silking session as the major ampullate samples. The second *A. trifasciata* did not produce minor ampullate fibers, despite repeated attempts at silking. An additional three samples of double stranded minor ampullate silk were collected from a third spider, an immature *A. trifasciata*, shortly after it was collected from the field.

The diameter of each fiber was measured using a Leica DLMB polarizing light microscope. Three digital photos were taken of each fiber at 1400× magnification, one near either end of the fiber and the third near the middle. NIH Image 1.63 (U.S. National Institutes of Health) was then used to determine the diameter of each fiber in three locations for each photograph. The resulting nine diameter estimates for each fiber were averaged and the cross-sectional area of each fiber was calculated by assuming the fibers are circular in cross-section.

Tensile tests were performed on samples of aciniform, major ampullate, and minor ampullate silk using a Nano Bionix tensile tester (Systems Corp., Oak Ridge, Tenn.). Fibers were extended at a constant rate of 1% strain/s, relative to their original lengths, until the samples failed. Engineering stress and strain were calculated from the force and extension data. Testworks 4.0 software (MTS Corp.) was used to visualize the resulting stress-strain curves, to calculate stiffness (Young's modulus E), and to calculate toughness by integrating the area under the stress-strain curve. One-way ANOVAs (Statistica 6.1, Statsoft, Inc., Tulsa, Okla.) were used to test for differences between the types of silks for ultimate strength, extensibility, toughness, and stiffness.

Results and Discussion

Characterization of AcSp1

No silk fibroin clones were discovered with the battery of probes designed from previously characterized araneoid fibroins (Gatesy et al. 2001). Instead, silk transcripts were identified in both the PLS and PMS aciniform gland cDNA libraries by size screening and subsequent hybridization with aciniform fibroin specific oligonucleotides. The 59 silk clones (34 from the PLS library, 25 from the PMS library) contained similar sequences, and like all previous spider silk cDNA studies (Xu and Lewis 1990; Hinman and Lewis 1992; Guerette et al. 1996; Beckwitt, Arcidiacono, and Stote 1998; Colgin and Lewis 1998; Hayashi and Lewis 1998; Gatesy et al. 2001), the cDNA inserts were partial-length. Based on selective sequencing and restriction enzyme fragment length analysis, the 59 silk clones appeared to differ only in length, primarily in the extent of sequence toward the 5′ end. In regions where the clones overlapped, there were no nucleotide differences, even in comparisons across libraries. These findings suggest that the same silk gene is expressed in both PLS and PMS aciniform glands.

The longest aciniform silk cDNA clone was 8,618 bp in length and was sequenced in its entirety (GenBank

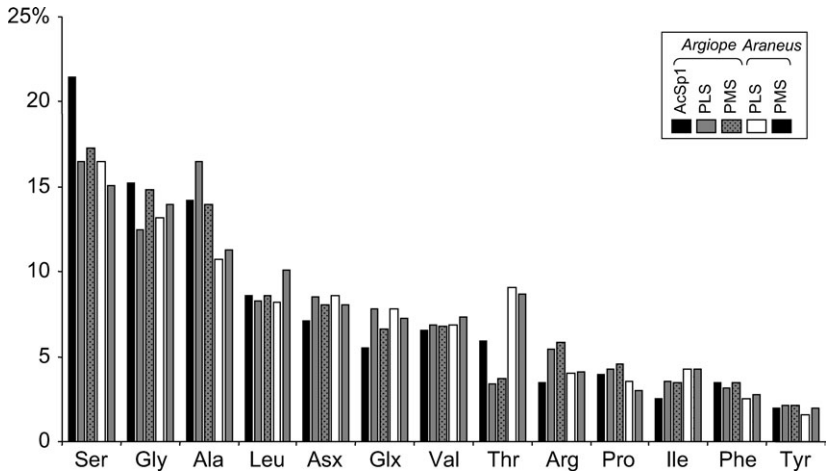


FIG. 2.—The predicted amino acid composition of AcSp1 (black) and the amino acid compositions of the aciniform glandular silk proteins from *Argiope trifasciata* (gray) and *Araneus diadematus* (white) are shown with bar graphs. For the glandular protein extracts, compositions from the aciniform glands associated with the posterior lateral spinnerets (PLS; solid bar) and posterior median spinnerets (PMS; speckled bar) are shown separately. Only the most abundant amino acids, which account for at least 97% of the total composition, are indicated. Note that because of the acid hydrolysis in the analytical method, Asn and Asp are indistinguishable (Asx), as are Gln and Glu (Glx).

accession number AY426339). Additional aciniform silk clones were completely sequenced and were nested within the 8,618-bp sequence with 100% nucleotide identity. The 8,618-bp transcript could be translated in just one reading frame without interruption from bases 1 through 8,505. The encoded fibroin is hereafter referred to as AcSp1, an abbreviation of "aciniform spidroin 1." This name follows the convention of MaSp1 (major ampullate spidroin 1; Xu and Lewis 1990) and MaSp2 (major ampullate spidroin 2; Hinman and Lewis 1992), where "spidroin" is a contraction of "spider fibroin."

The predicted amino acid composition of AcSp1 generally agreed with the composition of the PLS and PMS aciniform silk glands from *Argiope* (fig. 2). This correspondence was consistent with AcSp1 being a major constituent of the protein stored in the aciniform silk glands. There was a slight overabundance of serine and threonine in AcSp1 in comparison to the silk glands. This might have been due to additional proteins in the aciniform glands, to the contribution of uncharacterized sequence towards the amino-terminus of AcSp1, and/or to experimental error in the chemically determined glandular compositions. The similarity of the protein extracts from the PLS and PMS aciniform silk glands of *Argiope* is akin to the similarity observed by Andersen (1970) in the two sets of aciniform silk glands from the con-familial *Araneus diadematus* (Araneidae; fig. 2).

AcSp1 consists of a series of 200 amino acid long repeats that ends with a 99 amino acid long, nonrepetitive carboxy-terminus (fig. 3). In Blast searches with the AcSp1 carboxy-terminus, the top matches were to published spidroins. The other, lower probability matches were to two predicted proteins, one in rat (*Rattus norvegicus*; GenBank accession number XP_342841) and the other in rice (*Oryza sativa*; GenBank accession number AAR06336). The similarity of the carboxy-termini of AcSp1 with other spidroins and the repetitive genetic architecture of the AcSp1 cDNA (fig. 3) suggest that AcSp1 is a member of the spider silk fibroin gene family. The repetitive regions of spider silk fibroins do not allow precise hypotheses of basepair-to-basepair homology among paralogs, therefore an alignment of the nonrepetitive carboxy-terminal region was used to construct a fibroin gene tree for *A. trifasciata* and its close relatives, spiders from the family Araneidae. This inference of the evolutionary history of the gene family assumed that there has not been recombination among fibroin paralogs. Among characterized araneid fibroins, the AcSp1 carboxy-terminal is divergent, and it did not group robustly with other fibroin ortholog groups (fig. 4). Currently, precise rooting of this gene tree is difficult because spider fibroin sequences are poorly sampled (fig. 4).

Highly Similar AcSp1 Repeat Units

Although the nonrepetitive carboxy-terminus of AcSp1 gives some insights into the evolutionary history of aciniform silk, the repetitive portion of AcSp1 makes up the bulk of the sequence (fig. 3*a*). In pairwise comparisons, the individual repeat units are remarkably conserved at both the DNA and amino acid sequence levels (fig. 3*b*). Several of the repeats share an astounding 100% identity to each other. Although the repeat unit is 600 bp long, there were only 16 variable sites and one six-bp indel in the alignment of 14 repeats, with most of the variation clustered in the 3′ region of the terminal repeat (fig. 3*b*). There were minimally eight nonsynonymous differences among the repeats, not including the indel. Homogeneity of repeat units within a gene has been observed in all other characterized spider fibroins (summarized in Gatesy et al. 2001) and in convergent fibroins from lepidopterans (Mita, Ichimura, and James 1994; Sezutsu and Yukuhiro 2000; Fedič, Žurovec, and Sehnal 2003), but the very 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 (fig. 3*b*).

It could be argued that the conservation of repeats within AcSp1 is due to ancient duplications of the AcSp1

a.

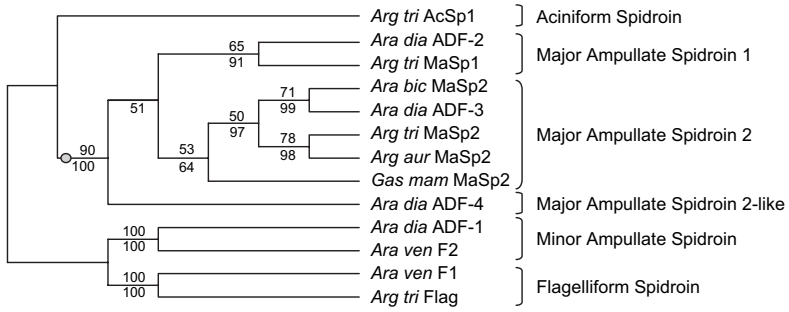
b.

```
                Gly Ser Ala Gly Pro Gln Gly Gly Phe Gly Ala Thr Gly Gly Ala Ser Ala Gly Leu Ile Ser Arg Val Ala Asn Ala Leu Ala Asn Thr Ser Thr Leu Arg
Consensus       GGA TCT GCT GGC CCT CAA GGT GGA TTC GGT GCC ACA GGT GGA GCG TCT GCT GGC CTT ATC TCC AGA GTA GCA AAC GCA CTT GCC AAT ACA TCA ACA TTG AGA
Repeat_1
Repeat_2        ... ... .A. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_3        ... ... .A. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_4        ... ... .A. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_6        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_7        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_8        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_9        ... ... .A. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_10       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_11       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_12       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_13       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_14       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                Thr Val Leu Arg Thr Gly Val Ser Gln Gln Ile Ala Ser Ser Val Val Gln Arg Ala Ala Gln Ser Leu Ala Ser Thr Leu Gly Val Asp Gly Asn Asn Leu
Consensus       ACT GTC CTC AGA ACT GGT GTA TCC CAA CAG ATT GCC TCC AGC GTG GTA CAG AGA GCC GCT CAG TCG TTG GCC AGT ACT CTC GGA GTC GAC GGA AAT AAC TTG
Repeat_1
Repeat_2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_4        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_6        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_7        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_8        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_9        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_10       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_11       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_12       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_13       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_14       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                Ala Arg Phe Ala Val Gln Ala Val Ser Arg Leu Pro Ala Gly Ser Asp Thr Ser Ala Tyr Ala Gln Ala Phe Ser Ser Ala Leu Phe Asn Ala Gly Val Leu
Consensus       GCC AGA TTC GCG GTA CAG GCC GTC TCT CGA CTG CCC GCC GGA TCA GAC ACT TCT GCT TAC GCT CAA GCA TTC TCT AGT GCG CTC TTC AAT GCC GGA GTT CTC
Repeat_1                                                                                                        ... ... ... ... ... ... ... ... ... ...
Repeat_2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_4        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_6        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_7        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_8        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_9        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_10       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_11       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_12       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_13       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_14       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                Asn Ala Ser Asn Ile Asp Thr Leu Gly Ser Arg Val Leu Ser Ala Leu Leu Asn Gly Val Ser Ser Ala Ala Gln Gly Leu Gly Ile Asn Val Asp Ser Gly
Consensus       AAT GCA AGC AAC ATT GAC ACA TTG GGA TCC CGA GTT CTC TCA GCA CTT TTG AAC GGA GTA TCA AGT GCG GCG CAA GGT CTT GGC ATC AAT GTA GAT AGC GGC
Repeat_1        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_4        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_6        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_7        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_8        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_9        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_10       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_11       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_12       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_13       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_14       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                Ser Val Gln Ser Asp Ile Ser Ser Ser Ser Ser Phe Leu Ser Thr Ser Ser Ser Ser Ala Ser Tyr Ser Gln Ala Ser Ala Ser Ser Thr Ser Gly Ala Gly
Consensus       AGT GTA CAA AGT GAC ATT AGT TCC AGT AGC AGC TTC CTC TCA ACA AGC TCG TCT TCG GCC AGT TAC TCT CAG GCA TCA GCT TCT TCG ACC AGC GGT GCC GGA
Repeat_1        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_3        ... ... ... ... ... ... ... ... ... ..T ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_4        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_6        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... A.. ...
Repeat_7        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_8        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_9        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_10       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_11       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_12       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_13       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_14       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                Tyr Thr Gly Pro Ser Gly Pro Ser Thr Gly Pro Ser Gly Tyr Pro Gly Pro Leu Gly Gly Gly Ala Pro Phe         Gly Gln Ser Gly Phe Gly
Consensus       TAC ACA GGA CCT TCT GGA CCT TCC ACT GGA CCT TCT GGC TAC CCT GGG CCT TTG GGT GGC GGA GCG CCG TTC --- --- GGT CAA TCA GGC TTT GGC
Repeat_1        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_2        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_3        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_4        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_5        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_6        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_7        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_8        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_9        ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_10       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_11       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_12       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_13       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Repeat_14       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
```

a.

```
Arg tri AcSp1       VGLRSASAASRLSQLTSSITNAVGANGVDANSLARSLQSSFSALRSSGMSSSDAKIEVLLETIVGLLQLLSNTQVRGVNPATASSVANSAARSFELVLA-*
Ara dia ADF-2       SRLSSPSAAARVSSAV-SLVSNGGPTSPAALSSSISNVVSQISASNPGLSGCDILVQALLEIISALVHILGSANIGPVNSSSAGQSASIVGQSVYRALS-*
Arg tri MaSp1       SRLSSPGAASRVSSAVTSLVSSGGPTNSAALSNTISNVVSQISSSNPGLSGCDVLVQALLEIVSALVHILGSANIGQVNSSGVGRSASIVGQSINQAFS-*
Ara bic MaSp2       SRLSSSAASSRVSSAVSSLVSSG-PTTPAALSNTISSAVSQISASNPGLSGCDVLVQALLEVVSALVHILGSSSVGQINYGASAQYAQMV??????????
Ara dia ADF-3       SRLSSPAASSRVSSAVSSLVSSG-PTKHAALSNTISSVVSQVSASNPGLSGCDVLVQALLEVVSALVSILGSSSIGQINYGASAQYTQMVGQSVAQALA-*
Arg tri MaSp2       SRLSSPQASSRVSSAVSTLVSSG-PTNPASLSNAISSVVSQVSSSNPGLSGCDVLVQALLEIVSALVHILGSSSIGQINYAASSQYAQLVGQSLTQALG-*
Arg aur MaSp2       SRLSSPQASSRVSSAVSTLVSSG-PTNPAALSNAISSVVSQVSASNPGLSGCDVLVQALLELVSALVHILGSSSIGQINYAAS?????????????????
Gas mam MaSp2       SRLSSPQAGARVSSAVSALVASG-PTSPAAVSSAISNVASQISASNPGLSGCDVLVQALLEIVSALVSILSSASIGQINYGASGQYAAMI??????????
Ara dia ADF-4       SVLSSPAASSRVSSAVSSLVSSG-PTNGAAVSGALNSLVSQISASNPGLSGCDALVQALLELVSALVAILSSASIGQVNVSSVSQSTQMISQALS-----*
Ara dia ADF-1       NRLSSAGAASRVSSNVAAIASAG----AAALPNVISNIYSGVLSS--GVSSSEALIQALLEVISALIHVLGSASIGNVSSVGVNSALNAVQNAVGAYAG-*
Ara ven F2          NRLSSSGAANRVSSNVAAIASGG----AAALPNVMSNIYSGVLGS--GVSSSEALIQALLEVISALMHVLGSASIGNVSSAGLDSTLNVVQNAVSQYAG-*
Ara ven F1          SRFPSLIN-GIMSSMQ------GGGFNYQNFGNVLSQFATGTGTCN--SNDLNLLMDALL---SAL-HTLSYQGMGTVPSYPSPSAMSAYSQSVRRCFGY*
Arg tri Flag        ERLPNLIN-GIKSSMQ------GGGFNYQNFGNILSQYATGSGTCN--YYDINLLMDALL---AAL-HTLNYQGASYVPSYPSPSEMLSYTENVRRYF--*
```

b.



c.

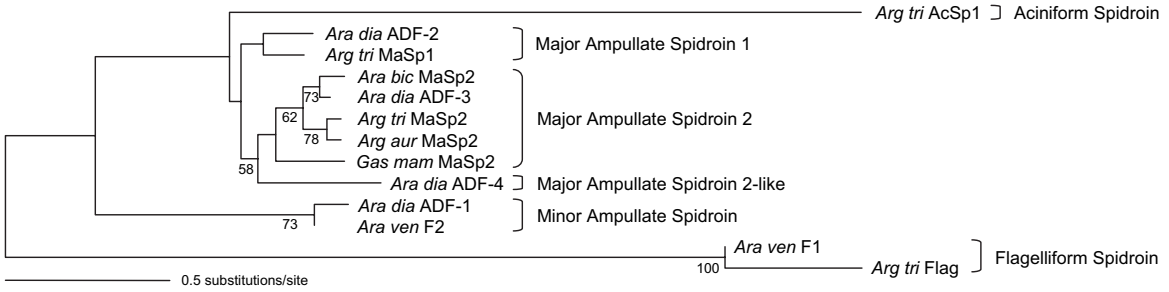


FIG. 4.—(*a*) The alignment of araneid silk fibroin carboxy-terminal sequences is depicted. Amino acids are indicated by one-letter abbreviations, gaps are shown by dashes, and missing data are represented by question marks. The sequences are identified as *Ara bic* (*Araneus bicentenarius*) MaSp2 (GenBank accession number U20328); *Ara dia* (*Araneus diadematus*) ADF-1 through ADF-4 (U47853 through U47856); *Ara ven* (*Araneus ventricosus*) F1 and F2 (AY174110, AY177203); *Arg aur* (*Argiope aurantia*) MaSp2 (AF350263); *Arg tri* (*Argiope trifasciata*) AcSp1 (AY426339), MaSp1 (AF350266), MaSp2 (AF350267), and Flag (AF350264); and *Gas mam* (*Gasteracantha mammosa*) MaSp2 (AF350272). A minor alteration was made in the GenBank entry for *Ara dia* ADF-4; see Hayashi and Lewis (1998) for an explanation of the change. (*b*) Parsimony analysis of the aligned carboxy-terminal amino acid sequences resulted in one minimum length tree. The gene family tree had 277 steps, a consistency index (CI; Kluge and Farris 1969) excluding uninformative characters of 0.8486, and a retention index (RI; Farris 1989) of 0.7765. Bootstrap percentages (Felsenstein 1985) greater than 50% are shown above internodes. The tree is midpoint rooted. Phylogenetic analysis of the aligned DNA sequences also resulted in one minimum length tree of 734 steps, a CI excluding uninformative characters of 0.6701, and a RI of 0.6337. Bootstrap percentages greater than 50% are shown immediately below internodes. The DNA tree differed from the amino acid tree in the placement of the minor ampullate spidroin clade at the internode marked with the gray circle. The resulting clade, an alliance of major and minor ampullate spidroins, was supported with a 61% bootstrap score. (*c*) Maximum-likelihood analysis of the aligned carboxy-terminal DNA sequences resulted in one minimum length tree. The tree is midpoint rooted and had a natural log likelihood of 3021.46973 for the HKY + G model. Branch lengths are scaled to the probability of change per site; note the long branches that connect to the aciniform and flagelliform fibroins. Bootstrap percentages greater than 50% are shown below internodes.

repeat unit, with subsequent conservation of the concatenated DNA sequence by demanding functional constraints. Because there are so few synonymous differences among repeats in this gene, however, stabilizing selection would have to extend to all synonymous nucleotide positions to explain the lack of variation among repeats. Inspection of the AcSp1 DNA sequence suggested some possible constraints on synonymous substitution; extreme codon

←

FIG. 3.—(*a*) AcSp1 contains at least 14 iterations of a single, complex repeat. An individual complex repeat is 200 amino acids in length and cannot be broken down into small subunits as in ensemble repeats from major ampullate, minor ampullate, tubuliform, and flagelliform fibroins. The array of complex repeats in AcSp1 is followed by a conserved, carboxy-terminal region ("C"). The ellipsis (. . .) before repeat 1 indicates additional upstream sequence that was not represented in the cDNA clones. (*b*) The nucleotide sequences of the 14 repeat units are aligned to each other. The first repeat is not complete, and dots signify matches to the majority-rule consensus sequence ("Consensus"). The translation of the consensus sequence is shown above each codon with the three-letter abbreviations of amino acids. The simple amino acid sequence motifs of Gly-Gly-Xaa and poly-Ala, which are abundant in other characterized araneoid fibroins (Hayashi, Shipley, and Lewis 1999; Gatesy et al. 2001) but rare in AcSp1, are underlined. The other repeated motifs, poly-Ser and Thr-Gly-Pro-Ser-Gly, are also underlined. The string of four nucleotide differences that is ten bases from the end of repeat 14 could be interpreted as a single mutational event; in our calculation of the minimum number of differences between repeats, however, these base changes were interpreted as four independent substitutions.

bias was observed for several amino acids. There were 55 tyrosines in the AcSp1 sequence, and TAC codon encoded all of these. The other codon choice, TAT, accounted for none of the tyrosines. Of 247 leucines in the AcSp1 sequence, none was encoded by CTA, and only three were encoded by TTA. The remaining 99% were encoded by TTG, CTT, CTC, and CTG. These biases in codon usage are striking, but the repetitive nature of this fibroin makes interpretations of codon biases problematic. Homogenization of repeat units could accentuate codon bias, by amplifying any small biases in a single unit. For example, an insignificant bias of five TAC codons versus one TAT codon in a single repeat unit could be amplified to a "significant" bias of 100 TAC codons versus 20 TAT codons, if 20 repeat units were homogenized by various evolutionary forces.

Although there is some evidence of constraints on synonymous substitutions within this gene, we suggest that the extreme similarity of repeat units is explained most simply as the result of intragenic gene conversion or unequal crossing-over events (Beckwitt, Arcidiacono, and Stote 1998). This explanation is most consistent with the patterns of sequence variation in other fibroin paralogs from spiders (e.g., Hayashi and Lewis 2000) and moths (e.g., Mita, Ichimura, and James 1994; Sezutsu and Yukuhiro 2000). The sequencing of additional *AcSp1* orthologs should help to clarify the processes that govern the intragenic similarity of AcSp1 repeats.

While the repeat units are highly conserved within AcSp1, Blast (NCBI) searches of the repetitive region found no matches in the nucleotide and protein databases. The Blast analysis also did not reveal any conserved protein domains. In contrast to the carboxy-terminal region, the AcSp1 repetitive region appears to have no substantial similarity to previously characterized genes and proteins. Unlike other silk fibroins, dot plot comparisons of the aciniform consensus repeat sequence to itself uncovered few repetitions of sequence motifs. The most common subrepeat in aciniform fibroin, poly-serine, accounted for only 8.5% of the consensus repeat unit (fig. 3*b*). A five amino acid motif, TGPSG, also is duplicated in the AcSp1 repeat (fig. 3*b*), but aciniform fibroin has a low content of the subrepeats that characterize fibroins from the major ampullate, minor ampullate, flagelliform, and tubuliform glands of araneoid spiders (Hayashi, Shipley, and Lewis 1999; Gatesy et al. 2001). Ensemble repeat units in these fibroins are composed primarily of four simple amino acid motifs, but these common motifs explain very little of the AcSp1 consensus repeat (6.5%; fig. 3*b*). For example, poly-alanine, a motif that has been hypothesized to account for the high tensile strength of major ampullate silk (Simmons, Michal, and Jelinski 1996; Gosline et al. 1999), is notably deficient in AcSp1; there are only two doublets of alanine in the aciniform repeat unit (fig. 3*b*).

Mechanical Properties of Aciniform Silk

The physical properties of aciniform, major ampullate, and minor ampullate silks from *A. trifasciata* were quantified with a mechanical testing system capable of measuring loads on very fine fibers (e.g., fig. 5). One-way ANOVAs demonstrated significant differences among the silks in ultimate strength, extensibility, and toughness, but did not for stiffness (fig. 6). Post hoc comparisons of the mean values for each type of silk using Tukey's Honest Significant Difference Tests for Unequal Sample Sizes revealed that ultimate strength, extensibility, and toughness of aciniform silk were significantly different from the values for major ampullate and minor ampullate silk (P at least <0.005). Major ampullate and minor ampullate silk also differed from one another in both ultimate strength and extensibility (P at least <0.005).

The most striking outcome of the mechanical tests was the extraordinary toughness of aciniform silk. Spider silks are renowned for their ability to absorb energy without failing (toughness). This toughness results from a combination of high strength and high extensibility. Toughness values for major ampullate dragline silk can approach 250 MPa (Köhler and Vollrath 1995). We found that aciniform silk is over 50% tougher than dragline silk. This is largely due to the almost fourfold greater extensibility of aciniform silk relative to major ampullate silk (86 ± 3% versus 22 ± 1%, respectively), rather than the ultimate strength of aciniform silk, which was only half that of major ampullate silk (687 ± 56 MPa vs. 1290 ± 29 MPa, respectively). Thus, aciniform silk is one of the toughest biological materials known (for comparisons see Gosline et al. 2002) and is over seven times as tough as Kevlar, a high-performance synthetic material (Gordon 1988).

Estimating stress from our tensile tests required an accurate measurement of the cross-sectional areas of fibers. We found that the diameters of aciniform fibers were very thin (0.35 ± 0.01 μm) compared to those of major ampullate (3.24 ± 0.10 μm) and minor ampullate (0.69 ± 0.10 μm) silk. Inaccurate measurement of these very fine fibers is a possibility when using light microscopy. Underestimation of diameter would result in increased values for stress and all the quantities calculated from stress, including both strength and toughness. However, comparison of diameters obtained with polarized light microscopy measurements to those obtained from scanning electron micrographs for a variety of silk samples demonstrates that polarized light microscopy yields measurements that average 12% higher than those from scanning electron micrographs (unpublished data). In contrast, to decrease the toughness of aciniform silk to the values published for major ampullate silk (e.g., Stauffer, Coguill, and Lewis 1994; Köhler and Vollrath 1995; Gosline et al. 2002), it would have been necessary to under-measure the diameters of aciniform fibers by 18%, rather than over-measure them.

The toughness and ultimate strength of *A. trifasciata* aciniform silk was also similar to aciniform silk spun by the closely related species *A. argentata*, although the latter silk was stiffer ($T_{11} = 3.1$, $P < 0.005$) and less extensible ($T_{11} = 4.9$, $P < 0.005$). Thus, the high toughness of aciniform that we measured is consistent across species. Finally, our characterization of toughness for major ampullate silk from *A. trifasciata* (mean = 145 ± 5 MPa) was similar to values obtained from studies of major

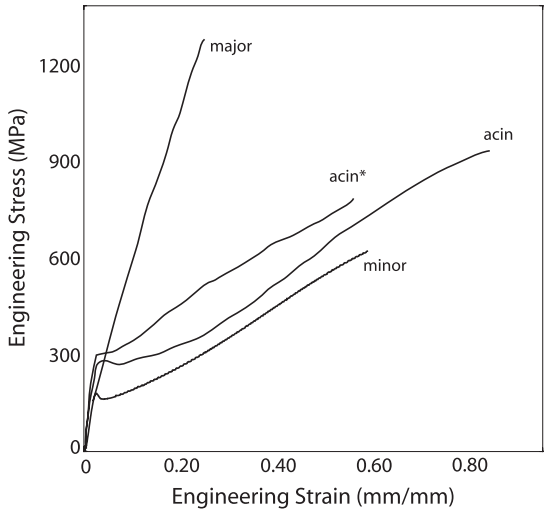


FIG. 5.—Stress-strain curves for aciniform (acin), minor ampullate (minor), and major ampullate (major) fibers from *Argiope trifasciata*, as well as aciniform silk from *A. argentata* (acin*).

ampullate silk from closely related orb weavers such as *Araneus diadematus* (131–211 MPa; Köhler and Vollrath 1995) and *Araneus sericatus* (160 MPa; Denny 1976). Pérez-Rigueiro et al. (2001) did measure a much lower toughness for major ampullate silk from a single *A. trifasciata* web (90 ± 10 MPa), but they also measured unusually low ultimate strengths for their silks (600 ± 50 MPa) compared to the published values for major ampullate silk from most studies (1–1.4 GPa; Denny 1976; Köhler and Vollrath 1995; this study).

The repetitive sequences of aciniform fibroin are inconsistent with sequences from previously characterized araneoid fibroins (fig. 3*b*), yet based on amino acid compositions for the aciniform glands (fig. 2) and our cDNA library screen, AcSp1 appears to be the primary fibroin component of aciniform silk. Furthermore, the mechanical properties of aciniform silk are quite different from other spider silks, especially in terms of toughness (fig. 6). These facts are relevant to recent debates concerning the importance of fibroin sequences versus the spinning process in the determination of silk mechanical

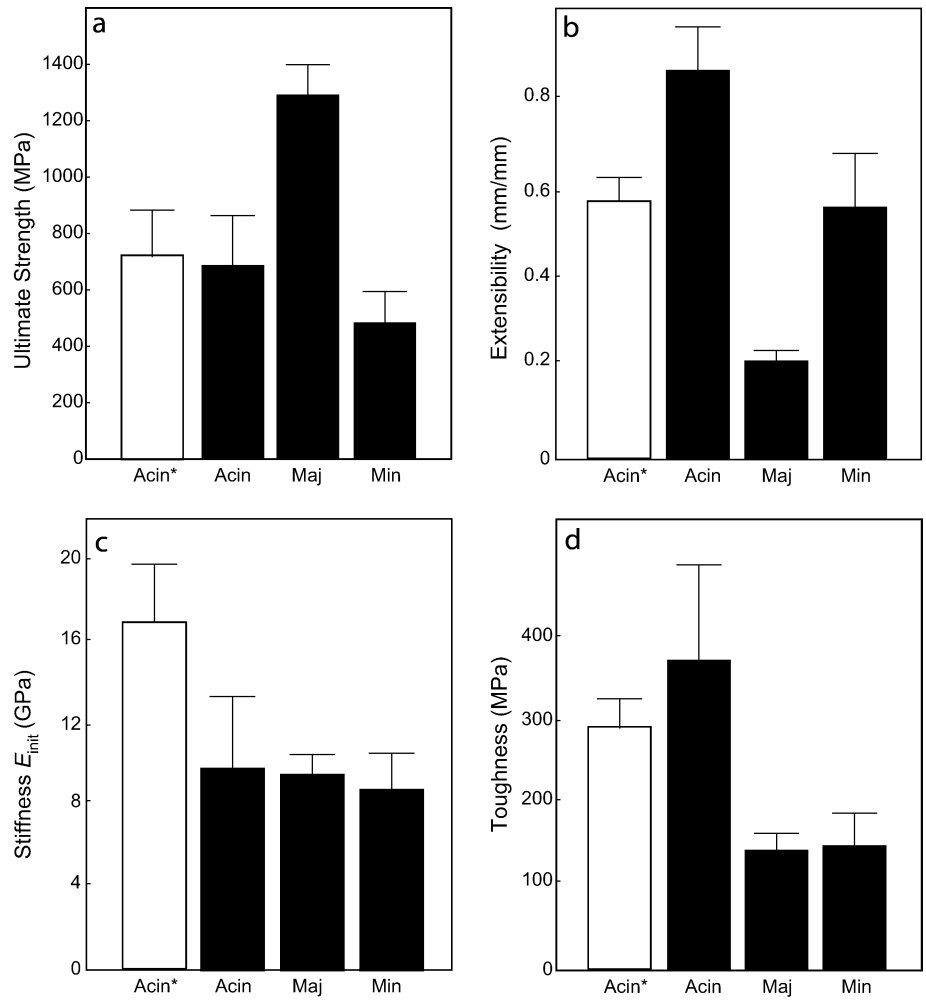


FIG. 6.—Comparison of the mean ± SD ultimate strength (*a*), extensibility (*b*), stiffness (*c*), and toughness (*d*) of aciniform (acin), minor ampullate (minor), and major ampullate (major) silks from *Argiope trifasciata*, as well as aciniform silk from *A. argentata* (acin*). Ultimate strength, extensibility, and toughness all differed significantly between the three types of *A. trifasciata* silk (one-way ANOVAs ($F_{2,32}$= 42–168, all $P < 0.0001$), but there was no significant difference in the stiffness of those silks ($F_{2,32} = 0.6$, P = n.s.).

properties (Calvert 1998; Vollrath 1999; Gatesy et al. 2001). Both sides of the debate acknowledge that fibroin sequences and specific processing of the fibroins are required to construct a silk fiber of exceptional toughness, extensibility, and tensile strength, but at this point it is not clear which factors yield the diversity of silk mechanical properties seen in nature (e.g., Stauffer, Coguill, and Lewis 1994; Köhler and Vollrath 1995; Gosline et al. 1999; Moore and Tran 1999; Shao and Vollrath 1999; Gosline et al. 2002).

While there is detailed information on the spinning process of major ampullate silk from the golden orb-weaver, *Nephila edulis* (Tetragnathidae; Vollrath and Knight 2001), very little is known on how the spinning process varies among spider species (Knight and Vollrath 2002), and differences among the various glands and spinnerets on a single spider remains unknown. In contrast, the extreme molecular diversity of paralogous fibroins from different spiders and glands has been characterized in some detail (e.g., Guerette et al. 1996; Gatesy et al. 2001; Hayashi 2002). Particular sequence motifs in divergent silk fibroin paralogs account for different molecular structures, and these molecular structures have been hypothesized to explain differences in mechanical performance among silk fibers produced by the various glands of araneoid spiders (e.g., Simmons, Michal, and Jelinski 1996; Gosline et al. 1999; Hayashi, Shipley, and Lewis 1999; Zhou, Wu, and Conticello 2001). Given the uniqueness of the AcSp1 sequence and the distinctive mechanical properties of aciniform silk fibers, it will be critical to determine the molecular structures formed by AcSp1 in future biophysical studies.

Significance of AcSp1

The characterization of AcSp1 and mechanical testing of aciniform silk fibers have resulted in several novel findings that impact our understanding of the evolution of spider silks. First, the same fibroin cDNA was found in expression libraries constructed from PLS and PMS aciniform glands of *A. trifasciata*. The expected product of this cDNA is the repetitive protein, AcSp1 (fig. 3*a*), and the predicted amino acid composition of AcSp1 matches the amino acid composition of protein extracted from the aciniform glands (fig. 2). Second, phylogenetic analyses of the carboxy-terminal regions of the characterized araneid fibroins suggest that AcSp1 represents a new ortholog group of spider silks (fig. 4*b* and *c*). The repetitive region of AcSp1 lacks the preponderance of simple subrepeats that typify the other described araneid fibroins. Instead, AcSp1 is composed of near perfect iterations of a complex 200 amino acid sequence repeat (fig. 3*b*). Finally, *A. trifasciata* aciniform silk has significantly different mechanical properties (ultimate strength, extensibility, and toughness) compared to *A. trifasciata* major and minor ampullate silks (figs. 5 and 6). These differences suggest that the repeats of AcSp1 assemble into different protein conformations than the major and minor ampullate fibroins. The structure-function relationships of the AcSp1 repeats warrant further investigation, as does the importance of the extreme sequence homogenization among AcSp1 repeats to fiber formation and mechanical properties.

Acknowledgments

This work was funded by the National Science Foundation (MCB-9806999) and the U.S. Army Research Office (DAAD19-02-1-0358). We thank R. Cardullo, J. Gatesy, M. Hinman, J. Jones, B. Kong, H. Lee, A. Summers, and J. Woods for laboratory and technical assistance.

Literature Cited

Andersen, S. 1970. Amino acid composition of spider silks. Comp. Biochem. Physiol. **35**:705–711.

Becker, N., E. Oroudjev, S. Mutz, J. Cleveland, P. Hansma, C. Hayashi, D. Makarov, and H. Hansma. 2003. Molecular nanosprings in spider capture-silk threads. Nat. Materials **2**:278–283.

Beckwitt, R., and S. Arcidiacono. 1994. Sequence conservation in the C-terminal region of spider silk proteins (Spidroin) from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentenarius* (Araneidae). J. Biol. Chem. **269**:6661–6663.

Beckwitt, R., S. Arcidiacono, and R. Stote. 1998. Evolution of repetitive proteins: spider silks from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentenarius* (Araneidae). Insect Biochem. Mol. Biol. **28**:121–130.

Calvert, P. 1998. Silk and sequence. Nature. **393**:309–310.

Colgin, M., and R. Lewis. 1998. Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like "spacer regions." Protein Sci. **7**:667–672.

Denny, M. 1976. The physical properties of spider silk and their role in the design of orb-webs. J. Exp. Biol. **65**:483–506.

Farris, J. 1989. The retention index and the rescaled consistency index. Cladistics. **5**:417–419.

Fedič, R., M. Žurovec, and F. Sehnal. 2003. Correlation between fibroin amino acid sequence and physical silk properties. J. Biol. Chem. **278**:35255–35264.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

Foelix, R. 1996. Biology of spider. Oxford University Press, New York.

Gatesy, J., C. Hayashi, D. Motriuk, J. Woods, and R. Lewis. 2001. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. Science. **291**:2603–2605.

Gordon, J. 1988. The science of structures and materials. Scientific American Library, New York.

Gosline, J., P. Guerette, C. Ortlepp, and K. Savage. 1999. The mechanical design of spider silks: from fibroin sequence to mechanical function. J. Exp. Biol. **202**:3295–3303.

Gosline, J., M. Lillie, E. Carrington, P. Guerette, C. Ortlepp, and K. Savage. 2002. Elastic proteins: biological roles and mechanical properties. Phil. Trans. R. Soc. Lond. B **357**: 121–132.

Guerette, P., D. Ginzinger, B. Weber, and J. Gosline. 1996. Silk properties determined by gland-specific expression of a spider fibroin gene family. Science **272**:112–115.

Hayashi, C. 2002. Evolution of spider silk proteins: insight from phylogenetic analyses. Pp. 209–223 *in* R. DeSalle, G. Giribet, and W. Wheeler, eds. Molecular systematics and evolution: theory and practice. Birkhaeuser, Cambridge, Mass.

Hayashi, C., and R. Lewis. 1998. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. J. Mol. Biol. **275**:773–784.

———. 2000. Molecular architecture and the evolution of a modular spider silk protein gene. Science **287**:1477–1479.

Hayashi, C., N. Shipley, and R. Lewis. 1999. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. Int. J. Biol. Macromol. **24**: 271–275.

Hinman, M., and R. Lewis. 1992. Isolation of a clone encoding a second dragline silk fibroin: *Nephila clavipes* dragline silk is a two-protein fiber. J. Biol. Chem. **267**:19320–19324.

Kluge, A., and J. Farris. 1969. Quantitative phyletics and the evolution of anurans. Syst. Zool. **18**:1–32.

Knight, D., and F. Vollrath. 2002. Spinning an elastic ribbon of spider silk. Phil Trans. R. Soc. Lond. B. **357**:219–227.

Köhler, T., and F. Vollrath. 1995. Thread biomechanics in the two orb-weaving spiders, *Araneus diadematus* (Araneae, Araneidae) and *Uloborus walckenaerius* (Araneae, Uloboridae). J. Exp. Zool. **271**:1–17.

Mita, K., S. Ichimura, and T. James. 1994. Highly repetitive structure and its organization of the silk fibroin gene. J. Mol. Evol. **38**:583–592.

Moore, A., and K. Tran. 1999. Material properties of cobweb silk from the black widow spider *Latrodectus hesperus*. Int. J. Biol. Macromol. **24**:277–282.

Pérez-Rigueiro, J., M. Elices, J. Llorca, and C. Viney. 2001. Tensile properties of *Argiope trifasciata* drag line silk obtained from the spider's web. J. Appl. Polym. Sci. **82**: 2245–2251.

Posada, D., and K. Crandall. 1998. ModelTest: testing the model of DNA substitution. Bioinformatics **14**:817–818.

Sezutsu, H., and K. Yukuhiro. 2000. Dynamic rearrangements within the *Antheraea pernyi* silk fibroin gene is associated with four types of repetitive units. J. Mol. Evol. **51**:329–338.

Shao, Z., and F. Vollrath. 1999. The effects of solvents on the contraction and mechanical properties of spider silk. Polymer **40**:1799–1806.

Simmons, A., C. Michal, and L. Jelinski. 1996. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. Science. **271**:84–87.

Stauffer, S., S. Coguill, and R. Lewis. 1994. Comparison of physical properties of three silks from *Nephila clavipes* and *Araneus gemmoides*. J. Arachnol. **22**:5–11.

Swofford, D. L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. Sinauer Associates, Sunderland, Mass.

Vollrath, F. 1999. Biology of spider silk. Int. J. Biol. Macromol. **24**:81–88.

Vollrath, F., and D. Knight. 2001. Liquid crystalline spinning of spider silk. Nature. **410**:541–548.

Xu, M., and R. Lewis. 1990. Structure of a protein superfiber: spider dragline silk. Proc. Natl. Acad. Sci. USA **87**:7120–7124.

Zhou, Y., S. Wu, and V. Conticello. 2001. Genetically directed synthesis and spectroscopic analysis of a protein polymer derived from a flagelliform silk sequence. Biomacromolecules **2**:11–125.

Michele Vendruscolo, Associate Editor

Accepted July 5, 2004