

Modifications of spider silk sequences in an attempt to control the mechanical properties of the synthetic fibers

Florence Teulé · William A. Furin · Alyssa R. Cooper ·
Joshua R. Duncan · Randolph V. Lewis

Received: 5 December 2006 / Accepted: 26 February 2007 / Published online: 17 July 2007
© Springer Science+Business Media, LLC 2007

Abstract Bacteria were genetically engineered to produce two spider silk protein variants composed of basic repeat units combining a flagelliform elastic motif ([GPGGX]₄) and a major ampullate silk strength motif ([linker/poly-alanine]. The secondary structures of the pure recombinant proteins in solution were determined by circular dichroism. The data presented suggest that the nature of the 5th and 10th amino acid (X) in the [GPGGX]₂ elastic motif and temperature have an impact on the amount of β -sheet structures present in the proteins. More specifically, increasing temperatures seem to be positively correlated with β -sheet formation for both proteins and this state is irreversible or reversible when both X (5th and 10th) in the elastic motif are hydrophilic or hydrophobic respectively. Moreover, each pure silk-like protein was able to spontaneously self-assemble into films from aqueous solutions. Two kinds of synthetic fibers were made by pulling fibers from these preassembled films as well as spinning fibers from each protein resolubilized in HFIP. The mechanical data show that the pulled fibers are far tougher than the spun fibers suggesting a better fiber organization.

Introduction

Spider silks have great potential to provide a new generation of bio-based materials for applications ranging from medical (micro-sutures, artificial ligaments, tendons, and drug-delivery coatings) to military (body armor, light weight gear) to civilian (textiles). In order to custom design and produce biomaterials, a critical step is to relate structure to function in these fiber proteins. For this purpose, spider silks are the model of choice because of the wide variety of silks spun, their different uses, and thus their differences in mechanical properties.

Of all the orb weaver-spun silks, the major ampullate silk, or dragline silk, from *Nephila clavipes* has been most extensively studied because of its extreme strength (4 GPa) [1, 2] and moderate elasticity (35%) [2, 3] resulting in an impressive toughness (160 MJ m⁻³) [3, 4]. Overall, this performance is believed to be dictated by the silk proteins' primary structures, more specifically, by the presence and combination of key amino acid motifs that have previously been identified. Crystalline forming motifs such as poly-alanine (A)_n ($n \leq 9$) present in the highly repetitive MaSp2 protein, one of the two proteins forming the major ampullate silk [5, 6], are thought to confer strength to the fiber. Structural studies have shown that in the fiber, these poly-alanine motifs form crystalline β -sheets aligned parallel to the fiber axis [7, 8]. Another silk of interest is the extremely elastic flagelliform silk, also called "viscid" silk, which is composed of a single highly repetitive protein containing no poly-alanine motifs [9, 10]. The basic flagelliform protein (Flag) repeat unit has been characterized, and it is mainly composed of a (GPGGX)_n motif ($43 \leq n \leq 63$, and usually X = A/V, or Y/S) juxtaposed to a (GGX)_n motif ($n = 12$) and a spacer sequence containing non-traditional silk amino acids [9, 10]. These (GPGGX)_n motifs are believed to adopt

Electronic supplementary material The online version of this article (doi:10.1007/s10853-007-1642-6) contains supplementary material, which is available to authorized users.

F. Teulé (✉) · W. A. Furin · A. R. Cooper ·
J. R. Duncan · R. V. Lewis
Department of Molecular Biology (Dept. 3944),
University of Wyoming, 1000 E. University Avenue,
Laramie, WY 82071, USA
e-mail: fteule@uwyo.edu

β -turn structures, and a series of these motifs would form spring-like spirals, or nanosprings, conferring elasticity to the fiber [11–13]. Structural studies of the recombinant polypeptide **1** showed that its sequence, [(GPGGS GPGGY)₂ GPGGK]₁₁, is composed of type II β -turns [13]. The central (PG) of the (GPGGX) consensus motif, also present in several elastomeric proteins such as elastin [14–16], and glutenin [17], has been found to be favorable for the formation of type II β -turns [18, 19]. Thus, the mechanism of silk elasticity is thought to resemble the one described for these β -turn forming polypeptides and seems to be entropically driven [3, 13]. Therefore, the extremely high number of (GPGGX)_n motifs in the Flag protein is believed to be responsible for the extraordinary elasticity measured for the flagelliform silk (200%) [1, 2]. This argument is reinforced by the fact that a lower amount of this motif in the dragline protein MaSp2, and its absence in the minor ampullate silk proteins Misp1 and Misp2, seem to result in much lowered elasticity of these two silk fibers [6].

We have engineered and produced recombinant repetitive silk-like proteins with basic repeats containing variants of the elastic (GPGGX) motifs found in the Flag protein, either (GPGGA)₄ (=A1 motif) or [(GPGGY) (GPGGS)]₂ (=Y1 motif) combined with a strength motif [Linker-(A)₈] (=S8 motif) found in the dragline silk MaSp2 protein. We have produced and purified the two proteins (A1S8₂₀ and Y1S8₂₀), confirmed their primary structures by amino acid analyses, characterized their secondary structures using circular dichroism, and generated two types of fibers from both proteins either from aqueous solutions or organic solvents. The synthetic fibers were mechanically tested to give us an evaluation of the performance resulting from the combination of either of the Flag protein elastic motifs, and the dragline silk protein strength motif.

Materials and methods

Gene construction and cloning in pBluescript®II SK+ in *E. coli*

Three sets of complementary synthetic oligonucleotides (roughly 60 bp each, Electronic Supplementary data or ESM, Fig. 1) were designed (Dr M. B. Hinman): (1) flagelliform putative elastic encoding motifs A1 and Y1 (coding respectively for (GPGGA)₄ and (GPGSGPGGY)₂), (2) the dragline silk putative strength encoding motif S8 (coding for a linker-polyalanine: (GGPSGPGS(A)₈)). These three types of synthetic oligonucleotides (Y1, A1 and S8) were synthesized and assembled as double stranded cassettes each cloned in pBluescript®II SK+ (Stratagene) at the *Hind* III/*Sma* I sites (Midland Certified Reagents Inc, Texas). These double stranded sequences were engineered with a 5'*Xma* I,

and a 3'*Bsp* I restriction sites allowing the use of a previously described multimerization or doubling strategy based on compatible but non-regenerable restriction sites [20]. The two synthetic genes constructed by successive doubling of each different basic repeat unit were called (A1S8)₂₀ and (Y1S8)₂₀. These sequences were cloned in a pBluescript®II SK+ plasmid and used to transform electrocompetent XL1-Blue cells (Stratagene). The recombinant plasmids containing the final (A1S8)₂₀ and (Y1S8)₂₀ silk-like sequences (2,300 bp each) were partially sequenced on both strands using vector specific primers.

Cloning of the silk-like sequences in the pET19b (Kan^R) expression vector

The pET19b (Kan^R) expression vector was previously obtained by replacing the *Bpu*1102/*Alw*NI fragment containing the ampicillin resistance gene in pET19b (Amp^R) (Novagen) by the *Bpu*1102/*Alw*NI fragment containing the kanamycin resistance gene (Kan^R) from pET26b (-) (Novagen) (Dr M. B. Hinman, unpublished data).

The plasmid DNAs of the bacterial pBluescript clones containing the confirmed (A1S8)₂₀ and (Y1S8)₂₀ silk-like sequences, as well as the pET19b (Kan^R) expression vector, were isolated using an alkaline lysis protocol [21].

The expression vector and each recombinant silk plasmid were subjected to a *Bam* HI restriction enzyme digestion (New England Biolabs Inc.). The totality of the restriction enzyme digestion sample was subjected to electrophoresis in a 0.8% agarose gel. The silk inserts and the pET19b (Kan^R) vector were purified by electroelution following standard protocols [22]. The DNA fragments were then recovered by addition of ammonium acetate to a final 1.5 M concentration, followed by the addition of two volumes of cold ethanol. The samples were placed at -80 °C for 45 min and the purified DNA fragments were recovered by centrifugation at 18,000g for 25 min at room temperature. The supernatant was discarded and each DNA pellet was dried and resuspended in 30 μ L of TE buffer (10 mM Tris-HCl/1 mM EDTA, pH 8).

Each purified *Bam* HI silk fragment was ligated to the purified *Bam* HI pET19b (Kan^R) expression vector using T4 DNA ligase following the recommendations of the Manufacturer (Promega). Each ligation reaction was used to transform electrocompetent *E. coli* XL1-Blue cells (Stratagene). The plasmid DNA from the recombinant clones were extracted by alkaline lysis and characterized by restriction enzyme digestion with *Bam* HI (New England Biolabs Inc.). These plasmids were also subjected to a restriction enzyme digestion with *Xma* I (New England Biolabs Inc.) to verify the insert's orientation. The two kinds of recombinant pET clones displaying the (A1S8)₂₀ or (Y1S8)₂₀ silk insert in the right orientation were

sequenced partially on both strands. The selected clones were named pE(A1S8)_{20x} and pE(Y1S8)_{20x}.

Once the silk insert sequence was confirmed, these two selected recombinant plasmid DNAs were introduced by electroporation in the *E. coli* cell line BL21 (DE3) (Novagen) for expression purposes. The recombinant clones were once again characterized by restriction enzyme digestion with *Bam* HI (New England Biolabs Inc.) to confirm the presence of the silk inserts. The selected recombinant plasmids containing the silk insert in the right orientation were named pET(A1S8)_{20x} and pET(Y1S8)_{20x}. These plasmids were partially sequenced on both strands.

Gene expression studies

The pET recombinant clones in BL21 (DE3) cells (pET(A1S8)_{20x} and pET(Y1S8)_{20x}) were cultured overnight in 10 mL of LB containing 50 µg/mL kanamycin in a shaking incubator at 37 °C. The overnight culture was used to inoculate 1 L of LB containing 50 µg/mL of kanamycin. The cultures were induced by addition of 1 mM IPTG (Isopropyl-β-D-thiogalactopyranoside; Biochemica & Synthetica, Switzerland), when the optical density at 600 nm (OD₆₀₀) of the culture had reached 0.6–0.8. The cells were then harvested by centrifugation at 5,300g/25 min about 3–4 h after induction. The media were discarded, and the cell pellets were washed once in sterile distilled water. After a second centrifugation at 3,300g for 15 min, and removal of the supernatant, the masses of the cell pellets were calculated and the samples were stored at –80 °C.

E. coli cell lysis

The cell pellets were resuspended in 1× lysis buffer (50 mM Tris–HCl pH 8, 10 mM MgCl₂, 10 mM NaCl) at 3 mL of buffer per every gram of cell. Lysozyme (Sigma) was added to each sample to a final concentration of 0.2 mg/mL and the samples were incubated on ice for 30 min swirling periodically. At this stage, 1 mM PMSF (phenylmethylsulphonyl fluoride; Sigma) was added to the lysates to prevent protein degradation. Then 1.5 g of deoxycholic acid (MP Biomedicals LLC) were added per gram of cells and the lysates were incubated 20 min at 37 °C. At this point, 0.02 mg of DNase I (Sigma) was added to the cell lysates per gram of cell and the samples were incubated at room temperature for 30 min on a platform shaker. The lysates were then subjected to centrifugation at 3,300g for 15 min to pellet the cellular debris. After a heat-treatment at 80 °C for 10 min, the recovered lysates, or supernatants, were again subjected to centrifugation at 3,300g for 15 min to pellet the denatured proteins. The cleared heat-treated lysates were stored at –80 °C until use for protein purification.

Protein purification

Sample preparation

The heat-treated cell lysates containing the recombinant His-tagged silk proteins were diluted in 1× binding buffer (20 mM Tris–HCl pH 8/0.5 M NaCl/5 mM imidazole) at a 1:4 ratio (v/v).

IMAC (immobilized metal ion affinity chromatography)

Each sample was loaded on a 10 mL polyrep chromatography column (Biorad) containing 2 mL (resin bed volume, BV) of Ni-NTA His-Bind resin (Novagen) previously washed once with 10 BV of water, and equilibrated with 5 BV of 1× binding buffer. The proteins were eluted using an imidazole step gradient. The concentrations of imidazole of the 1× binding buffer used in the successive washes were adjusted to 20 mM, then 40 and 50 mM. The recombinant His-tagged silk proteins were eluted when using imidazole concentrations equal or greater to 60 mM (80 and 250 mM were best). The resin was stripped of the nickel ions using a 1× strip buffer (100 mM EDTA, 0.5 M NaCl, 20 mM Tris–HCl pH 8). This strip fraction along with the other fractions were collected and saved for analyses.

Protein characterization

SDS-PAGE analyses

The heat-treated protein extracts and the purified protein fractions were analyzed by SDS-PAGE. For all SDS-PAGE analyses, 4% stacking and 10% separating polyacrylamide gels containing 0.1% SDS were made in Tris–HCl buffers (Mini PROTEAN[®]3 Cell protocol for SDS-PAGE buffer system, Biorad). The composition of the sample buffer and the 5× electrode buffer used are the ones recommended by the manufacturer (Biorad). Typically, 10 µL of protein sample were subjected to SDS-PAGE analysis. A Kaleidoscope[™] Prestained Standards (Biorad, 10 µL) or a Precision Plus Protein Dual Color Standard (Biorad, 8 µL) were used as a molecular weight marker. All electrophoreses were performed using the Mini PROTEAN[®]3 Cell (Biorad) at a constant voltage of 80 V.

Staining of polyacrylamide gels

After SDS-PAGE analysis, the gels were stained with Coomassie Brilliant Blue (R-250) dye according to the method published [23]. The stained gels were placed in a 10% glycerol solution for 1 h before being dried between two sheets of Ultra Clear Cellophane (Research Products

International Corp., Mount Prospect, Illinois) using a plexiglass frame.

Western blot analyses

The proteins samples separated by SDS-PAGE were transferred to a PVDF/Immobilon™ P transfer membrane (Millipore) by electroblotting using the Mini Trans-Blot Electrophoretic Transfer Cell (Biorad). Blots were set up as specified by the manufacturer (Biorad). All transfers were performed overnight at room temperature under a constant current of 25 mA. After fixation of the proteins, the membranes were subjected to Western blot analyses using the *His-Tag*® antibody (Novagen) directed against the (histidine)₁₀ tag ((His)₁₀) present in the amino terminus of the fusion proteins. The secondary antibody used was the Goat Anti-Mouse IgG HRP conjugate (H + L) (Novagen). All Western blot analyses were performed according to the protocols described by the manufacturers. Chemiluminescent detections were performed using the ECL™ Western Blotting Detection Reagents (Amersham Biosciences). The membranes were then wrapped in plastic wrap and placed in a light proof cassette against a high performance chemiluminescence film (Hyperfilm™ ECL, Amersham Biosciences) to visualize the proteins of interest.

Amino acid analyses

The purified protein samples were dialyzed against 5 mM ammonium bicarbonate and lyophilized. The samples were treated using the Waters AccQ-Tag™ Method for hydrolysate amino acid analysis (Millipore) and then subjected to HPLC using a Hitachi LC 6500. The amino acid compositions were expressed in mole % (ESM, Table 1).

Circular dichroism (CD)

For CD, the purified proteins samples were dialyzed extensively against 5 mM Tris–HCl pH 8 or 0.1× PBS (Phosphate Buffered Saline: 13.7 mM NaCl/0.27 mM KCl/1 mM Na₂HPO₄/0.18 mM KH₂PO₄) five times and subsequently concentrated by dialysis against a 20% PEG (polyethylene glycol)/5 mM Tris–HCl pH 8 or 20% PEG/0.1× PBS using a dialysis tubing with a smaller molecular weight cut-off (MWCO: 3,500 Daltons). The final CD samples concentrations were estimated by BCA assay (Pierce) using a BSA standard as well as “home made” standards of A1S8₂₀ and Y1S8₂₀ proteins (2 mg of each pure lyophilized A1S8₂₀ or Y1S8₂₀ protein were resolubilized in 1 mL of 5 mM Tris–HCl pH 8 or 0.1× PBS). The concentrations of the pure soluble A1S8₂₀ dilutions used for CD were 0.98 mg/mL in Tris–HCl pH 8 and 0.35 mg/mL in 0.1× PBS. The concentrations of the

pure soluble Y1S8₂₀ dilutions used for CD were 0.75 mg/mL in Tris–HCl pH 8 and 0.54 mg/mL in 0.1× PBS. The CD spectra were recorded on a JASCO-810 spectropolarimeter using the Spectra Manager for Windows 95/NT software (Version 1.18.00). A total of 8 spectra accumulations were recorded from 185 nm to 260 nm with a resolution of 0.1 nm and a path length of 0.01 cm. Melt and anneal experiments were conducted successively from 0 °C to 85 °C and 85 °C to 0 °C at 5 °C intervals. The data obtained were submitted to the following online site for analyses: www.cryst.bbk.ac.uk/cdweb/html/home.html [24]. We used the SELCON 3, CONTIN and CDSSTR methods for protein secondary structure predictions (see [24] for full references).

Production of synthetic fibers

Two methods of fiber production were used for both types of synthetic proteins (Fig. 1). The first one relied on the natural ability of these purified synthetic fibers to spontaneously form fibers in aqueous solutions, and the second relied on the wet spinning/extrusion of a silk dope made in an organic solvent.

Hand pulled fibers

The Y1S8₂₀ pure protein spontaneously formed an oily looking film at the surface of the elution or strip solutions collected in a glass dish, thus naturally separating from the solution. We used forceps to pull the edge of the Y1S8₂₀ film that lifted into single fibers (Fig. 1a). While A1S8₂₀ did not spontaneously form a film at the surface of the elution or strip fractions collected, we noticed a film-like structure at the bottom of the dish. Thus shaking vigorously

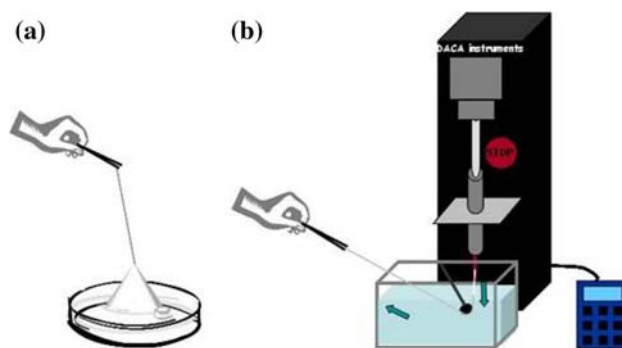


Fig. 1 Synthetic fiber production. (a) Hand pulled fibers from the protein pure fraction: the fibers were hand pulled from a self-assembled film using forceps. (b) Wet spinning of fibers using an extruder/spinneret: the spinning dope loaded in a syringe was placed in a computer controlled spinneret and extruded at a constant speed into a coagulation bath. The extruded fibers were recovered using forceps

the dish to mix the contents resulted in the surfacing of preformed A1S8₂₀ films that could then be pulled into fibers using forceps (Fig. 1a).

Silk dope preparation and fiber wet spinning

The pure lyophilized proteins were resolubilized in 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP; TCI America, Portland OR) to make a 25–30% (w/v) spinning dope. The fibers were extruded using a DACA Instruments SpinLine system (DACA Instruments, Santa Barbara CA) provided by Nexia Biotechnologies (Montreal, Canada). The spinneret was composed of a 1 mL Hamilton Gastight[®] syringe (Hamilton company) mounted with a 10 cm long “red” PEEK tubing (Upchurch Scientific) with an internal diameter of 0.127 mm. The syringe was loaded with the silk dope and placed in the barrel of the extruder. Prior to spinning, we added about 15% of H₂O to the dope and slowly mixed it with the dope in the syringe. The silk fibers were extruded at a plunger speed of 0.6 mm/minute into a 90% isopropyl alcohol coagulation bath (Fig. 1b). The extruded synthetic fibers were collected from the bath with forceps, cut and mounted on testing cards.

Mechanical testing and data analyses

All fibers (pulled and spun) were mounted on testing cards. The initial length of the fiber to be tested was 15 mm. Three to five pictures of each synthetic fiber were taken along the fiber axis under a light microscope using the 20× and 40× magnifications. The average diameters of the fibers were calculated after the pictures were imported into the ImageJ software prior to testing. We tested all the fibers produced using a custom made 10 g load cell (Transducer Techniques, Temecula CA) mounted on a MTS Synergie 100 system (MTS Systems Corporation, Eden Prairie MN). The mechanical data was recorded at a strain rate of 5 mm/min using a frequency of 250 MHz. Parameters such as load and elongation were recorded and allowed the successive calculation of the engineered stresses and strains. The stress/strain curves as well as the polynomial regressions of the curves necessary to calculate the toughness (or energy to break/work of rupture), Young's Modulus (stiffness = initial slope of the curve), and determination of parameters such as maximum stress and maximum extension (=maximum % strain), were performed using MATLAB (Version 7.1). Any mechanical data obtained having a polynomial fit for the data points with an $R^2 \leq 0.9$ was eliminated from the study. The best fits for the data points recorded were obtained using polynomial regressions of 8th or 9th order.

Results and discussion

Gene expression and protein production

We used sequences encoding flagelliform silk-like putative elastic motifs (A1 or Y1) and a dragline silk putative strength motif (S8) to build two spider silk-like genes. These three synthetic spider silk-like sequences were used to build two types of basic repeat units: (A1S8) and (Y1S8). The only difference between the (A1S8) and (Y1S8) basic repeat units is the nature of the amino acids present in the 5th and 10th positions in the [GPGGX]₄ elastic motif. More specifically, the elastic motif encoded by the A1 sequence is [(GPGGA) (GPGGA)]₂ whereas the one encoded by the Y1 sequence is [(GPGGY) (GPGGS)]₂. Such motifs naturally occur in the native flagelliform sequence [10]. However, in the putative elastic motif of the native flagelliform silk protein, when there is an alanine (A) in 5th position, there is usually a valine (V) in the 10th position. We chose to alternate alanine (A) in the 5th position with alanine (A) in the 10th position, as this motif is also present in the native flagelliform putative elastic motifs, and since this change does preserve the hydrophobic nature of the amino acid in the 10th position.

We used the doubling strategy described earlier [20] to rapidly increase the sizes of these two kinds of silk basic units. The spider silk-like constructs were sequenced after each doubling experiment to avoid the selection of mutations that could occur during cloning. The final sizes of the synthetic silk repeats were 2,300 bp for both (A1S8)₂₀ and (Y1S8)₂₀. These two spider silk-like constructs, initially built in the pBluescript[®]II SK vector, were successfully cloned in the pET19b (Kan^R) expression vector and introduced in the *E. coli* cell line BL21 (DE3) for expression: clones pET(A1S8)_{20x} and pET(Y1S8)_{20x}. The cell lysates, or protein extracts, were heat-treated at 80 °C for 10 min as a first purification step. Spider silk-like proteins are usually fairly heat-stable, and such heat-treatments were first investigated as initial steps (95 °C/5 min) in the purification of spider silk-like proteins produced in plants [25]. This heat-treatment was also used successfully as a first step during the purification of MaSp 1-like homopolymer and copolymers produced in yeast (*Pichia pastoris*) [26]. In this case, during heat-treatment, most of the *E. coli* native proteins were denatured and could be eliminated by centrifugation. All of the recombinant spider silk-like proteins produced in *E. coli* are fusion proteins possessing a histidine tag ((His)₁₀) in their carboxyl terminus facilitating purification by ion metal affinity chromatography. The primary structures, or amino acid sequences, of the two spider silk-like recombinant proteins A1S8₂₀ (57.64 kDa) and Y1S8₂₀ (61.96 kDa) that were produced in *E. coli* are shown in Fig. 2a.

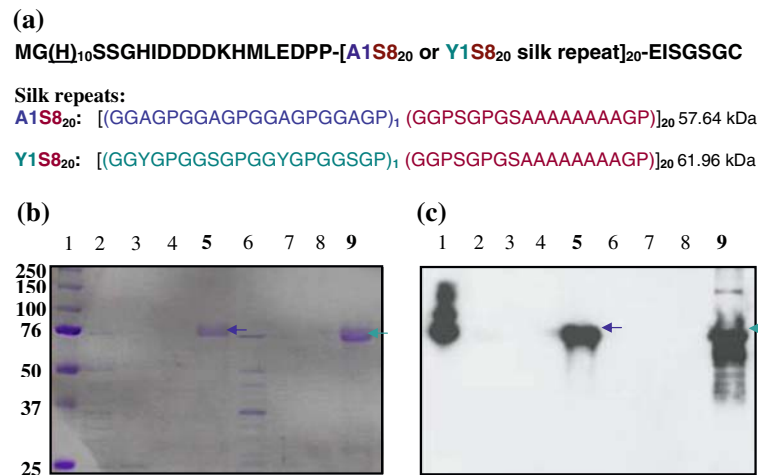


Fig. 2 Primary structures and Western Blot analyses of the Y1S8₂₀ and A1S8₂₀ recombinant proteins. (a) Sequences and molecular weights of the spider silk-like proteins produced in *E. coli*. The molecular weights of the fusion proteins are indicated next to each silk repeat. (b) SDS-PAGE analysis of the purification of A1S8₂₀ and Y1S8₂₀ from heat-treated extracts (stained with Coomassie Brilliant Blue). (c) Western blot analysis of (b) using the *His-Tag*[®] antibody as primary antibody and the Goat Anti-Mouse IgG HRP conjugate as

secondary antibody. In (b) and (c): Lane 1 is a Molecular weight marker; Lanes 2–5 show the A1S8₂₀ purification; Lanes 6–9 show the Y1S8₂₀ purification; Lanes 2 and 6 show the unbound protein fractions; Lanes 3 and 7 show the 40 mM imidazole washes; Lanes 4 and 8 show the 50 mM imidazole washes; Lanes 5 and 9 show the eluted/strip fractions. In both (b) and (c), the blue and green arrows point at the purified A1S8₂₀ and Y1S8₂₀ proteins respectively

Protein purification and characterization

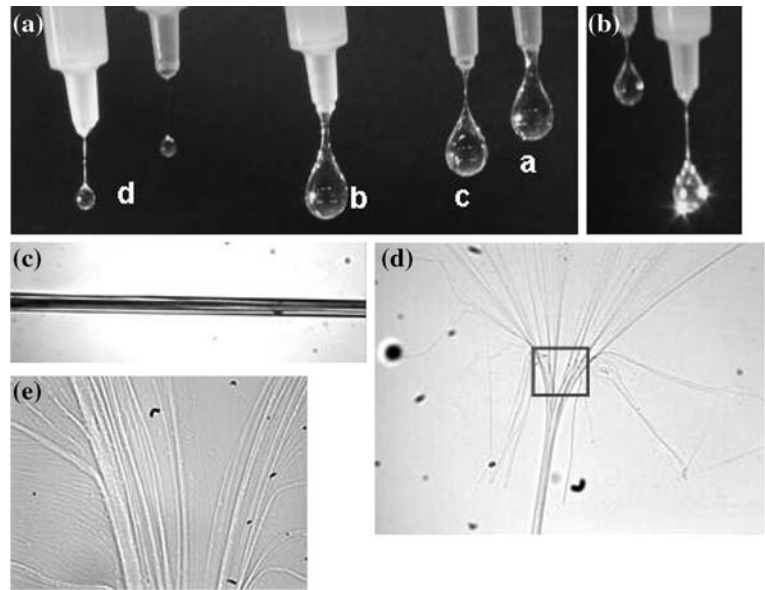
The A1S8₂₀ and Y1S8₂₀ proteins were easily purified with classic metal affinity chromatography and were recovered using as low as 60 mM imidazole in the elution buffer. The SDS-PAGE and Western Blot analyses of the purification of Y1S8₂₀ and A1S8₂₀ are shown in Fig. 2 (b and c). The two pure full-size proteins were visible when stained with Coomassie blue (Fig. 2b). Note that on the Western blot analyses, we can see both full-size proteins, as well as truncation products only for Y1S8₂₀ (Fig. 2c). Truncated proteins are likely the direct result of translation from a prematurely terminated transcript and are commonly observed for native [3, 10] or recombinant spider silk proteins [20, 27–30]. After expression, heat-treatment and affinity chromatography purification, we were able to recover of 7–10 mg/L of pure A1S8₂₀ or Y1S8₂₀ proteins. The results of the amino acid analyses of the purified A1S8₂₀ and Y1S8₂₀ proteins confirmed their identities and their high purity (ESM, Table 1).

While purifying Y1S8₂₀ and A1S8₂₀, we observed the formation of fibers at the bottom of the affinity chromatography columns as the elution or strip fractions were being collected (Fig. 3a and b). Samples of these in situ formed fibers were collected on a slide and observed under a light microscope (Fig. 3, c through e). It looked as though the proteins were forming a film at the periphery of an extremely viscous droplet. These fibers were drawn out of the column by the weight of the droplet that dangled at the

bottom of the forming fiber thus allowing the film at the periphery of the droplet to fuse into a single fiber (Fig. 3, d and e). Moreover, we determined that the aggregation of these proteins was concentration dependent as dilution of the original heat-treated protein extract loaded on the column resulted in no such fiber formation. We believe that fiber formation is also purity dependant as these fractions were extremely pure. It is important to stress the fact that these fibers were able to form in an aqueous solution close to physiological pH rather than in harsh solvents. More importantly, these aqueous solutions might help preserve the original secondary structures of the recombinant silk proteins that might be critical for the resulting mechanical properties of the fiber.

After further observation, we noticed that the eluted Y1S8₂₀ protein could self-assemble to form a very uniform film with an oily appearance that floated at the surface of the purified protein fraction. A major difference in behavior between the two pure proteins was that A1S8₂₀ did not form an apparent film at the surface of the pure protein fraction like the one observed for Y1S8₂₀ but would rather form a viscous layer at the bottom of the dish covered by a more fluid layer. However, by vigorously shaking the dish containing the pure A1S8₂₀ fraction, we were able bring the viscous layer to the surface as a somewhat broken up film. We used forceps to pull each type of films made by these two proteins into fibers (Fig. 1a) and mounted the “pulled fibers” on testing cards for mechanical testing. The film formed by Y1S8₂₀ could be reeled in and be wound

Fig. 3 Pictures of the *in situ* fibers formed on the column and their observation of under a light microscope. (a) The different stages of Y1S8₂₀ fiber formations are shown (a through d) while recovering the extremely pure fraction (strip fraction) from the affinity chromatography columns. (b) Close up on the fiber formed (40×). (c), (d), and (e) are pictures of the observations under a light microscope of the types of fibers collected from the column and mounted on a slide. (c) shows a Y1S8₂₀ fiber (40×), (d) shows an A1S8₂₀ fiber (20×), (e) is an enlargement (40×) of the part of the fiber in (d) indicated by the square box



around a tube very easily without breaking. The fibers pulled from the A1S8₂₀ small films were shorter, however, limited by the size of the broken film pieces.

Structural analyses and hypotheses correlating secondary structure and self assembly

Circular dichroism was used to determine the secondary structures of the A1S8₂₀ and Y1S8₂₀ silk-like proteins in aqueous solutions. The CD spectra of both melting (from 0 °C to 85 °C) and successive annealing (from 85 °C to 0 °C) for both proteins in the 5 mM Tris–HCl pH 8 are shown in Fig. 4. We chose to use a Tris–HCl buffer pH 8 for CD analyses since we observed protein self-assembly in the elution and strip buffers containing Tris–HCl pH 8. Even though these two proteins only differ by two amino acids (in the 5th and 10th positions in the [GPGGX]₄ repeat), there were substantial differences in the secondary structures observed for both proteins.

For A1S8₂₀, the CD analyses show that at 0 °C, the protein is mostly unordered (about 76.30% and 64.26% of random coils observed respectively in Tris–HCl pH 8 and PBS). Random coils are noticeable in the melt profile of A1S8₂₀ in Tris–HCl (Fig. 4, top left) by the presence of a lower maximum near 200 nm. As the temperature increases, the SELCON 3, CONTIN and CDSSTR analyses of the CD spectra for the A1S8₂₀ protein in Tris–HCl show a slight increase in helices (to a maximum of 10%). Although there seems to be a slight increase in sheets and turns, the percentages of these two species at higher temperatures indicated by the SELCON 3 and CONTIN methods are substantially lower than indicated by the CDSSTR method. Indeed, for both SELCON3 and CONTIN, the results of the analyses for A1S8₂₀ in Tris–HCl

from 65 °C to 85 °C show an average of 6.33% sheets, 9.10% turns, 8.30% helices, and 77.68% random coils while the ones obtained for the same temperature bracket with CDSSTR show 21.20% sheets, 17.80% turns, 8.60% helices, and 53.00% random coils. From the CD spectra of the A1S8₂₀ melt presented in Fig. 4, we can confirm that there is indeed an increase in sheets and turns (see the increase at 200 nm and the decrease at 220 nm for the sheets, and the increase at 210 nm for the turns). However, there is still a high content of random coils (lower maximum before 200 nm) thus we feel that the results given by SELCON 3 and CONTIN seem most probable for the A1S8₂₀ protein. The same trends were observed for this protein in PBS (ESM, Fig. 2). The successive annealing of the A1S8₂₀ protein from 85 °C to 0 °C clearly shows that the heat induced changes in secondary structure observed during the melting are almost totally reversible (Fig. 4, top right).

For Y1S8₂₀, the CD analyses show that at 0 °C in Tris–HCl, the protein is mostly unordered (about 55% in both buffers) but it contains less random coil than was observed for A1S8₂₀ (Fig. 4, bottom left). According to the melt spectra and the results of the SELCON 3, CONTIN, and CDSSTR analyses however, there are substantial changes in the secondary structures when the temperature increases to 85 °C. For Y1S8₂₀, in the melt profiles shown in Fig. 4, the maximum peak observed at 200 nm accompanied by a minimum at 220 nm are indicative of a noticeable increase in sheets (up to 31.32%), and the maximum peak observed at 210 nm (broadened by the sheet peak occurring at 200 nm) is indicative of an increase in turns (about 20%) that is visible as soon as the temperature reaches 15 °C. The same trend was observed for this protein in PBS (ESM, Fig. 2). The successive anneal experiment for Y1S8₂₀

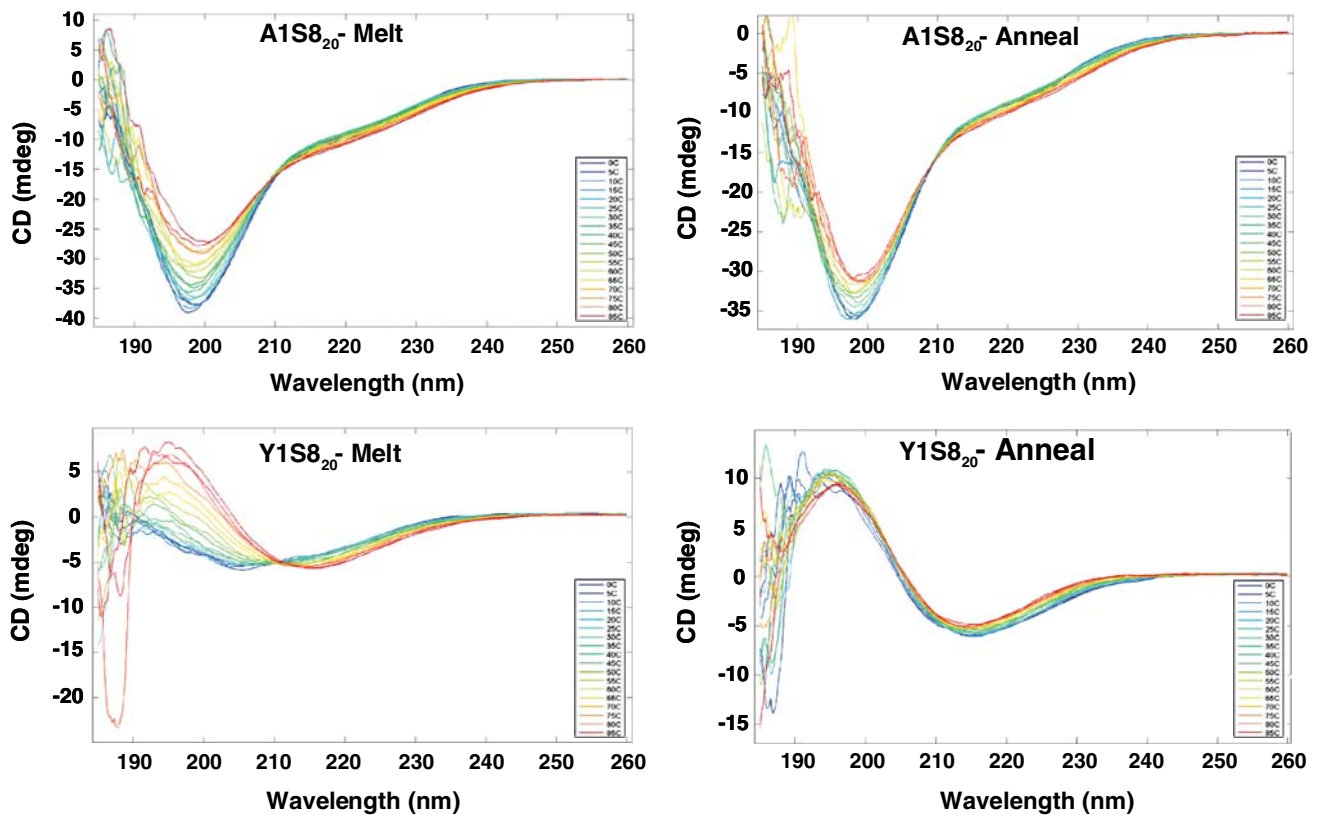


Fig. 4 Circular dichroism spectra for A1S8₂₀ and Y1S8₂₀. Melt (from 0 °C to 85 °C) and successive anneal (from 85 °C to 0 °C) spectra obtained for both proteins in 5mM Tris–HCl pH 8. Note that

the scale from one graph to another is not the same for the CD [mdeg] units (vertical axes)

(Fig. 4, bottom left) shows that once the secondary structures have formed during the melt, there is no reversion to the initial structure of the protein.

According to this data, for Y1S8₂₀, higher temperatures promote irreversible sheet and turn formation through hydrogen bonding that remain stable. Such irreversible temperature-induced β -sheet formation has also been observed in melt experiments of major and minor ampullate silk gland contents [31]. These structures for A1S8₂₀, though less pronounced, are reversible. Moreover, the fact that the amount of turns increase and remain stable for Y1S8₂₀ may indicate that the [(GPGGY)(GPGGS)]₂ motif is able to adopt the β -turns structures proposed earlier for this motif [6] and may be stabilized by intra-molecular hydrogen bonds involving tyrosine (Y) and serine (S). We hypothesize that by achieving a better organization or folding of this latter motif, the protein might be able to better self organize, thus allowing the poly-alanine segments to interact with one another and lock into β -sheets by a nucleation process. This nucleation process may initiate the self-assembly of the Y1S8₂₀ molecules into a supra-molecular structure. The fact that we do get spontaneous self-assembly of the pure Y1S8₂₀ proteins that were

subjected to a heat-treatment prior to purification reinforces this argument (Fig. 3). Moreover, changing the hydrophilic 5th (Y) and 10th (S) residues and replacing them by hydrophobic residues such as alanine (A) in A1S8₂₀ deprives the protein of the ability to retain large amounts of stable turn and sheet structures even when tested at a higher concentration than Y1S8₂₀. This may be the reason shear forces are needed to force film formation from a liquid crystalline phase from the pure A1S8₂₀ fractions.

Regarding the native Flag protein in the flagelliform gland, by having both [GPGGY GPGGS] alternating with [GPGGA GPGGV/A], the protein may have parts of its elastic segments in a stable β -turn conformation, possibly under salt and pH control as suggested for other silk glands [32, 33], while the rest of the elastic components are unfolded, or more disordered, thus allowing the molecules to remain in a liquid crystalline phase avoiding premature aggregation in the gland. Although in the Flag protein there is no poly-alanine segments [6, 9, 10], the GGX repeats are known to promote self-assembly in proteins such as lamprin [34] and might play the same role as the poly-alanine segments, or poly-(glycine/alanine) in other silks, in initiating the self-assembly of the Flag molecules. A film

similar to the ones seen in the A1S8₂₀ and Y1S8₂₀ pure fractions may self-assemble in the spider's silk glands. Then, shear forces resulting from a combination of the spider's pull on the fiber and the extremely small size of the spinning duct the proteins travel through might be enough to induce the proper folding of the [GPGGA GPGGV] motifs. By locking all its [GPGGX]_n motifs into the proper secondary structures, the rest of the molecules would then be able to self organize and assemble into a fiber. Moreover in our case, by pulling the structure together, and having properly folded motifs, the (GPS)₂ linker sequence directly preceding the poly-alanine segments might be able to form stable intermolecular hydrogen bonds that would stabilize the supramolecular structure. In the native Flag protein, we can hypothesize that this stabilizing role of the supramolecular structure may be fulfilled by the spacer region, positioned next to the (GGX)₁₂ sequence, containing both highly hydrophobic and highly hydrophilic amino acids.

Recent molecular dynamic simulation studies on the (GVPGV)₇ elastin pentapeptide suggest that such a sequence has a higher propensity for PPII (poly-proline II) structures since they cannot stabilize any turns through hydrogen bonding [35]. It also retains a high backbone hydration level that would constrain these peptides in disordered conformations that are unable to exclude water when aggregated, thus providing elastomeric properties to the matrix formed [35]. Such a model might be applicable to the more hydrophobic elastic motifs found in flagelliform silk as well as the elastic motif of the A1S8₂₀ protein. Moreover in elastins, glycine-rich repeats deprived of

proline seem to promote self aggregation in an amyloidogenic fashion [35], thus reinforcing the argument stated above about the role of GGX sequences in the self-assembly of the native Flag protein.

Mechanical performances of the synthetic pulled and spun fibers

Using an extruder, we were able to spin fibers out of both the A1S8₂₀ and Y1S8₂₀ proteins that were resolubilized in 100% HFIP. However, we determined that adding 15% water to the dope prior to spinning dramatically improved the mechanical properties as well as the appearance of the fibers. This can be explained by the fact that although HFIP might promote intramolecular hydrogen bonding [36, 37], water was necessary to force the hydrophobic poly-alanine segments away from the water phase and into proper sheet structures, thus initiating self-assembly. The stress/strain curves for A1S8₂₀ and Y1S8₂₀, both pulled (P) and spun (S) are shown in Fig. 5 while fiber specifics (number of fibers used and diameter) and mechanical performance data (maximum stress, maximum extension, Young's modulus, and toughness) are shown in Table 1. Before tensile test, pictures of each type of fibers used in this study were taken, and the pictures of the fibers displaying the best extension, best maximum stress, and average maximum stress are shown in Fig. 6. According to our data, the pulled fibers, which had a much smaller diameter on average than the spun ones, outperformed the spun fibers in average maximum extension, maximum stress and toughness (Table 1

Fig. 5 Stress/strain curves of the synthetic fibers. P1-P12 = Pulled fibers 1 through 12; S1-S12 = Spun fibers 1 through 12. Note that the scales for both stress and strain differ from one graph to another

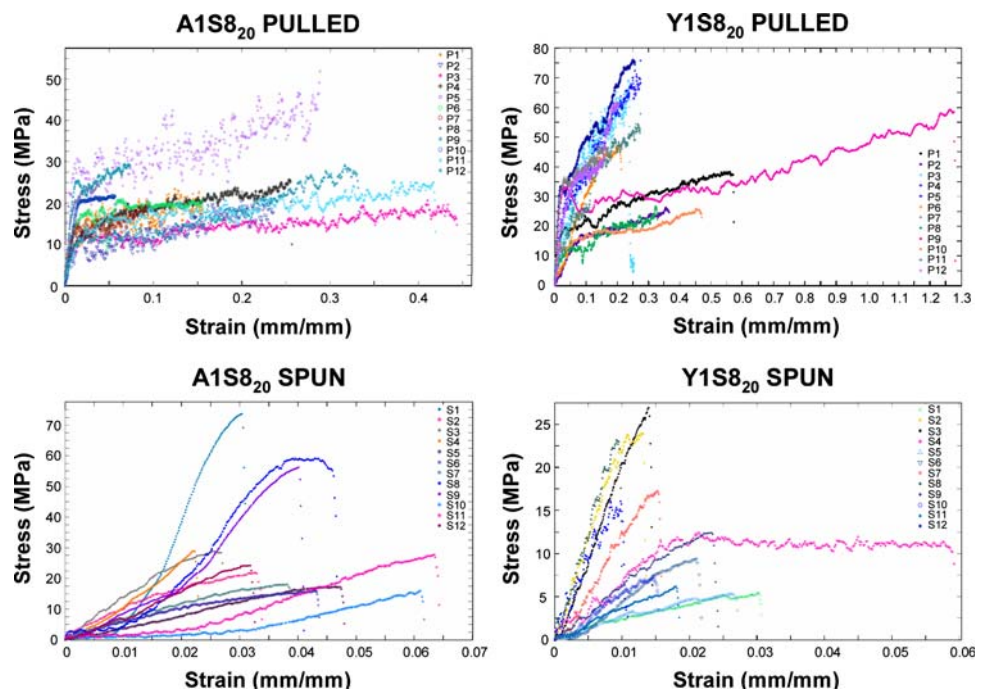


Table 1 Mechanical testing data

Fiber type	Total of fibers	Diameters (μm)	Young's modulus (MPa)	Maximum stress (MPa)	Maximum extension (%)	Toughness (MJ/m^3)
A1S8 ₂₀ P	15	12.20 \pm 4.99	1,706.8 \pm 791.87	28.64 \pm 8.41	18.99 \pm 12.88	3.41 \pm 2.61
A1S8 ₂₀ S	19	32.15 \pm 16.24	759.68 \pm 540.27	28.58 \pm 17.18	3.72 \pm 1.24	0.464 \pm 0.30
Y1S8 ₂₀ P	31	15.79 \pm 6.05	1,081.49 \pm 1,000	49.64 \pm 19.35	34.06 \pm 25.30	10.6 \pm 10.2
Y1S8 ₂₀ S	18	28.4 \pm 11.32	933.62 \pm 727.14	10.21 \pm 7.32	1.59 \pm 1.03	0.089 \pm 0.11

Average values measured for all types and kinds of synthetic fibers. The standard deviation of each value is indicated (\pm STD). P = pulled; S = spun

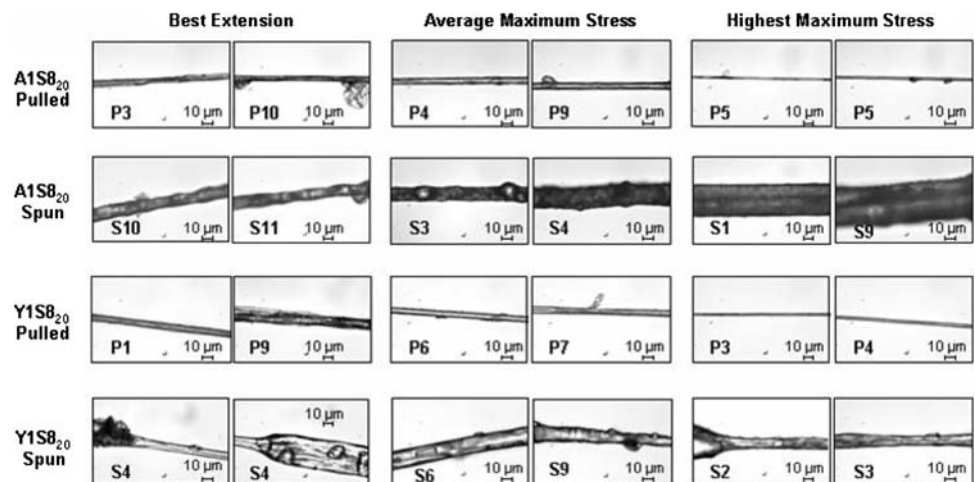
and Fig. 5). The appearance of these pulled fibers was also strikingly different than the spun ones (Fig. 6). Not only were their surfaces smoother, but they were also more even and had greater sheen than any of the spun fibers. This can be explained by the fact that the action of pulling the preassembled film into a fiber was probably close to post spin drawing that would improve the overall organization of the fiber, hence its performance.

Within the pulled group, the Y1S8₂₀ fibers were the toughest (average toughness = 10.6 MJ/m^3). They were also twice as extensible (average maximum extension = 34.06%), and could withstand almost twice as much maximum stress (average maximum stress = 49.64 MPa) as the A1S8₂₀ fibers which as a result displayed a higher stiffness (average Young's modulus = 1.706 GPa). This is understandable if we consider the secondary structures present in these Y1S8₂₀ proteins in an aqueous buffer at pH 8 (more β -sheets and more turns than for A1S8₂₀), and the fact that the elastic motifs in Y1S8₂₀ can be stabilized by hydrogen bonds due to the presence of Y and S in the elastic motif.

Of the two synthetic constructs studied here, the primary structure of the Y1S8₂₀ protein consensus repeat more closely resembles the native *N. clavipes* dragline MaSp2 consensus repeat ($[(\text{GPGQQ GPGGY})_2 \text{GPSGPS} (\text{A})_n]_n$)

[6], and it is interesting to note that the Y1S8₂₀ pulled fibers exhibit the same elasticity as the native dragline silk (34% vs. 35%), thus confirming the role of the MaSp 2 protein in the overall elasticity of the dragline. The Y1S8₂₀ pulled fibers, however, cannot withstand as much stress as the native dragline silk (0.050 GPa vs. 4 GPa), or flagelliform silk (0.050 GPa vs. 0.5 GPa). Consequently, their toughness (10.6 MJ/m^3 , Table 1) is lower than the ones reported for both dragline silk (160 MJ/m^3) and flagelliform silk (150 MJ/m^3) [3, 4]. We can speculate that this difference in stress threshold is the result of two major factors. The first one is that the synthetic proteins are much smaller than the native silk proteins forming the fiber thus limiting the number of intermolecular chain interactions necessary to stabilize the overall fiber structure and allowing weaker spots by lack of molecular overlap. Indeed, the sizes of both dragline silk proteins are between 300 kDa and 350 kDa [38], and although the size of Flag has not been determined, the fact that the Flag mRNA is bigger (15 kb) [6] than both MaSp mRNAs (11 kb and 12 kb) [39, 5] suggests that this protein is at least bigger than 350 kDa. Additionally, it is possible that a difference in sizes of the crystals formed and their lack of orientation affect the strength of the synthetic fiber. The second may be attributed to the fact that the native dragline contains a

Fig. 6 Pictures of pulled and spun A1S8₂₀ and Y1S8₂₀ fibers. For each fiber type, the pictures taken before tensile tests show the fibers that achieved the best extension, best maximum stress and average maximum stress. The fibers shown here correspond to the ones plotted in the stress/strain curves (Fig. 5): fiber type (P = pulled fiber; S = spun fiber) and identity (number) of each individual fiber is indicated in bold in the bottom left corner of each picture



second protein, MaSp1 [39], rich in GGX repeats and containing both (GA) and (A)_n crystalline forming motifs that can impart additional strength to the fiber.

The comparison between the structures and performances of both the Y1S8₂₀ pulled fibers and the dragline silk might give some insight as to the necessity of two proteins in the native dragline silk (MaSp 1 and MaSp2). The amount of (GPGXX)₄ motifs adjacent to crystalline forming motifs in the MaSp 2 consensus repeat, as well as their spacing, seem to be optimal for retaining reasonable elasticity (35%) while maintaining high strength. Therefore, a second protein is necessary for additional strength possibly because adding extra crystalline forming motifs to the MaSp2 consensus repeat might decrease the overall elasticity of fiber. Moreover, water is a known plasticizer of silks. It has been shown that native dragline silk extruded in water is stronger thus tougher than those extruded in air [40]. In the case of flagelliform silks, the presence of a natural aqueous glue coating is critical in promoting extreme elasticity (200%) [1]. It is possible that the mechanical performances of these ‘pulled’ fibers may be modified, even improved, when wet although no such mechanical data is available to support this statement.

Regarding the performances of the spun fibers (Fig. 5 and Table 1), we found, not surprisingly, that the A1S8₂₀ fibers on average outperformed the Y1S8₂₀ fibers for maximum stress (28.58 MPa vs. 10.21 MPa), maximum extension (3.72% vs. 1.59%), and toughness (0.464 MJ/m³ vs. 0.089 MJ/m³). In HFIP, we think that the elastic motifs of the Y1S8₂₀ protein may not be able to properly form stabilized β -turns. Indeed, in this solvent, having bulky residues present in the 5th and 10th positions (Y and S) of the elastic motif may impair its folding, while the A1S8₂₀ protein has the advantage of small residues (A) in these two positions in the elastic motif. This may better accommodate the proper folding of the molecule providing better overall organization allowing the poly-alanine to be locked together into sheets and promote self assembly. However, the extension of the spun A1S8₂₀ fibers is poor compared to the pulled A1S8₂₀ fibers reinforcing the idea that additional shear forces, such as post-spin draw, might be necessary for this protein to achieve better mechanical properties.

From our results, it appears that the Y1S8₂₀ protein may be more suitable to generate fibers with reasonable mechanical properties when using aqueous solutions while it is not a good candidate to generate fibers from organic solvents.

During these studies, we noticed a high variability in the performance of the pulled fibers as well as the spun ones within each type of fiber (Fig. 5). Such variability, although not as substantial, exists in native silk fibers [41, 42] as well as synthetic polymeric fibers. In our experiments, no discernible factors such as the use of different

protein batches to make the dopes, the ‘age’ of the dope, the diameters of the fibers for instance, could account for the distinct mechanical behaviors observed within a given fiber group. This variability is important to notice as it defines the upper and lower limits for the mechanical performance of any fiber. Refining these limits, that is lowering the variability in the mechanical performance of any synthetic fibers, should be a goal when making such synthetic fibers and may be achieved by mastering the spinning conditions. Hence, we want to stress the importance of repetitions when testing synthetic materials to know the full spectra of their performance.

Conclusion

We demonstrated a temperature dependent β -sheet induction in aqueous solution irreversible for Y1S8₂₀ and reversible for A1S8₂₀, possibly explaining the difference in spontaneous film/fiber formation observed for the two proteins in aqueous solutions.

The mechanical data obtained from both types of synthetic fibers made from the recombinant silk protein analogs (A1S8₂₀ and Y1S8₂₀) clearly demonstrates that the [GPGGX]_n motif is indeed responsible for the elasticity displayed by both native dragline and flagelliform silks. Moreover, the nature of the [GPGGX]_n motif, that is the hydrophobicity or hydrophilicity of the 5th and 10th amino acid residues, has a large impact in the level of elasticity probably due to stability issues (presence or absence of internal hydrogen bonding in the motif) as well as in the self organization process of the proteins into films or fibers.

Acknowledgements We would like to thank Dr Michael B Hinman (Department of Molecular Biology, University of Wyoming) for providing the original A1, Y1 and S8 pBluescript clones, and for performing the amino acid analyses of the recombinant proteins. We also want to thank Professor Michael S Ellison (School of Materials Science and Engineering, Clemson University) for his valuable advice on the mechanical testing methods and analyses. This research was supported by grants from NIH, NSF and AFOSR.

References

- Gosline J, Denny M, DeMont M (1984) *Nature* 309:551
- Stauffer S, Coguill S, Lewis R (1994) *J Arachnol* 22:5
- Gosline J, Guerette P, Ortlepp C, Savage K (1999) *J Exp Biol* 202:3295
- Denny M (1976) *J Exp Biol* 65:483
- Hinman M, Lewis R (1992) *J Biol Chem* 267:19320
- Hayashi C, Shipley N, Lewis R (1999) *Int J Biol Macromol* 24:271
- Simmons A, Michal C, Jelinski L (1996) *Science* 271:84
- Termonia Y (1994) *Macromolecules* 27:7378
- Hayashi C, Lewis R (1998) *Mol Biol* 275:773
- Hayashi C, Lewis R (2000) *Science* 287:1477

11. Becker N, Oroudjev E, Mutz S, Cleveland J, Hansma P, Hayashi C, Makarov D, Hansma H (2003) *Nat Mater* 2:278
12. Hayashi C, Lewis R (2001) *BioEssays* 23:750
13. Zhou Y, Wu S, Conticello V (2001) *Biomacromolecules* 2:111
14. Chang D, Ventachalam C, Prasad K, Urry D (1989) *J Biomol Struct Dynamics* 6(i5):851
15. Urry D, Shaw R, Prasad K (1985) *Biochem Biophys Res Commun* 130(i1):50
16. Urry D, Hugel T, Seitz M, Gaub H, Sheiba L, Dea J, Xu J, Parker T (2002) *Phylos Trans R Soc London, Ser B* 357(i1418):169
17. Tatham A, Hayes L, Shewry P, Urry D (2001) *Biochim Biophys Acta* 1548:187
18. Hutchinson E, Thornton J (1994) *Protein Sci* 3:2207
19. Wilmot C, Thornton J (1988) *J Mol Biol* 5:221
20. Lewis R, Hinman M, Kothakota S, Fournier M (1996) *Protein Expr Purif* 7:400
21. Birnboim H, Doly J (1979) *Nucleic Acids Res* 7:1513
22. Sambrook J, Fritsch E, Maniatis T (1989) *Molecular cloning. A laboratory manual. Second Edn. Cold Spring Harbor Laboratory* 28:426
23. Wong C, Sridhara S, Bardwell J, Jakob U (2000) *BioTechniques* 28:426
24. Whithmore L, Wallace B (2004) *Nucleic Acids Res, Web Server issue* 24:W668
25. Scheller J, Guhrs K, Grosse F, Conrad U (2001) *Nature* 19:573
26. Teulé F, Jung S, Wood J, Marcotte W, Ellison M, Abbott A (2002) In: Brebbia C, Sucharov L, Pascolo P (eds) *Design and nature*, Udine, Italy. WIT press, p 379
27. Prince J, McGrath K, DiGirolamo C, Kaplan D (1995) *Biochemistry* 34:10879
28. Fahnestock S, Irvin S (1997) *Appl Microbiol Biotechnol* 47:23
29. Fahnestock S, Bedzyk L (1997) *Appl Microbiol Biotechnol* 47:33
30. Arcidiacono S, Mello C, Kaplan D, Cheley S, Bayley H (1998) *Appl Microbiol Biotechnol* 49:31
31. Dicko C, Knight D, Kenney J, Vollrath F (2004) *Biomacromolecules* 5:2105
32. Vollrath F, Knight D (2001) *Nature* 410:541
33. Knight D, Vollrath F (2001) *Naturwissenschaften* 88:179
34. Robson P, Wright G, Sitarz E, Maiti A, Rawa M, Youson J, Keeley F (1993) *J Biol Chem* 268(2):1440
35. Rauscher S, Baud S, Miao M, Keeley F, Pomès R (2006) *Structure* 14:1667
36. Creighton T (1993) *Proteins: structures and molecular properties, Second Edn. Freeman, New York*
37. Haq S, Khan R (2005) *Int J Biol Biomacromolecules* 36:47
38. Sponner A, Schlott B, Vollrath F, Unger E, Grosse F, Weisshart K (2005) *Biochemistry* 44:4727
39. Xu M, Lewis R (1990) *Proc Natl Acad Sci USA* 87:7120
40. Chen X, Shao Z, Vollrath F (2006) *Soft Matter* 2:448
41. Blackledge T, Hayashi C (2006) *J Exp Biol* 209:2452
42. Blackledge T, Hayashi C (2006) *J Exp Biol* 209:3131