Spider Glue Proteins Have Distinct Architectures Compared with Traditional Spidroin Family Members*

Received for publication, July 10, 2012, and in revised form, August 23, 2012. Published, JBC Papers in Press, August 27, 2012, DOI 10.1074/jbc.M112.399816

Keshav Vasanthavada‡1,2, Xiaoyi Hu†11, Tiffany Tuton-Blasingame∥11, Yang Hsia‡, Sujatha Sampath‡, Ryan Pacheco‡, Jordan Freeark∥, Arnold M. Falick, Simon Tang∥, Justine Fong‡, Kristin Kohler‡, Coby La Mattina-Hawkins‡, and Craig Vierra‡3

From the ‡Department of Biological Sciences and the §School of Engineering and Computer Science, University of the Pacific, Stockton, California 95211 and the ¶Department of Chemistry and Biochemistry, Magnetic Resonance Research Center, Arizona State University, Tempe, Arizona 85287-1604

Background: Spiders extrude adhesive glues to form connection joints that mediate web construction and prey wrapping.

Results: DNA microarray analysis and mass spectrometry reveal new protein glue constituents that comprise connection joints.

Conclusion: Spider glue proteins represent a diverse group of polypeptides with distinct molecular architectures.

Significance: Learning how spider glues mediate the fusion of fibers is crucial for understanding adhesion mechanisms in biology.

Adhesive spider glues are required to perform a variety of tasks, including web construction, prey capture, and locomotion. To date, little is known regarding the molecular and structural features of spider glue proteins, in particular bioadhesives that interconnect dragline or scaffolding silks during three-dimensional web construction. Here we use biochemical and structural approaches to identify and characterize two aggregate gland specific gene products, AgSF1 and AgSF2, and demonstrate that these proteins co-localize to the connection joints of both webs and wrapping silks spun from the black widow spider, Latrodectus hesperus. Protein architectures are markedly divergent between AgSF1 and AgSF2, as well as traditional spider silk fibroin family members, suggesting connection joints consist of a complex proteinaceous network. AgSF2 represents a nonglycosylated 40-kDa protein that has novel internal amino acid block repeats with the consensus sequence NVNVN embedded in a glycine-rich matrix. Analysis of the amino acid sequence of AgSF1 reveals pentameric QPGSG iterations that are similar to conserved modular elements within mammalian elastin, a rubber-like elastomeric protein that interfaces with collagen. Wet-spinning methodology using purified recombinant proteins show AgSF1 has the potential to self-assemble into fibers. X-ray fiber diffraction studies performed on these synthetic fibers reveal the presence of noncrystalline domains that resemble classical rubber networks. Collectively, these data support that the aggregate gland serves to extrude a protein mixture that contains substances that allow for the self-assembly of fiber-like structures that interface with dragline silks to mediate prey capture.

Spider silk is a lightweight material that has a combination of high-tensile strength and extensibility. Anatomical studies of spider abdominal glands have revealed the presence of seven distinct silk and/or glue-silk producing glands, which include the major and minor ampullate, tubuliform, flagelliform, aciniform, pyriform, and aggregate gland (1). The silk producing glands have been hypothesized to serve as specialized biofactories that produce proteins spun into different silk types with distinct mechanical properties (2, 3). These structural proteins share common protein architectures, including the presence of internal block repeats, and nonrepetitive conserved N and C termini (4–6). Six of the seven silk-producing glands have been shown to produce fibers with the exception of the aggregate gland (7–13). Due to the morphology of the aggregate gland, its spinning duct, and the anatomical features of its sigmoids, it has been debated whether this gland synthesizes structural proteins that self-assemble into fibers or whether it exclusively secretes nonstructural adhesive proteins.

Chemical analysis of web constituents cast by orb-weavers has demonstrated that the aggregate gland manufactures a complex aqueous solution containing water-soluble compounds related to neurotransmitters, inorganic salts, small peptides, and glycoproteins (14, 15). These studies have shown that ecribellate spider glue consists of microscopic nodules made of glycoproteins (15). Protein sequences for two of the glycoproteins reveal these products are expressed from opposite strands of an identical DNA sequence (16). Thus, at least for orb-weavers, spider glue glycoproteins have been proposed to function as adhesive substances that facilitate prey capture; however, they have not been shown to have intrinsic properties for self-assembly into structures that resemble fibers. Because spider glues and cements are used in conjunction with dragline silk (scaffolding threads), we hypothesized that the expression of glue

* This work was supported by National Science Foundation RUI Grants MCB-0950372 and DMR-1105310.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) JX262195, JX262192, and EH072326.1.

‡ Joint first co-authors.

1 Present address: Discovery Technologies, Hoffmann-La Roche, 340 Kingsland St., Nutley, NJ 07110.

2 To whom correspondence should be addressed: Dept. of Biological Sciences, University of the Pacific, 3601 Pacific Ave., Stockton, CA 95211. Tel.: 209-946-3024; Fax: 209-946-3022; E-mail: cvierra@pacific.edu.
proteins could be subject to coordinate regulation or energetically linked to the synthesis of the MaSp1 and MaSp2 dragline silk proteins via inter-silk-gland communication. To test this hypothesis, we coupled DNA microarray analysis, mass spectrometry, and other biochemical approaches to investigate the presence of new materials that interface with dragline silk fibers to perform essential biological tasks. For this, dragline silk was collected by allowing the spiders to free fall and changes in gene expression profiles in the silk-producing glands were monitored.

In this study, we describe the identification and characterization of two gene products aggregate gland silk factors 1 and 2 (AgSF1 and AgSF2). AgSF1 and AgSF2 are expressed in high quantities in the aggregate gland and extruded as adhesive materials that function to interconnect scaffolding fibers in the three-dimensional architecture of the web as well as wrapping silk for prey immobilization. We also provide evidence that recombinant AgSF1 protein chains can self-assemble into synthetic fibers. Structural studies of these fibers reveal a material that is predominantly noncrystalline.

**EXPERIMENTAL PROCEDURES**

**DNA Microarray Analysis**—A cDNA library was prepared as previously described (17). After plating the cDNA library, over 3300 individual plaques were removed and placed into SM (100 mM NaCl, 8 mM MgSO4, 50 mM Tris-Cl) buffer to facilitate diffusion. The diffused recombinant viruses were used as templates for PCR. cDNAs were amplified using the universal forward and reverse primers with the sequences 5’TAA TAC CAC TAC-3 and 5’-GTC GAC GAG TTA CAT GGC CAA G-3’, respectively. After confirming amplification of cDNAs using agarose gel electrophoresis, the samples were desalted and then resuspended in 3× SSC (0.9 M NaCl and 0.09 M sodium citrate, pH 7.2) and printed onto Corning UltraGAPS™-coated slides using an OmniGrid Accent microarrayer (GeneMachines). Three different spiders were silked once every 24 h over a 5-day period. The quantity of dragline silk collected each day ranged from 150 to 200 cm. Control spiders were not silked but housed under identical conditions. Following silking, both groups were dissected and the cDNAs were subject to DNA sequencing to confirm the accuracy of the sequences (University of Florida). For structural studies, the cDNAs were excised from the pBAD TOPO ThioFusion vectors using the restriction enzymes EcoRI and SalI and ligated into pET-19b-Sumo.

The full-length AgSF2 cDNA was amplified using PCR with the forward and reverse primers: 5’-GAA TTC CGA CCT GAA CCT GGT GTT CAT G-3’ and 5’-GTC GAC GAG TTA CAT GGC CAA G-3’, respectively. After confirming amplification, the cDNA was inserted into the pBAD-TOPO TA expression vector site for signal peptidase. Following amplification, the cDNA was ligated into pET-19b-Sumo. The forward primer was designed to remove the first 24 amino acid residues, which corresponded to the predicted cleavage site for signal peptidase. Following amplification, the cDNA was ligated into pET-19b-Sumo.

**DNA Constructs**—For expression studies, the AgSF1 cDNA was amplified by PCR using the forward and reverse primers

5’-GAA TTC GGC TAC AAA AAA ACC GTT GGA AAA GAT GGA CAA ATT-3’ and 5’-GTC GAC GAG TTA CAT GGC CAA G-3’, respectively. EcoRI and SalI restriction sites were engineered on the 5’-termini of the forward and reverse primers to facilitate subcloning of the cDNAs. AgSF1 cDNA fragments were amplified using PCR with a cDNA library prepared from the silk-producing glands of the black widow spider, Latrodectus hesperus. Following amplification, two cDNAs, one ~600 bp and another 1000 bp, were extracted and ligated into the pBAD TOPO ThioFusion expression vector (Invitrogen). The 600-bp cDNA fragment coded for two copies of the 64-amino acid repeat, which was denoted AgSF1G. The 1000-bp fragment encoded for two copies of the 64-amino acid repeat and 3.5 copies of the 36-amino acid ensemble repeat, which was dubbed AgSF1G+XPGXG. Both cDNAs were subject to DNA sequencing to confirm the accuracy of the sequences.

**Development of Polyclonal Antibodies**—Anti-AgF1 polyclonal antibodies were generated using the synthetic peptide GKYKTVKGDQIVC (EZBiolab Inc., Carmel, IN). The anti-AgF2 polyclonal antiserum was generated against two synthetic peptides with the sequences IQSRDPGVHHVRYC and GYKKTVGKDGQIVC (EZBiolab Inc., Carmel, IN). The antibodies were generated using the synthetic peptide ligands via inter-silk-gland communication. To test this hypothesis, we coupled DNA microarray analysis, mass spectrometry, and other biochemical approaches to investigate the presence of new materials that interface with dragline silk fibers to perform essential biological tasks. For this, dragline silk was collected by allowing the spiders to free fall and changes in gene expression profiles in the silk-producing glands were monitored.

In this study, we describe the identification and characterization of two gene products aggregate gland silk factors 1 and 2 (AgSF1 and AgSF2). AgSF1 and AgSF2 are expressed in high quantities in the aggregate gland and extruded as adhesive materials that function to interconnect scaffolding fibers in the three-dimensional architecture of the web as well as wrapping silk for prey immobilization. We also provide evidence that recombinant AgSF1 protein chains can self-assemble into synthetic fibers. Structural studies of these fibers reveal a material that is predominantly noncrystalline.
Protein Expression and Purification—Recombinant proteins were produced using 1 liter of LB. For induction experiments, each clone was grown overnight until saturation. Saturated cultures were then combined with 4 liters of fresh LB, supplemented with 0.02% arabinose or 1 mM isopropyl-β-D-galactopyranoside for induction purposes, and grown an additional 4 h. Following induction, the bacterial cells were pelleted at 6,000 × g and then lysed in 13.3 ml of FastBreak™ Cell Lysis Reagent (Promega) per 1.0 liter of culture. To promote further lysis, the solution was sonicated at maximum setting for 1 min using a VirSonic 60 instrument (VirTis). After sonication, samples were clarified by spinning for 5 min at 8,000 × g. Supernatants were combined with 1 ml of nickel-nitrilotri-acetic acid-agarose resin (50% slurry; Qiagen) and gently rocked at room temperature for 1 h to facilitate binding of the fusion proteins via their His6 tag. The resin was washed using a buffer (50 mM NaH2PO4, 300 mM NaCl) supplemented with 20 mM imidazole, followed by elution of the fusion proteins with 20 ml of same buffer containing 250 mM imidazole. Protein samples were dialyzed against water for 2–3 days and then subjected to freeze-drying using a FreeZone12plus instrument (Labconco). Typical yields for truncated AgSF1 recombinant proteins after purification using 10 liters of culture were ~50–100 mg, respectively.

Collection and Solubilization of Silk Fibers—Spiders were housed and silk-producing glands were removed as previously described (19). For Western blot studies, silk samples were dissolved in 8 M urea, incubated at 55 °C for 2 h, and then heated at 95 °C for 5 min before SDS-PAGE analysis. Whole cell lysates were generated by solubilizing the silk-producing glands in 8 M urea and treated in a similar fashion as the fibers, with the exception that the samples were briefly sonicated prior to electrophoresis.

Mass Spectrometric Analysis—In-solution tryptic digests of the solubilized wrapping silk as well as other fiber types were performed as previously described (9, 10). The MS map and peptide MS/MS spectra were acquired with a MALDI/TOF/TOF mass spectrometer (4700 Proteomics Analyzer, Applied Biosystems, Foster City, CA) operated in reflector mode or a TOF mass spectrometer (4700 Proteomics Analyzer, Applied Biosystems, Foster City, CA) operated in reflector mode or a TOF mass spectrometer (4700 Proteomics Analyzer, Applied Biosystems, Foster City, CA). De novo peptide sequences were identified over 3,300 cDNAs from individual plaques carrying recombinant DNA. DNA Microarray Analysis Reveals Novel Gene Products That Are Linked to Dragline Silk Production—To identify genes that encode proteins that have functional ties with dragline silk constituents, we generated spider DNA chips for microarray analysis. Using a cDNA library prepared from the silk-producing glands of the black widow spider, along with a universal set of primers that flanked the region of cDNA insertion, we amplified over 3,300 cDNAs from individual plaques carrying recombinant viral chromosomes. Spider cDNAs were printed onto glass slides as unknown features for DNA microarray analysis. Two distinct treatments were used to collect total RNA from spiders for the DNA microarray analysis. One group of spiders, the “free fall-silked group,” was used to collect dragline silk. The other group, which represented the nonsilked or control group, was maintained under identical conditions except no dragline silk was removed from these individuals. Because the major amylace gland is known to produce MaSp1 and MaSp2, we hypothesized that dragline silk would influence the expression pattern of fibroins and other silk components in different silk-producing glands because coordinate regulation between the glands has been proposed (1). This strategy proved successful, as ~44 of the printed genes were observed to be down-regulated by as much as 5-fold following the silking procedure (Table 1). These cDNAs were selected for sequencing and their identification.
identities determined using the computer algorithm BLASTn to search the sequences against the NCBInr database (Table 1). Several gene products that have been demonstrated by MS/MS studies to be constituents of silks, which include the tubuliform fibroins TuSp1 (7), ECP-1 (18), and ECP-2 (24), as well as gene products previously implicated in the silk production pathway or spiral glue droplets, such as a peroxidase, arginine kinase (25), and aggregate spider glue factor (16), respectively, were identified during our DNA microarray analysis.

Two cDNAs were found multiple times on the DNA microarray; these corresponded to clones 549, 2078, 2548, 2860, and 3019 as well as clones 98, 947, 958, 2525, 2691, and 3009. These clones were identified as being related to genes that are expressed in the silk glands and that showed unique expression patterns. Based on these findings, we decided to further characterize their molecular properties and function. To obtain full-length cDNAs, we screened a cDNA library prepared from silk-producing glands using conventional nucleic acid-nucleic acid hybridization assays (data not shown). From this screen, we isolated longer cDNAs that contained 1213 amino acids with a predicted molecular mass of 112 kDa and pI 6.61. Theoretical amino acid composition profiling demonstrated glycine (Gly) and asparagine (Asn) were the most abundant amino acids, comprising 28.4 and 14.5%, respectively. Using the program REPRO, we identified a 33-amino acid block repeat that was repeated 4 times within the internal region of the translated sequence (Fig. 1A). Within each block repeat two copies of the sequence NXXNXN (X = Val, Asp, Leu, Phe, or Met) were found interrupted by GGX iterations, where X represents a small subset of amino acids (X = Ala, Gly, Leu, or Ser). Occasional histidine residues were embedded within the GGX iterations (Fig. 1C). GGX motifs have been reported in major amylase fibroin sequences to adopt helical structures (28). In addition, 2 cysteine residues were present within the C terminus of translated cDNA. Protein BLAST searches against the nrNCBI database with the predicted full-length sequence revealed no significant matches to other proteins. However, consistent with proteins that are secreted, the algorithm SignalP predicted the presence of an N-terminal secretion signal.

**AgSF1 and AgSF2 Display Aggregate-Gland Restricted Patterns of Expression**—Expression patterns for these two genes were determined using quantitative real-time PCR and Western blot analyses. For quantitative RT-PCR, total RNA was isolated from the seven different silk-producing glands. Quantitative RT-PCR studies demonstrated that the aggregate gland expressed the highest quantities of mRNAs for both genes (Fig. 2, A and B). Very low levels of AgSF1 mRNA were detected in the pyroform gland; however, these levels are likely insignificant given that the aggregate gland expresses ~19-fold more mRNA. Based upon their restricted patterns of expression, we named the two gene products AgSF1 and AgSF2. To correlate mRNA levels with endogenous protein amounts, peptides corresponding to translated regions of the AgSFs were synthesized.

**Table 1: Genes that decrease after dragline silking spiders for 5 days**

| Clone number | Description | Putative Function | Accession # |
|--------------|-------------|-------------------|-------------|
| 2            | Arginine kinase | Metabolism | JX26190 |
| 10           | Cysteine protease inhibitor | Metabolism | JX26190 |
| 2412         | Glutathione reductase | Metabolism | JX26190 |
| 39, 64       | Heat shock protein 70 | Stress response | JX26190 |
| 41, 557      | Eukaryotic initiation factor 4A | Translation | JX26190 |
| 44           | Aggregate spider glue 2 | Translation | JX26190 |
| 1921         | Hexosamine II | Metabolism | JX26190 |
| 3024         | Integral membrane protein | Signal transduction | JX26190 |
| 590, 666, 2424 | TuSp1 | Tubulin-like | JX26190 |
| 2593         | Peroxidase | Stress response | JX26190 |
| 131, 639, 2677, 2913, 2745, 3146 | ECP-1 | Tubulin-like | JX26190 |
| 2913         | ECP-2 | Tubulin-like | JX26190 |
| 2890         | F1,6-Bisphosphate aldehyde | Metabolism | JX26190 |
| 98, 947, 958, 2525, 2691, 2860, 3009 | Unknown | Glue protein (AgSF2) | JX26190 |
| 2891         | Elongation factor alphaTu | Translation | JX26190 |
| 1309         | Ras-related protein Rab-7 | Translation | JX26190 |
| 2602         | Sec61 alpha subunit | Protein transport | JX26190 |
| 2410         | Cytochrome oxidase subunit II | Metabolism | JX26190 |
| 549, 2078, 2548, 2860, 3019 | Unknown | Glue protein (AgSF1) | JX26190 |
| 3029         | 3145         | Unknown | JX26190 |
| 2004         | RNA-dependent RNA polymerase | Nucleotidytransferase | JX26190 |
| 1117         | Unknown | Unknown | JX26190 |
| 908          | Retinol-tube | Unknown | JX26190 |
| 2499         | PPR repeat domain | Unknown | JX26190 |
| 907          | Mollidine fasciculin | Adhesion | JX26190 |
| 65           | Unknown | Unknown | JX26190 |
and used to generate anti-AgSF1 and anti-AgSF2 polyclonal antibodies for Western blot analysis using lysates from the silk-producing glands. For expression studies, bacterially produced AgSF1 His-Thio-tagged or His-Sumo-tagged proteins engineered with either two copies of the 64-amino acid repeat alone (AgSF1G), or two copies of the 64-amino acid module plus 3.5 copies of the 36-amino acid repeat (AgSF1G\*\(X\)PGXG) were readily detected by Western blot analysis using the anti-AgSF1 antibodies (Fig. 3A, lanes 3, 5, 9, and 10). The same recombinant proteins were detected using a monoclonal anti-His antibody (Fig. 3A, lanes 1–2 and 7–8). Bacterially expressed His-Thio-tagged full-length AgSF2 was also detected using an anti-His monoclonal antibody and anti-AgSF2 polyclonal antisera from arabinose-induced culture extracts (Fig. 3A, lanes 11 and 13). No immunoreactive bands were detected from lysates under nonarabinose conditions (Fig. 3A, lanes 4, 6, 12, and 14). In total, these studies demonstrate that the AgSF1 and AgSF2 antisera recognize the recombinant proteins in a highly specific manner.

To examine the expression pattern of AgSF1 and AgSF2 in tissues, we tested whole cell lysates collected from the seven different silk-producing glands. Western blot analysis using an anti-AgSF1 antibody revealed dominant immunoreactive species in the aggregate, aciniform, and flagelliform glands (Fig. 3B, lanes 6, 10, and 11). The molecular masses of these species were \(\sim 245\) kDa or larger. Using the anti-AgSF2 antibody, a single immunoreactive band was detected in the aggregate gland extract (Fig. 3C, lane 7). The experimental molecular mass of AgSF2 closely correlated to the predicted mass, which was \(\sim 40\) kDa. No other immunoreactive species were detected in the major and minor ampullate glands, tubuliform, aciniform, or flagelliform glands (Fig. 3C, lanes 8–12). A minor immunoreactive, slower migrating species was detected in the pyriform glandular extract (Fig. 3C, lane 13).

**AgSF1 and AgSF2 Are Constituents of Spider Glues**—To investigate whether AgSF1 and AgSF2 were spun or deposited on fibers, we collected different fiber types from black widow
spiders for biochemical studies. Dragline silk, egg sacs, wrapping silk, gumfooted lines, scaffolding fibers, and connection joints were obtained and solubilized in urea or guanidinium hydrochloride. Dissolved protein mixtures were analyzed using Western blot analysis for the presence of AgSFs using the anti-AgSF1 and anti-AgSF2 antisera. Because connection joints are difficult to individually remove from the web, we collected scaffolding fibers containing connection joints (scaffolding fibers plus connection joints) and scaffolding fibers alone. Immunoreactive AgSF1 proteins were detected in scaffolding fibers containing connection joints (scaffolding fibers plus connection joints) and scaffolding fibers alone. Immunoreactive AgSF1 proteins were detected in scaffolding fibers containing connection joints and gumfooted lines, but were absent in dragline, egg sacs, and scaffolding fibers without connection joints (Fig. 3B, lanes 1–5). The immunoreactive protein detected in the scaffolding fibers plus connection joints displayed a migration profile similar to the species detected in the aggregate gland lysate (Fig. 3B, compare lanes 2 and 6). The molecular mass of the immunoreactive species detected in the gumfooted lines was considerably larger, exceeding 300-kDa (Fig. 3B, lane 4). Similar fiber localization profiles were observed for AgSF2, which revealed the presence of an immunoreactive species in scaffolding fibers with connection joints, but not in the scaffolding fibers alone (Fig. 3C, lanes 2 and 3). AgSF2 was also detected in wrapping silk (Fig. 3C, lane 6). However, it was not detected in dragline silk, gumfooted lines, or egg sacs (Fig. 3C, lanes 1, 4, and 5). Taken together, the mRNA expression results and Western blot data support that AgSF1 and AgSF2 are expressed in the aggregate gland and extruded into the connection joint structures.

In addition to using Western blot analyses to investigate the presence of AgSF1 and AgSF2 in the different fiber types, we...
Spider Glues Interconnect Dragline and Wrapping Silk Fibers

Several species of spider glue interconnect the loss of scaffolding threads to create a matrix connecting Dragline Silk Fibers and AgSF2 are also extruded into wrapping silk constituents. Acidic sequences of previously reported spider silk family members represent hot spots for nonenzymatic degradation of proteins. Cysteine residues were identified by MS/MS analysis of tryptic peptides. These peptides were shown to be identical to amino acid sequences following theoretical tryptic cleavage. These peptide masses correspond to the regions of the AgSF2 cDNA sequence, with the exception that, in some cases, de novo sequences showed Leu instead of Ile. It is sometimes difficult to distinguish these two isobaric amino acid residues by MS/MS.

In addition to enzymatic digestion of solubilized wrapping silk, we discovered two precursor ion masses that were identical to proteins from the translated AgSF1 cDNA sequence following theoretical tryptic cleavage. These peptide ion masses corresponded to m/z 813.4 and 1580.7. After MS/MS analysis, these peptides were shown to be identical to regions found in the predicted sequence of AgSF1 (Fig. 4, B and C). One of the peptides, GGHHPGFFGK, resulted from two tryptic cleavage events, whereas the other peptide, GGHHPGFFGK, likely resulted from a combination of acid hydrolysis (nonenzymatic) and tryptic cleavage (Fig. 4, B and C). Both aspartic acid and asparagine have been reported to be present in post-translational modifications to AgSF2 (data not shown).

Within a region of the junction where a thin layer of joint material is observed, large numbers of connection joints within the ultrastructure reveal the presence of post-translational modifications to AgSF2 (data not shown). Enzymatic digestion of solubilized wrapping silk, we discovered two precursor ion masses that were identical to proteins from the translated AgSF1 cDNA sequence following theoretical tryptic cleavage. These peptide ion masses corresponded to m/z 813.4 and 1580.7. After MS/MS analysis, these peptides were shown to be identical to regions found in the predicted sequence of AgSF1 (Fig. 4, B and C). One of the peptides, GGHHPGFFGK, resulted from two tryptic cleavage events, whereas the other peptide, GGHHPGFFGK, likely resulted from a combination of acid hydrolysis (nonenzymatic) and tryptic cleavage (Fig. 4, B and C).

Both aspartic acid and asparagine have been reported to be present in post-translational modifications to AgSF2 (data not shown). Within a region of the junction where a thin layer of joint material is observed, large numbers of connection joints within the ultrastructure reveal the presence of post-translational modifications to AgSF2 (data not shown). Enzymatic digestion of solubilized wrapping silk, we discovered two precursor ion masses that were identical to proteins from the translated AgSF1 cDNA sequence following theoretical tryptic cleavage. These peptide ion masses corresponded to m/z 813.4 and 1580.7. After MS/MS analysis, these peptides were shown to be identical to regions found in the predicted sequence of AgSF1 (Fig. 4, B and C). One of the peptides, GGHHPGFFGK, resulted from two tryptic cleavage events, whereas the other peptide, GGHHPGFFGK, likely resulted from a combination of acid hydrolysis (nonenzymatic) and tryptic cleavage (Fig. 4, B and C). Both aspartic acid and asparagine have been reported to be present in post-translational modifications to AgSF2 (data not shown).

In addition to enzymatic digestion of solubilized wrapping silk, we discovered two precursor ion masses that were identical to proteins from the translated AgSF1 cDNA sequence following theoretical tryptic cleavage. These peptide ion masses corresponded to m/z 813.4 and 1580.7. After MS/MS analysis, these peptides were shown to be identical to regions found in the predicted sequence of AgSF1 (Fig. 4, B and C). One of the peptides, GGHHPGFFGK, resulted from two tryptic cleavage events, whereas the other peptide, GGHHPGFFGK, likely resulted from a combination of acid hydrolysis (nonenzymatic) and tryptic cleavage (Fig. 4, B and C). Both aspartic acid and asparagine have been reported to be present in post-translational modifications to AgSF2 (data not shown).

Within a region of the junction where a thin layer of joint material is observed, large numbers of connection joints within the ultrastructure reveal the presence of post-translational modifications to AgSF2 (data not shown). Enzymatic digestion of solubilized wrapping silk, we discovered two precursor ion masses that were identical to proteins from the translated AgSF1 cDNA sequence following theoretical tryptic cleavage. These peptide ion masses corresponded to m/z 813.4 and 1580.7. After MS/MS analysis, these peptides were shown to be identical to regions found in the predicted sequence of AgSF1 (Fig. 4, B and C). One of the peptides, GGHHPGFFGK, resulted from two tryptic cleavage events, whereas the other peptide, GGHHPGFFGK, likely resulted from a combination of acid hydrolysis (nonenzymatic) and tryptic cleavage (Fig. 4, B and C). Both aspartic acid and asparagine have been reported to be present in post-translational modifications to AgSF2 (data not shown).

FIGURE 3. AgSF1 and AgSF2 proteins are expressed in the aggregate gland and localize to the connection joints of fibers. Western blot analysis of recombinant proteins expressed in Escherichia coli, extracts generated from silk-producing glands, or spun fibers. Plus or minus designations represent induced or noninduced conditions, respectively. Fat tissue was also included as a negative control. Denatured proteins were size fractionated using a 4–20% Tris-HCl precast polyacrylamide gel. A, bacterially expressed AgSF1 or AgSF2 recombinant proteins were detected using an anti-His6 antibody (α6xHis), anti-AgSF1 (αAgSF1), or anti-AgSF2 (αAgSF2) polyclonal antisera. B, solubilized proteins from different fiber types and silk-producing glands were probed with the anti-AgSF1 polyclonal antisera. Molecular masses for the protein ladder are shown on the left in kDa.
FIGURE 4. MALDI-TOF or MALDI-tandem TOF analysis of precursor peptide ions from in-solution tryptic digestion of solubilized wrapping silk. A, MS analysis reveals the presence of precursor ion masses that are identical to products from the theoretical tryptic digestion of the translated AgSF2 cDNA. Subsequently, MS/MS analyses confirmed precursor ions with m/z 1144.5 and 1574.8 have the sequences HFLGGGLGNMAK and GGHLLGPLGGLLNTAGK, respectively (boxed region). In these peptide sequences, the underlined Leu residues were determined to be the amino acid Ile (an isobaric residue) in the translated AgSF2 sequence. B, high energy collision-activated dissociation spectrum of precursor ion with m/z 813.4 (MH\(^+\), monoisotopic). The sequence of this peptide was found to be GGHLPFGGGK. C, high energy collision-activated dissociation spectrum of precursor ion with m/z 1580.7 (MH\(^+\), monoisotopic). The sequence of this peptide was found to be GGHLPFGGTGGGGFGGGK. Both precursor ions with m/z 813.4 and 1580.7 after de novo sequencing were found to be present within the translated AgSF1 cDNA sequence.
peak at \(d \sim 10\ \text{Å}\) decreases from the as-spun (FWHM 74.9°) to the 3.5× fiber (FWHM 56.7°), which also suggests a higher degree of orientation of the protein chains in the 3.5× fiber (21, 31). Taken together, the decrease in the widths of the azimuthal peaks combined with the shift of radial peaks to lower Q values in the 3.5× fiber confirm that the post-spin draw has increased local structural order and a greater axial alignment of the protein chains. However, the post-spin draw did not result in the formation of nano-crystallites in these fibers. We also observed a strong diffraction intensity in the small angle x-ray scattering region of \(\sim 28–34\ \text{Å}\) in the XRFD pattern (Fig. 7, C and D), which is shadowed by the beam stop in the XRFD experiment.

**FIGURE 5.** Black widow spider connection joints hold web- and prey-wrapping fibers together. A, digital photograph of connection joints in a web spun by a caged spider. B and C, dissection scope images of a web segment and single connecting joint at \(\times 10\) and \(\times 115\), respectively. D, S.E. of a web connecting joint at \(\times 1500\). E, dissecting scope image of a cricket immobilized with wrapping silk at \(\times 115\). F, S.E. of a synthetic fiber spun from purified recombinant AgSF1>GxPGxG protein.
After dragline silking female black widow spiders for 5 days, the DNA microarray analysis revealed gene expression profiles that could be separated into 2 categories. One class comprised genes that showed no significant alterations in their mRNA levels, which included *masp1* and *masp2*, whereas the second class involved genes that displayed decreased mRNA quantities. The largest changes in gene activity ranged from 3.5 to 5-fold repression. Genes that were down-regulated the most included *agsf1*, *agsf2*, *tuspl*, *ecp-1*, and *ecp-2*. These levels were all ~5-fold lower relative to nonsilked spiders. The presence of a collection of genes that can be down-regulated after silking implies that the different silk-producing glands are capable of shunting resources to the major ampullate gland to maintain the synthesis of dragline silk proteins, MaSp1 and MaSp2. Consistent with this assertion is that following 1 day of dragline silking, we observed elevated levels of MaSp1 and MaSp2 mRNA (data not shown); however, after 5 days of silking, whereas some gene expression profiles for other genes were declining in different silk-producing glands, MaSp1 and MaSp2 transcripts were maintained at similar levels relative to non-silked spiders. The reshuffling of resources to the major ampullate gland is also supported by the decreased expression of *tuspl*, as well as the *ecps*, which are all genes highly expressed in females during egg case laying season and whose resources for synthesis are logically interconnected and balanced with dragline silk production. Thus, the down-regulation of egg case silk components after dragline silk removal demonstrates that our approach can be applied to identify genes that are actively expressed at high levels in other silk-producing glands, but are sensitive to the energetic costs of dragline silk synthesis.

In orb-weavers, the aggregate gland has been shown to extrude adhesive substances that include small molecular weight glycoproteins aggregate spider glue 1 and 2 (ASG1 and ASG2) (16). These glycoproteins have been reported to function as central constituents of spider viscid glue; however, whether these proteins have fiber forming properties has not been established. Sequence analysis of the ASG1 and ASG2 protein sequences showed no similarities to the connection joint proteins AgSF1 or AgSF2. Intriguingly, our DNA microarray studies did uncover one cDNA (clone 44) that contained a 65-amino acid region that showed 54% identities to a portion of the ASG2 sequence. Thus, the translated clone 44 was ASG2-like. This segment, somewhat surprisingly, did not correspond to the repetitive domain within ASG2, a region reported to share similarities with elastin and flagelliform spider silk protein repetitive sequence domains (16). Presently, it is unclear whether the product of clone 44 localizes to a specific fiber type or which silk gland expresses this gene. Attempts to use either MS or MS/MS analyses of tryptic peptides obtained from dissolved attachment discs, wrapping silk, egg case fibers, gum-footed lines, and dragline silk did not reveal peptide masses or amino acid sequences that matched theoretical digestion of the translated cDNA from clone 44. Further biochemical analysis will be required to determine the expression pattern of this gene and whether its product is extruded to coat fibers.

For the production of the anti-AgSF1 polyclonal antiserum, we used a synthetic peptide that was located within the N-terminal 64-amino block repeat. Based upon protein-protein BLAST searches of the NCBInr database, this region showed no similarities to other spidroin family members. Western blot analyses performed with this antibody revealed immunoreactive species in the aggregate, aciniform, and flagelliform glands despite the aggregate gland restricted pattern of expression of AgSF1. The molecular mass of the immunoreactive species detected in the aggregate gland co-migrated with the protein repetitive sequence domains (16). Presently, it is unclear whether these proteins have fiber forming properties has not been established. Sequence analysis of the ASG1 and ASG2 protein sequences showed no similarities to the connection joint proteins AgSF1 or AgSF2. Intriguingly, our DNA microarray studies did uncover one cDNA (clone 44) that contained a 65-amino acid region that showed 54% identities to a portion of the ASG2 sequence. Thus, the translated clone 44 was ASG2-like. This region, somewhat surprisingly, did not correspond to the repetitive domain within ASG2, a region reported to share similarities with elastin and flagelliform spider silk protein repetitive sequence domains (16). Presently, it is unclear whether the product of clone 44 localizes to a specific fiber type or which silk gland expresses this gene. Attempts to use either MS or MS/MS analyses of tryptic peptides obtained from dissolved attachment discs, wrapping silk, egg case fibers, gum-footed lines, and dragline silk did not reveal peptide masses or amino acid sequences that matched theoretical digestion of the translated cDNA from clone 44. Further biochemical analysis will be required to determine the expression pattern of this gene and whether its product is extruded to coat fibers.
species within the gumfooted lines. Based upon our MS studies of solubilized gumfooted silk, however, it suggests this immunoreactive species is also distinct from AgSF1. In the future, it will be interesting to determine the identity of these other immunoreactive species and their sequence relationships to AgSF1.

The AgSF2 protein sequence has a single internal block repeat that is present in 4 copies. This block repeat is also distinct from other reported spider silk fibroin modules. Further structural studies will need to be performed to determine how this primary sequence folds into a secondary structure. Based upon our Western blot data, AgSF2 was shown to have a substantially smaller molecular mass relative to traditional fibroin family members, which have masses in the range of 250 to 350 kDa. Thus, despite the presence of internal blocks repeats, whether AgSF2 can self-assemble into fibers will require further investigation. We have not yet attempted to spin recombinant AgSF2 proteins into synthetic fibers. The observation that AgSF2 is extruded into a fast drying glue-like interface material that serves to fuse individual scaffolding fibers could imply that it functions as a novel bioadhesive during web construction to help enhance the structural integrity of the web. Although we made numerous attempts to determine whether AgSF2 is N- or O-glycosylated in vivo, our MS and MS/MS studies did not support the presence of glycans covalently attached to AgSF2 (data not shown). Because highly glycosylated proteins, such as mucins are known to have bioadhesive properties, the lack of detectable glycans on AgSF2 tryptic peptides implies it has evolved a different molecular mechanism to interact with the protein chains of major ampullate silk fibroins. AgSF2 was observed to contain glycine and asparagine content that approaches 43%. One possibility could involve the side chain atoms of asparagine, which could lend themselves to hydrogen bond with other protein chains. In addition, two cysteine residues were present within the C-terminal region of AgSF2, but whether these residues participate in interdisulfide bond formation with

---

**FIGURE 7. Synthetic fibers spun from purified recombinant AgSF1 contain amorphous domains.**

A. XRD two-dimensional pattern of a single as-spun AgSF1 fiber. B. XRD two-dimensional pattern of the AgSF1 synthetic fiber after 3.5 × post-spin draw treatment. C and D, one-dimensional radial intensity profile of the azimuthally integrated, whole two-dimensional pattern from the as-spun and 3.5 × post-spin fibers, respectively. E and F, one-dimensional azimuthal intensity profile of the radially integrated diffraction rings at d ~ 4.5 Å and 10 Å for the as-spun and 3.5 × post-spin fibers, respectively.
Spider Glues Interconnect Dragline and Wrapping Silk Fibers

the conserved cysteine residue in the C terminus of MaSp1 or MaSp2 remains to be determined.

Although AgSF1 and AgSF2 are co-extruded into connection joints, their protein chain lengths and internal block repeat sequences, in addition to their N and C termini, were substantially divergent from each other and spidroin family members. For example, spidroin family members, such as PySp1 and PySp2, which are known as glue silk fibroins, contain significantly distinct repeat modules (9, 33, 34). Based upon these differences, it implies distinct roles for the AgSFs and the presence of at least two distinct glue protein families in the connection joints. Furthermore, the presence of two distinct block repeats in the AgSF1 protein chain also implies these regions perform distinct biological functions. One of the internal block repeats, the 36-amino acid segment, could be divided into the following submotifs: 1) five tandem copies of the pentameric sequence X1_PGGX2; 2) one hexameric sequence RPGGKG; and 3) one pentameric sequence QPGYY. These 36-amino acid repeats, the 36-amino acid segment, could be divided into the following submotifs: 1) five tandem copies of the pentameric sequence X1_PGGX2; 2) one hexameric sequence RPGGKG; and 3) one pentameric sequence QPGYY. These 36-amino acid block repeats were iterated 11 times, comprising over 55 copies of the X1_PGGX2 subrepeats; these pentameric sequences resemble VPGXG repeats found within elastin (35). Structurally, VPGXG iterations have been proposed to adopt a long-range β-spiral conformation with spring-like structures (36, 37). Whether iterations of X1_PGGX2 sequences fold into a similar secondary structure is unclear, but the PG residues and their presence in the pentameric modules strikingly resemble the VPGXG and GPGQQ modules of elastin and the flagelliform fibroin in spiral capture silk, respectively (13, 35). Despite these similarities, there are some differences between the VPGXG and X1_PGGX2 modules. One difference is the hydrophobic nature of the valine elastin relative to the hydrophilic properties of X1 (X1 = Gln, Glu, or Arg) in AgSF1. Differences in chemical properties are also present within the last residue of the pentameric sequence, where elastin contains the hydrophobic glycine residue and X2 was polar (X2 = Asn or Ser). Elastin also has been established to form intermolecular cross-linking via conserved lysine residues within its amino acid sequence. The conservation and spacing of the lysine residues within the RPGGKG motifs of AgSF1 could point to a similar function, providing an intermolecular cross-linking hub to enhance molecular alignments. However, based upon our biochemical studies, we have not been able to provide direct evidence that these residues are oxidized via chemical or enzymatic processes. The third submodule QPGYY, which contains dityrosine residues in tandem, was highly conserved throughout the 36-amino acid block repeats. Aromatic residues have been reported to be important mediators of protein assembly, and specifically tyrosine residues, have been reported to participate in intermolecular cross-linking (38–41). Thus, it raises the question whether these residues have an important role in controlling the nucleation, aggregation, polymerization, and/or intermolecular cross-linking of AgSF1 monomers during the extrusion process.

The exact mechanism that spiders use to assemble glue proteins into fibrous-like materials is highly debatable. Previous morphological studies of the spinneret silk spigots from black widow spiders show the posterior lateral spinneret contains two aggregate gland spigots with large flap-like openings, suggesting the glue-like proteinaceous material exits via a spraying process (20). Our microdissections of the aggregate glands support this report, revealing a large duct that exits the spinneret through a large flap-like spigot structure (19). Thus, unlike traditional spigots, which extrude threads, the proteins released from the aggregate gland appear to experience a different mechanism of self-assembly relative to fibroins spun by the major ampullate gland. Spraying the liquid substance through the large flap-like structure implies the secreted protein mixture is capable of undergoing assembly after extrusion. Preliminary data that removes the Sumo-tag from the recombinant AgSF1G+XPGXG followed by gentle agitation, resulted in protein aggregation in vitro.6 Thus, the observation that AgSF1G+XPGXG monomers can spontaneously assemble is consistent with the assertion that extrusion, followed by extension of the liquid via the movement of the abdomen, can lead to fiber formation. In agreement with this proposal is the ability of AgSF1G+XPGXG recombinant proteins to spontaneously form aggregates without shear as well as form synthetic fibers with shear, and the presence of small diameter fibers emerging from centralized regions of the connection joint material. In light of our findings, these studies provide the first insight into the diverse architectures of spider glue protein sequences that function to fasten web and wrapping fibers, uncovering a rich source of new structural biomaterials.

Acknowledgments—We thank Dr. Jason Li, Asylum Research, for technical assistance on the AFM imaging and Jeff Yarger, Arizona State University, Department of Chemistry & Biochemistry, for help with the x-ray diffraction data collection. Use of the Advanced Photon Source was supported by the United States Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. DE-AC02-06CH11357. Use of the BioCARS Sector 14 was also supported by grants from the National Center for Research Resources Grant 5P41RR007707 and NIGMS Grant 8P41GM103543 from the National Institutes of Health.

REFERENCES
1. Foelix, R. (1996) Biology of Spiders, Oxford University Press, New York
2. Kovoor, J. (1987) in The Ecophysiology of Spiders (Nentwig, W., ed) pp. 160–186, Springer-Verlag, New York
3. Guerette, P. A., Ginzinger, D. G., Weber, B. H., and Gosline, J. M. (1996) Silk properties determined by gland-specific expression of a spider fibroin gene family. Science 272, 112–115
4. Gatesy, J., Hayashi, C., Motriuk, D., Woods, I., and Lewis, R. (2001) Extreme diversity, conservation, and convergence of spider silk fibron sequences. Science 291, 2603–2605
5. Motriuk-Smith, D., Smith, A., Hayashi, C. Y., and Lewis, R. V. (2005) Analysis of the conserved N-terminal domains in major ampullate spider silk proteins. Biomacromolecules 6, 3152–3159
6. Sponner, A., Unger, E., Grosse, F., and Weisshart, K. (2004) Conserved C-termini of Spidroins are secreted by the major ampullate glands and retained in the silk thread. Biomacromolecules 5, 840–845
7. Hu, X., Lawrence, B., Kohler, K., Falick, A. M., Moore, A. M., McMullen, E., Jones, P. R., and Viera, C. (2005) Araneoid egg case silk. A fibroin with novel ensemble repeat units from the black widow spider, Latrodectus hesperus. Biochemistry 44, 10020–10027
8. Hayashi, C. Y., Blackledge, T. A., and Lewis, R. V. (2004) Molecular and mechanical characterization of aciniform silk. Uniformity of iterated se-

6 K. Vasanthavada, X. Hu, T. Tuton-Blasingame, Y. Hsia, S. Sampath, R. Pacheco, J. Freeark, A. M. Falick, S. Tang, J. Fong, K. Kohler, C. La Mattina-Hawkins, and C. Viera, unpublished data.
Spider Glues Interconnect Dragline and Wrapping Silk Fibers

TuSp1, ECP-1, and ECP-2. Biochemistry 45, 3506–3516
25. Pouchkina, N. N., Stanchev, B. S., and McQueen-Mason, S. J. (2003) From EST sequence to spider silk spinning. Identification and molecular characterization of Nephila senegalensis major ampullate gland peroxidase NsPox. Insect Biochem. Mol. Biol. 33, 229–238
26. Sandberg, L. B., Leslie, J. G., Leach, C. T., Alvarez, V. L., Torres, A. R., and Smith, D. W. (1985) Elastin covalent structure as determined by solid phase amino acid sequencing. Pathol. Biol. 33, 266–274
27. George, R. A., and Heringa, J. (2000) The REPRO server. Finding protein internal sequence repeats through the Web. Trends Biochem. Sci. 25, 515–517
28. van Beek, J. D., Hess, S., Vollrath, F., and Meier, B. H. (2002) The molecular structure of spider dragline silk: Folding and orientation of the protein backbone. Proc. Natl. Acad. Sci. U.S.A. 99, 10266–10271
29. Geiger, T., and Clarke, S. (1987) Deamidation, isomerization, and racemization at asparagine and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. J. Biol. Chem. 262, 785–794
30. Hsia, Y., Gnesa, E., Pacheco, R., Kohler, K., Jeffery, F., and Vierra, C. (2012) Synthetic spider silk production on a laboratory scale. J. Vis. Exp. 65, 4191
31. Keum, J. K., and Song, H. H. (2005) Thermal deformations of oriented noncrystalline poly(ethylene terephthalate) fibers in the presence of mesophase structure. Polymer 46, 939–945
32. Murthy, S. (2004) Recent developments in polymer characterization using x-ray diffraction. Rigaku J. 21, 15–24
33. Geurts, P., Zhao, L., Hsia, Y., Gnesa, E., Tang, S., Jeffery, F., La Mattina, C., Franz, A., Larkin, L., and Vierra, C. (2010) Synthetic spider silk fibers spun from pyriform spidroin 2, a glue silk protein discovered in orb-weaving spider attachment discs. Biomacromolecules 11, 3495–3503
34. Perry, D. J., Bittencourt, D., Siltberg-Liberles, J., Rech, E. L., and Lewis, R. V. (2010) Pyriform spider silk sequences reveal unique repetitive elements. Biomacromolecules 11, 3000–3006
35. Kushner, A. M., and Guan, Z. (2011) Modular design in natural and bio-mimetic soft materials. Angew. Chem. Int. Ed. Engl. 50, 9026–9057
36. Urry, D. W. (1988) Entropic elastic processes in protein mechanisms. II. Simple (passive) and coupled (active) development of elastic forces. J. Protein Chem. 7, 81–114
37. Urry, D. W. (1988) Entropic elastic processes in protein mechanisms. I. Elastic structure due to an inverse temperature transition and elasticity due to internal chain dynamics. J. Protein Chem. 7, 1–34
38. Brown-Augusburger, P., Tisdale, C., Broekelmann, T., Sloan, C., and Mechem, R. P. (1995) Identification of an elastin cross-linking domain that joins three peptide chains. Possible role in nucleated assembly. J. Biol. Chem. 270, 17778–17783
39. Reches, M., and Gazit, E. (2003) Casting metal nanowires within discrete self-assembled peptide nanotubes. Science 300, 625–627
40. Gazit, E. (2002) A possible role for π-stacking in the self-assembly of amyloid fibrils. FEBS J. 16, 77–83
41. Souza, J. M., Giasson, B. I., Chen, Q., Lee, V. M., and Ischiropoulos, H. (2000) Dityrosine cross-linking promotes formation of stable α-synuclein polymers. Implication of nitrate and oxidative stress in the pathogenesis of neurodegenerative synucleinopathies. J. Biol. Chem. 275, 18344–18349