



# NIH Public Access

## Author Manuscript

*Biomacromolecules*. Author manuscript; available in PMC 2014 June 10.

Published in final edited form as:

*Biomacromolecules*. 2013 June 10; 14(6): 1751–1760. doi:10.1021/bm400125w.

## ***Nephila clavipes* Flagelliform Silk-like GGX Motifs Contribute to Extensibility and Spacer Motifs Contribute to Strength in Synthetic Spider Silk Fibers**

**Sherry L. Adrianos<sup>\*,1</sup>, Florence Teulé<sup>2</sup>, Michael B. Hinman<sup>3</sup>, Justin A. Jones<sup>3</sup>, Warner S. Weber<sup>4</sup>, Jeffery L. Yarger<sup>4,5</sup>, and Randolph V. Lewis<sup>3</sup>**

<sup>1</sup>Department of Molecular Biology, University of Wyoming, Laramie, Wyoming 82071

<sup>2</sup>Department of Zoology and Physiology, University of Wyoming/Casper College Center, Casper, Wyoming 82601

<sup>3</sup>Department of Biology, Synthetic Biomanufacturing Center, Utah State University, Logan, Utah 84422-5305

<sup>4</sup>Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 83287-1604

<sup>5</sup>X-ray Science Division, Advanced Photon Source, Argonne National Laboratories, Argonne, Illinois, 60439

### **Abstract**

Flagelliform spider silk is the most extensible silk fiber produced by orb weaver spiders, though not as strong as the dragline silk of the spider. The motifs found in the core of the *Nephila clavipes* flagelliform Flag protein are: GGX, spacer, and GPGGX. Flag does not contain the polyalanine motif known to provide the strength of dragline silk. To investigate the source of flagelliform fiber strength, four recombinant proteins were produced containing variations of the three core motifs of the *Nephila clavipes* flagelliform Flag protein that produces this type of fiber. The as-spun fibers were processed in 80% aqueous isopropanol using a standardized process for all four fiber types, which produced improved mechanical properties. Mechanical testing of the recombinant proteins determined that the GGX motif contributes extensibility and the spacer motif contributes strength to the recombinant fibers. Recombinant protein fibers containing the spacer motif were stronger than the proteins constructed without the spacer that contained only the GGX motif or the combination of the GGX and GPGGX motifs. The mechanical and structural X-ray diffraction analysis of the recombinant fibers provide data that suggests a functional role of the spacer motif that produces tensile strength though the spacer motif is not clearly defined structurally. These results indicate that the spacer is likely a primary contributor of strength with the GGX motif supplying mobility to the protein network of native *N. clavipes* flagelliform silk fibers.

### **Keywords**

Flagelliform; spacer; Flag; spider silk; recombinant protein; X-ray diffraction

---

\*Corresponding author: Sherry L. Adrianos; 7SherryA@gmail.com.

**CONFLICTS:**

The authors have no known conflicts.

**SUPPORTING INFORMATION AVAILABLE:**

The p-values for the statistical analysis of the fiber group comparison for tensile strength and % extensibility for Table 2 (Table S1). This information is available free of charge via the Internet at <http://pubs.acs.org/>.

## INTRODUCTION

Native flagelliform spider silk is a tough proteinaceous fiber produced by orb weaver spiders to make the capture spiral of the orb web<sup>1, 2</sup>. Orb weavers can produce 6 different types of silk fibers, each originating from different silk-specific glands with different mechanical and biochemical properties<sup>3–6</sup>. Amino acid sequences of various silks of different species of orb weaver spiders revealed repetitive elastic motifs and modules<sup>7–9</sup>. The consensus sequence motifs are similar between species however actual amino acid sequences differ. Our understanding of spider silk protein is limited, yet there is much to be learned<sup>10</sup>. The flagelliform silk sequences from *Nephila clavipes* were used in our study.

The strength and remarkable extensibility of the flagelliform fiber produces mechanical properties tougher than most natural and manmade materials, three times tougher than Kevlar<sup>4, 5</sup>. Much of what is known about spider silk comes from the studies of major ampullate (dragline) silk, the strongest spider silk<sup>4, 11–19</sup>. However, unlike dragline silk, flagelliform silk is difficult to obtain in appreciable quantities for fiber analysis for two reasons. First, flagelliform silk is one fourth the diameter of dragline silk and second, it cannot be easily collected by pulling the fiber from the spinnerets located on the abdomen of the spider. Therefore recombinant protein techniques provide tools that will further our understanding of the motif structure and functional contributions to the native flagelliform silk protein fibers.

The prey-capturing adhesiveness of flagelliform silk originates from the application of an aqueous glycoprotein “glue” coating to the core flagelliform silk fiber. The flagelliform core fiber and the applied sticky coating are two separate proteins produced in different glands and synthesized individually (Fig 1). Hygroscopic components in the glycoprotein glue attract water from the environment. The water keeps the glycoprotein and silk fibers of the web hydrated<sup>20, 21</sup>. Hydration of native flagelliform silk by the aqueous glue has been shown to supplement extensibility<sup>22, 23</sup>. Water naturally plasticizes silk proteins and thus plays a critical role in the mechanical properties of native silk fibers<sup>3, 24</sup>. Regenerated spider silk and chimeric spider silk-like fibers displayed improved mechanical properties when processed in aqueous alcohol solutions that plasticized the amorphous regions and promoted secondary structural transitions that reorganized the fiber network<sup>25–27</sup>. Aside from the role of water, the strength and extensibility of spider silks has been correlated with the primary sequences of the proteins of the various spider silk fibers<sup>7</sup>.

Flagelliform silk is composed of a single very large repetitive protein, Flag. The *N. clavipes* Flag protein is approximately 360 kDa in size with a glycine rich protein core that forms 90% of the protein<sup>7, 8</sup>. Three motifs dominate the Flag protein: GPGGX, GGX (X=A, S, V, and Y in GPGGX or T in GGX), and a glycine poor “spacer” (Fig 2). Iterations of the GPGGX motif, (GPGGX)<sub>n</sub>, provide elasticity and extensibility produced from type II β-turns<sup>29–31</sup> generating β-turn nanospring structures<sup>18, 19, 32, 33</sup>. The β-turn generated structures are similar to those found in the elastic proteins elastin and gluten<sup>29, 30, 34</sup>. The (GPGGX)<sub>43–62</sub> extensibility regions comprise 75% of the total Flag protein and are linked to a non-iterated spacer by a comparably smaller string of the iterated GGX motifs (GGX)<sub>7–12</sub> (Fig. 2). The negatively charged spacer motif, composing collectively 7.5% of the protein core in near equal proportion to the (GGX)<sub>7–12</sub> motifs, interrupts the long glycine rich arrays. Each Flag protein contains up to fourteen repeats of the [(GPGGX)<sub>n</sub> (GGX)<sub>n</sub> spacer]<sub>n</sub> ensemble<sup>7, 8</sup>.

The conserved non-repetitive flanking C- and N- termini regions of dragline and flagelliform proteins were suggested to control solubility and fiber formation in native

spider silk<sup>35</sup> and also in recombinant silk proteins in solution<sup>36–38</sup>. Recombinant tubuliform silk fibers that contained the C-terminal domain had increased extensibility over fibers that did not contain the terminal domain<sup>39</sup>. Several recombinant spider silk-like proteins have been shown to form fibers without the inclusion of the terminal sequences<sup>27, 40–43</sup>. Materials with varying degrees of re-solubility were produced from “tunable” silk-elastin protein polymers that were constructed with variations of silk and elastin sequences<sup>44</sup>.

Flagelliform silk exhibits six times the elasticity and one fourth the strength of dragline silk. The strength of dragline silk is attributed to the polyalanine “strength” motifs, (A)<sub>n</sub>/(GA)<sub>n</sub>, that form stacked crystalline β-sheet structures that align parallel to the axis of the fiber<sup>15, 45–49</sup>. However, the Flag protein does not contain this typical polyalanine strength motif found in other silks. Raman studies suggest β-sheet structure attributed to the spacer motif confers strength to the flagelliform fibers<sup>50</sup>. The extensibility of flagelliform silk is associated with other elastic proteins such as the resilin protein that was found to be an amorphous random network that may consist of a sliding β-turn model structure<sup>51</sup>. Understanding the extensibility of the Flag protein is important, yet the source of its strength has remained a mystery.

In this study, four *N. clavipes* Flag-like protein variants were designed to investigate the motif contributions to mechanical properties of the spacer motif of the flagelliform Flag protein. Iterations of the GGX motif were designed and engineered into all four encoding DNA constructs to create the GGX series of Flag-like sequences. The Flag-spacer and/or iterations of the (GPGGX)<sub>8</sub> motif module were designed into three of the encoding (GGX)<sub>7</sub> module containing constructs and produced in *Escherichia coli*. The purified recombinant silk proteins generated liquid spin dopes for fiber spinning. A standard spinning method was used to generate all fibers. The processed fibers were stretched in aqueous alcohol and air dried. The fibers were tested mechanically and the resulting fiber data was compared to each other. X-ray diffraction analysis was used to investigate structure. Based on the Flag-like motif content and function in the recombinant proteins, the data were correlated with the similar motif content of the native Flag protein. These studies help demonstrate the motif contributions of the native Flag protein fibers through the use of the recombinant Flag-like protein fibers.

## EXPERIMENTAL SECTION

### DNA Sequence design

Four double stranded (ds) DNA sequences were designed encoding four different proteins that each contained a Flag-like (GGX)<sub>7</sub> structural motif module as the constant motif component designated as “G” in the new proteins (Fig.3). The Flag-spacer is designated as “F” and the (GPGGX)<sub>8</sub> module as “Y”. Three dsDNA sequences were engineered encoding separately the (GGX)<sub>7</sub> 5'-

actagtggatccgcataatggcccccgggtggcgctggcgtagcggtagggcgccgggtggttctgggtgtggcggtctggcgca  
ccggatccggagatatctggatccaaactg 3', the [(GGX)<sub>7</sub> spacer] 5'-

actagtggatccccatatggcccccggagggtgcgggggaagtgggggagcaggagggtcagggtgggtccggcggcagccgg  
ggcaccaccatcattgaggatctggacatcacaatcgacggcgcggacggccctatcacaattagcgaagaactcactattagcgg  
agctgggggtccggcggatctggatccaagctt 3', and the (GPGGX)<sub>8</sub><sup>41</sup> motifs. All dsDNA sequences were optimized for an *E. coli* codon usage bias. The G and GF sequences were synthesized and cloned directionally in the pBluescript®II SK+ plasmid (Stratagene) between the *SpeI* and *HindIII* restriction enzyme sites, with the introduction of new flanking *BamHI* and *NdeI*, restriction sites by Bio S&T Inc. (Fig.4; Montreal, Quebec). The new G and GF plasmid clones encoded the motif sequences: GGAGGSGGAGGSGGVGGSGGT for ‘G’ and TIIEDLDITIDGADGPITISEELTISGAGGS for ‘F’ (Fig. 4). The three dsDNA

sequences, G, GF, and Y, were used to build the constructs for the production of the new proteins.

The Y plasmid clone encoding the (GPGGSGPGGY)<sub>4</sub> sequence was produced from the same core sequence that was designed and supplied as the “Y2” clone previously reported<sup>27, 41, 52</sup>. The Y2 dsDNA core sequence was released from the pBluescript SK+ plasmid utilizing the 5' *Xma*I and the 3' *Bsp*EI restriction sites and then inserted into a BIO S&T modified pBluescript plasmid vector that had been prepared by removing the G sequence at the same restriction sites. This method generated the dsDNA sequences of the Y clone in pBluescript SK+ with the additional restriction sites that had been designed into the G and GF sequences thus preparing the Y clone sequence for compatible manipulation with the G and GF clone sequences.

The structural motif combinations that form the module basis for each of the new proteins (Fig. 5) were constructed through the manipulation of their encoding dsDNA. The restriction enzymes *Bsp*EI or *Xma*I, in conjunction with *Alw*NI, were used utilizing an established compatible non-regenerable restriction site method<sup>41, 43, 52, 55, 56</sup>. The DNA sequences were manipulated to assemble the G, GY, GF, and GFY encoding modules of the basic repeat ensembles. These modules were then iterated 32×, 12×, 12×, and 8× to get final DNA constructs with sizes of 2.30, 2.37, 2.05, and 2.37 kbp respectively, using the compatible non-regenerable method<sup>55, 56</sup>. All intermediate and final plasmid constructs were maintained in *E. coli* strain XL1 Blue (Stratagene). Recombinant plasmid DNAs were analyzed by enzyme restriction digestion followed by agarose gel electrophoresis. The DNA sequence of each silk insert was verified by standard Sanger sequencing methods (Nucleic Acid Exploration Facility, University of Wyoming, WY), using the T3 and T7 vector specific primers (Stratagene).

The common feature for all four silk protein variants was the presence of the (GGX)<sub>7</sub> domain or the G module, in their primary sequences. The role of the selected Y motif has been previously confirmed to contribute extensibility to synthetic fibers made with proteins containing a Y motif half the size of the one used here<sup>57</sup>. Thus, all of these newly engineered Flag-like protein variants were specifically designed to address the functional role of both the (GGX)<sub>7</sub> and the spacer structural domains (*i.e.* the G and F modules) found in the native silk protein.

### Recombinant Silk Gene Expression and Protein Purification

The four newly constructed sequences were separately inserted into the plasmid vector pET19b-k (kan<sup>R</sup>) at the *Nde*I/*Bam*HI restriction sites and cloned into *E. coli* strain BL21DE3 cells (Stratagene). This pET19b-k vector is a re-engineered pET 19b plasmid vector from Novagen that was previously modified to provide kanamycin resistance (see<sup>41</sup> for details on the construction of this vector).

The proteins were expressed, harvested, and purified using methods previously reported<sup>43, 52</sup>. All recovered cell extracts were heat treated to remove most native *E. coli* proteins<sup>41, 58</sup>. The His-tagged proteins were then purified using automated immobilized metal ion affinity chromatography (IMAC) techniques performed on an AKTA™ Avant 150 instrument (GE Healthcare, Piscataway, NJ). The unbound proteins were removed using successive wash buffers containing imidazole concentrations of 20 mM and 60 mM or 80 mM. The proteins were recovered with an elution buffer containing 150 mM imidazole. All separate eluted protein fractions were dialyzed against water containing 5 mM ammonium bicarbonate using standard methods, frozen, and lyophilized<sup>59</sup>. The dry recombinant silk (rSilk) proteins were solubilized for spin dopes and characterized by SDS-PAGE and

Western blot analyses using a 6× His mAb-HRP conjugate (Clontech) anti-His tag antibody as described previously<sup>41</sup>.

### Preparation of the Spin Dopes

Fifteen % (w/v) rSilk-HFIP protein spin dopes were generated by dissolving each of the purified lyophilized proteins into 1,1,1,3,3,3-hexafluoroisopropanol (HFIP; TCI America, Portland, OR) to prepare a liquid spin dope. Refinement of fiber formation and spinning conditions for this group of proteins was obtained through spin testing from HFIP utilizing a toluene solvent addition to the rSilk-HFIP dope following previous trials<sup>60</sup>. The addition of 5% (v/v) toluene to the prepared 15% (w/v) rSilk-HFIP spin dope was determined to be optimal for fiber spinning for this group of proteins. To allow for complete protein solubilization, the prepared spin dopes were vortexed and then placed on a rocking platform (Adams Nutator 1105, Parsippany, NJ) for 2–4 days. The viscous spin dopes were then filtered using U-prep spin columns (Genesee Scientific, San Diego, CA) and centrifuged at 10K g for 2 minutes prior to fiber spinning. All spin dopes made from the different proteins (G, GY, GF, and GFY) were prepared separately and processed in the same standardized manner to eliminate dope preparation variables.

### Fiber Spinning

The synthetic fibers were spun as previously reported<sup>41, 43, 52, 61</sup> on a DACA Instruments SpinLine system (Santa Barbara, CA) using a 1mL Hamilton Gastight syringe extruder (Hamilton, Reno, NV) equipped with a 0.005" inner diameter PEEK tubing (Upchurch Scientific) custom made needle. Initial spinnability testing of the GY spin dope was investigated to obtain standardized spinning and coagulation conditions under which all four recombinant proteins G, GY, GF, and GFY could produce fibers. Methods of fiber extrusion into an isopropanol (IPA) coagulation bath containing water, like that reported earlier were investigated<sup>27, 41, 52</sup>. However, the addition of 10%, or 5%, or as low as 2% (v/v) water to the coagulation bath caused the extruded forming fibers to disintegrate. Therefore the spin dopes were coagulated directly into a 100% IPA coagulation bath as previously reported<sup>43, 62, 63</sup>.

During fiber spinning, a glass submersion rod was used to guide the nascent fibers through the coagulation bath and to prevent fiber flotation and entanglement at the surface of the bath. The newly formed fibers were too fragile to be stretched directly using the godets on the DACA spinneret. The fibers emerging from the coagulation bath were collected directly onto a spool without slack or tension on the fiber. The as-spun fibers were removed from the spool and dried overnight without physical restraint at room temperature, before being further processed.

### Fiber Post-Spin Drawing

The as-spun (AS) fibers were individually processed to generate the post-spin drawn (PSD) fibers by submersion in an 80% (v/v) aqueous isopropanol bath for 0.1–1 minute until malleable. The immersed fibers were then stretched from both ends with forceps to 3 times their original lengths in the isopropanol bath. All AS fibers were brittle and resisted stretching with lower water concentrations. Water concentrations greater than 20% in the draw bath caused the fibers to dissolve. The processed fibers were removed from the bath and held constrained in the air at the stretched length for 20 seconds to dry. The fibers were further dried unconstrained overnight at room temperature.

## Mechanical Testing and Analysis

Both as-spun (AS) and stretched (PSD) fibers were cut and mounted on testing cards with an initial uninhibited fiber length ( $L_0$ ) of 19.05 mm for mechanical testing. The fibers were visualized under a Nikon Eclipse E200 microscope (Nikon Metrology Inc., Brighton, MI) at a 40 $\times$  magnification. A total of 9 diameter measurements along the axis of each fiber were taken from 3 different images utilizing Image J 1.42 software (National Institute of Health, USA). These measurements were used to determine the average fiber diameters and to calculate the cross sectional area of each individual fiber. The average diameters and cross sectional areas were used in conjunction with the stress/strain data to compile the mechanical properties of the new silk fibers.

The fibers were tested until failure at ambient conditions (25° C and 19–22% relative humidity) on an MTS Synergie 100 system (MTS Systems Corporation, Eden Prairie, MN) using a custom made 10 gram load cell (Transducer Techniques, Temecula, CA) at a constant strain rate of 2 mm/min. The mechanical data for ten fibers per group were recorded at a frequency of 30 Hz utilizing the TestWorks4 software (Software Research Inc., San Francisco, CA). The strain and load data were used to calculate engineering stress ( $\sigma_e$ ) and strain ( $\epsilon_e$ ) using the following formulas:  $\sigma_e=F/A$  and  $\epsilon_e=L-L_0/L_0$ ; where, F=load, A=cross-sectional area of the fiber; L=length and  $L_0$ =original fiber length.

Stress/strain curves were plotted and utilized to determine additional parameters such as toughness and % strain ( $\epsilon_e \times 100$ ) also called percent (%) extension. Fiber toughness corresponds to the area under the stress/strain curve and was calculated using the standard trapezoidal method. The mean data with standard deviations were calculated for the mechanical data (Table 1). To compare the mechanical properties of the different groups of fibers (G, GY, GF, and GFY), an unpaired t-test was performed between all PSD fiber groups. Statistically significant differences ( $p<0.05$ ) in tensile strength and extensibility were calculated using a confidence level of  $\alpha = 0.05$ .

## Structural Analysis: X-ray Diffraction

Wide angle X-ray scattering (WAXS) experiments were performed in sector 14 BM-C/BioCARS of the Advanced Photon Source (APS) at Argonne National Laboratory (Argonne, IL, USA)<sup>64</sup>. The WAXS data were recorded with a large area 9-chip ccd detector (ADSC Quatum-315) placed 33 mm behind the sample. The monochromatic X-ray beam was focused to 150  $\times$  200 microns at an incident energy of 12.67 keV and with an X-ray wavelength of 0.978 Å. The bundled fiber samples were placed vertically, perpendicular to the X-ray beam, in the same geometry as the beam stop. For each sample, 5 frames were collected with a 50 mm beamstop to detector distance and an exposure time of 60 seconds. The background (air scattering) was subtracted from all X-ray intensities shown.

## RESULTS AND DISCUSSION

### Recombinant Proteins

This series of four Flag-like protein variants were designed to investigate structure/function relationships in the flagelliform spider silk protein, Flag of *N. clavipes*. Although the nature of the elasticity of this silk is attributed to the presence and abundance of the (GPGGX)<sub>n</sub> motif, the source of its high tensile strength as well as its mode of self-assembly are largely unexplored. The genetically engineered G, GY, GF, and GFY Flaglike repetitive DNA sequences were successfully cloned into *E. coli* producing the corresponding His-tagged proteins with expected molecular weights ranging from 54–66 kDa (Table 1). The protein production levels varied between the various clone types: G 50 µg/L; GY 500 µg/L; GF 5 µg/L; and GFY 50 µg/L. The spacer containing GF and GFY proteins were only observed as

dimers with Western blot analysis, while the GY protein produced predominantly monomers with minimal truncated product. Occasionally monomers were observed for the G protein but it primarily produced dimers.

The primary structure of these proteins included repeating modules containing all (*i.e.* GFY) or some (*i.e.* G, GY and GF) of the following native Flag structural motifs (Fig. 2): the G module or (GGX)<sub>7</sub>, the F module or ‘spacer’, and the Y module which is one of the putative Flag elastic (GPGGX)<sub>8</sub> motifs, (GPGGY GPGGS)<sub>4</sub><sup>7</sup>. In this study, only the roles of structural domains present in the large highly repetitive core of the native Flag protein were investigated, as this region constitutes the majority (90%) of the primary sequence of this native silk protein. As such, this repetitive core is thought to be the main contributor to the mechanical properties of the fibers<sup>7,8</sup>. The pH gradient of the spider silk gland was previously shown to affect or promote fiber formation<sup>38,65,66</sup>. Structural work on the non-repetitive flanking regions has suggested an *in vivo* aggregation role in fiber formation<sup>35</sup>. Recently, it was reported that the flanking regions of recombinant proteins in solution drive fiber formation<sup>36,37</sup>. However, many studies have shown that these flanking regions are not necessary in the recombinant silk proteins to allow fiber formation<sup>27,40,41,43,52,67</sup>. For the sake of targeted structural motif study, the N-and C-terminal ends were not included in our silk protein sequences design.

### Synthetic Silk Fibers

Each of the four protein variants was successfully spun into synthetic fibers. All newly as-spun (AS) fibers were brittle, therefore post-spin draw processing was necessary to produce fibers with improved mechanical properties (Table 2)<sup>40</sup>. No mechanical data was collected for the G and GY AS fibers as these were too brittle to test. The AS fibers averaged 27.24 to 36.81 microns in diameter. After the G, GY, GF, and GFY fibers were spun and dried, stretching the fibers in aqueous isopropanol (IPA) containing no more than 20% (v/v) water was the best processing treatment for all four different fiber types (Fig. 6). These fibers would dissolve completely in aqueous IPA with higher than 20% water content. The average fiber diameter of the PSD fibers ranged from 13.31 to 21.46 microns (Table 2).

The wetted AS fibers became malleable when submersed in 80% IPA and were easily processed. However, all four synthetic fiber types were still soluble in water after PSD processing, unlike what has been reported for both native regenerated silks and synthetic Flag/MaSp 2 silk fibers<sup>26,27</sup>. The solubility of the processed fibers suggests that the proteins in the PSD fibers may not have yet reached the structural transitions and optimal molecular alignment found in the native fibers.

The resulting synthetic fibers clearly indicate that under these artificial spinning conditions, the structural domains present in the repetitive core of the Flag protein are able to drive recombinant silk protein aggregation. Fiber formation either is driven by the domains independently, as seen for the G protein with only the G module or synergistically for proteins that contain both the F and G modules. These results do not exclude the potential additional involvement of the conserved N- and/or C-termini regions in the *in vivo* mechanism of silk protein regulation, aggregation, and fiber formation<sup>36–38</sup>. However, the investigation of these conserved regions was not the focus of our study. The results of our fiber spinning investigation are in line with those published on fibers produced from MaSp-like or Flag/MaSp2 chimeric proteins which also focused on the structural study of the repetitive silk core<sup>27,41,43,63</sup>.

## Mechanical Properties of the Fibers

Variability in the mechanical data within each fiber group is noticeable judging by the standard deviations of all mean values (Table 2). This is not unexpected as similar observations were reported for the mechanical performance of native<sup>68</sup> and synthetic silk fibers<sup>27, 43</sup>. Little to no variation stems from the spinning process, as these conditions were constant for all fiber groups. Fiber heterogeneity and manual processing of the AS fibers may be the most likely identifiable contributing factors of data variation as previously suggested<sup>27</sup>. This variation generated by processing is further supported by recent work showing that protein processing influences native spider silk performance<sup>69</sup>.

Analyses of the mechanical data show that statistically significant differences in mechanical properties exist between the synthetic PSD fibers. Statistical analyses of the tensile strength data divided the four fiber types into two groups, the weaker GY and G fibers (*p*-value=0.18), and the stronger GFY and GF fibers (*p*-value=0.26) strongly suggesting that strength is derived from the spacer. Additionally, the extensibility for the PSD fibers were significantly different except between the GY and GF PSD fibers (*p*-value=0.28; Table 2 & Supplemental Table 1). To better appreciate these differences, the stress/strain performances of the strongest or the most extensible PSD fibers for each variant protein fiber type were compared (Fig. 7). The data suggest that the differences in mechanical properties observed are associated with the variation of the structural motif content within the G, GY, GF, and GFY proteins (Table 2 & Fig. 7) thus corresponding with variations in the silk protein primary structure.

Interestingly, the presence of the (GGX)<sub>7</sub> motif module provides superior extensibility, over two-fold, as shown best by the differences in extensibility between the top performing G PSD fibers and both the GY and GF PSD fibers (Fig. 7). In spite of its low tensile strength, the G PSD fibers display enhanced extensibility, which is near that of native flagelliform fibers and makes them the second toughest synthetic fibers behind the GFY PSD fibers (Table 2). Furthermore, the addition of the (GPGGX)<sub>8</sub> motif in the GY proteins reduces the extensibility of the PSD GY fibers three-fold, compared to that of the G PSD fibers, though it does not significantly affect their tensile strength (Fig. 7, Table 2). The GY protein, with its  $\beta$ -turn/ $\beta$ -spiral forming (GPGGX)<sub>8</sub> motif was actually expected to produce a more extensible PSD GY fiber than PSD fibers made from the G protein. However, these GY PSD fibers were among the least extensible fibers and had the lowest toughness of all. This indicates that an integral part of the protein function is missing in the GY fibers, limiting what could have been superior extensibility from the high percentage of (GPGGX)<sub>8</sub> content comparatively.

The dual role of the spacer in the recombinant protein is apparent in the increased tensile strength and moderately restricted extensibility of the fibers. This is best demonstrated by the higher tensile strength achieved by the GF and GFY PSD fibers (up to three-fold) compared to those recorded for the G and GY PSD fibers (Table 2 & Fig. 7). The trade-off for these stronger GF and GFY PSD fibers seems to be a pronounced reduction in fiber extensibility suggestive of intermolecular interactions attributed to the spacer motif. The GFY PSD fibers achieved the second greatest extensibility. This result reinforces the idea that the key role played by the spacer (F) promotes or stabilizes the proper secondary structures in the two adjacent modules in the GFY protein, producing the toughest PSD fibers.

It is important to note that the toughest fibers were produced from the GFY protein, which had a primary sequence that most closely resembled that of the native Flag protein. The combinations of the three motif modules, G, F, and Y, in the GFY protein, may complete a more functional molecular unit. This organization may provide the complex, synergistic role

that concertedly provides optimal tensile strength and extensibility similar to that found in the native silk fibers. In the correct combination, the function of the synthetic molecules orchestrates the optimal fiber, yet independently, or in combinations such that when one motif element is missing, the mechanical properties of the fibers are significantly affected.

The largely amorphous structure of flagelliform silk may be partly explained by the elasticity mechanism for elastomeric protein polymers such as resilin<sup>70</sup>. The random elastomeric model or the sliding  $\beta$ -turn model describing the random network structure of recombinant resilins may be useful in better understanding the Flag protein in the future<sup>51</sup>. While it appears that the Flag spacer induces increased order of the overall protein structure in the fiber based on the mechanical property contributions, further investigations are needed to better understand the structure and function relationship of flagelliform silk.

### Structural Analysis of the Synthetic Fibers

The WAXS patterns of the GY, GF and GFY fibers are all characterized by the presence of a diffuse ring centered at 4.4 Å (Fig. 8). These isotropic patterns are consistent with the presence of an amorphous phase composed of largely disoriented structures such as helices or of largely disoriented  $\beta$ -sheet nanocrystals formed upon extrusion as indicated previously in the case of regenerated dragline silk and synthetic AS silk-like fibers<sup>26, 27</sup>.

For both the G and the GY AS sample, the formation of imperfect polyglycine-type nanosheets (PGI) may contribute to this component as none of these proteins possess any other classical  $\beta$ -sheet motifs in their sequences. The formation of  $\beta$ -sheet in flagelliform silk was reported in quantities high enough to crosslink the molecular network but not reinforce the network<sup>71</sup>. Other glycine-rich spider silk peptides formed  $\beta$ -sheets when cast into films<sup>72</sup>. These data are interesting but do not fully elucidate the protein structures.

In the WAXS patterns of only the GF and the GFY fibers, an additional reflection at 10 Å is present (Fig. 8). The reflection must result from the presence of the spacer in these proteins. These specific double reflection WAXS patterns are characteristic of stacked  $\beta$ -sheet type structures (4.7 Å and 8–11 Å) observed in amyloid fibers<sup>73, 74</sup>. Additionally, these stacked nanocrystalline structures are highly isotropic in the GF and GFY AS fibers, although their initial presence may ultimately set the stage for the higher tensile strengths achieved by their respective PSD fibers (no WAXS data).

It has just recently been reported that indeed, small quantities of  $\beta$ -sheet were observed by Raman spectra in native flagelliform spider silk fiber<sup>50</sup>. However, in that same study, the five-component deconvolution method that has been so widely used was changed by Lefèvre<sup>50</sup> and a different decomposition pattern containing only four components was developed specifically for flagelliform silk simplifying an appropriate and reasonable fit. The amide I deconvolution 1670 cm<sup>-1</sup> component observed in these fibers is attributed to  $\beta$ -sheets based in prior studies<sup>50</sup>. If this structure is not  $\beta$ -sheet, it may originate from some other stacked or packed structure perhaps stacked PGII helices that also emits at signal ~ 1670 cm<sup>-1</sup> but whose structure is currently unknown or difficult to determine due to the solid state of the fibers. There is also the possibility that in this solid state another secondary structure can overlap this region besides  $\beta$ -sheet. The difficulty in obtaining samples of native flagelliform silk inhibits supportive structural analysis such as NMR to more clearly determine the structure of the proteins.

## CONCLUSIONS

The *N. clavipes* Flag-like silk proteins designed with four motif-variation combinations of the flagelliform Flag GGX (G), spacer (F) and GPGGX (Y) motifs, produced fibers with

differing mechanical properties. Fibers generated from proteins containing the spacer motif clearly produced the strongest fibers and maintained high extensibility. Fibers produced from proteins composed of only the GGX motif produced the most extensible fibers likely due in part to the reduced interactions between the proteins facilitating molecular mobility. This is evidenced by the great extensibility of the protein fibers containing solely GGX motif repeats. The GGX motif has the potential to form helices and  $\beta$ -sheets and facilitate structural alignment of other motifs that may form  $\beta$ -sheets, nanosprings, or other structures, tethering the structures together while acting as a flexible connector<sup>13, 75</sup>. This would support an amorphous random network polymer containing nanostructures<sup>51</sup>. More research on the GGX motif, focusing on the variations in the amino acids of the X position may find multiple structures with varying functions attributed to this single GGX motif.

The toughest fiber, GFY, with the greatest strength and moderate extensibility was produced from the only protein containing all three of the Flag core motifs. Combinations of only two of the three core motifs (GY and GF) exhibited restricted extensibility and reduced strength. This suggests a more optimal concerted molecular arrangement of the GFY proteins over those proteins in the G, GY, and GF fibers. The spacer motif therefore may function as a “strength-motif” in these Flag-like proteins due in part to  $\beta$ -sheet formation and diverse sequence complexity. Further studies are needed to determine more specifically how the spacer motif confers mechanical properties to the silk fibers.

In the native flagelliform silk, the roll of the spacer motif is likely functioning in a similar fashion as the polyalanine strength-motif functions in dragline silk, in agreement with recent findings from the analysis of native Flag<sup>50</sup>. The strength-motif found in *N. clavipes* flagelliform sequences may also hold true in other orb weaver species with similar motifs. We would expect variations in the sequences to account for variations in the mechanical properties of different flagelliform silks.

These data suggest the Flag GGX motif contributes extensibility and the spacer motif contributes strength to *N. clavipes* flagelliform spider silk fibers. Unfortunately, the limitations of current technology do not allow the elucidation of fine structural differences for proteins within thin fibers<sup>50</sup>. Further studies will be necessary to elucidate and confirm the specific structures as technology and methods advance.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors would like to thank the DOD, DOE, and NIH for funding. The data and views presented here are solely those of the authors and not of the agencies listed above.

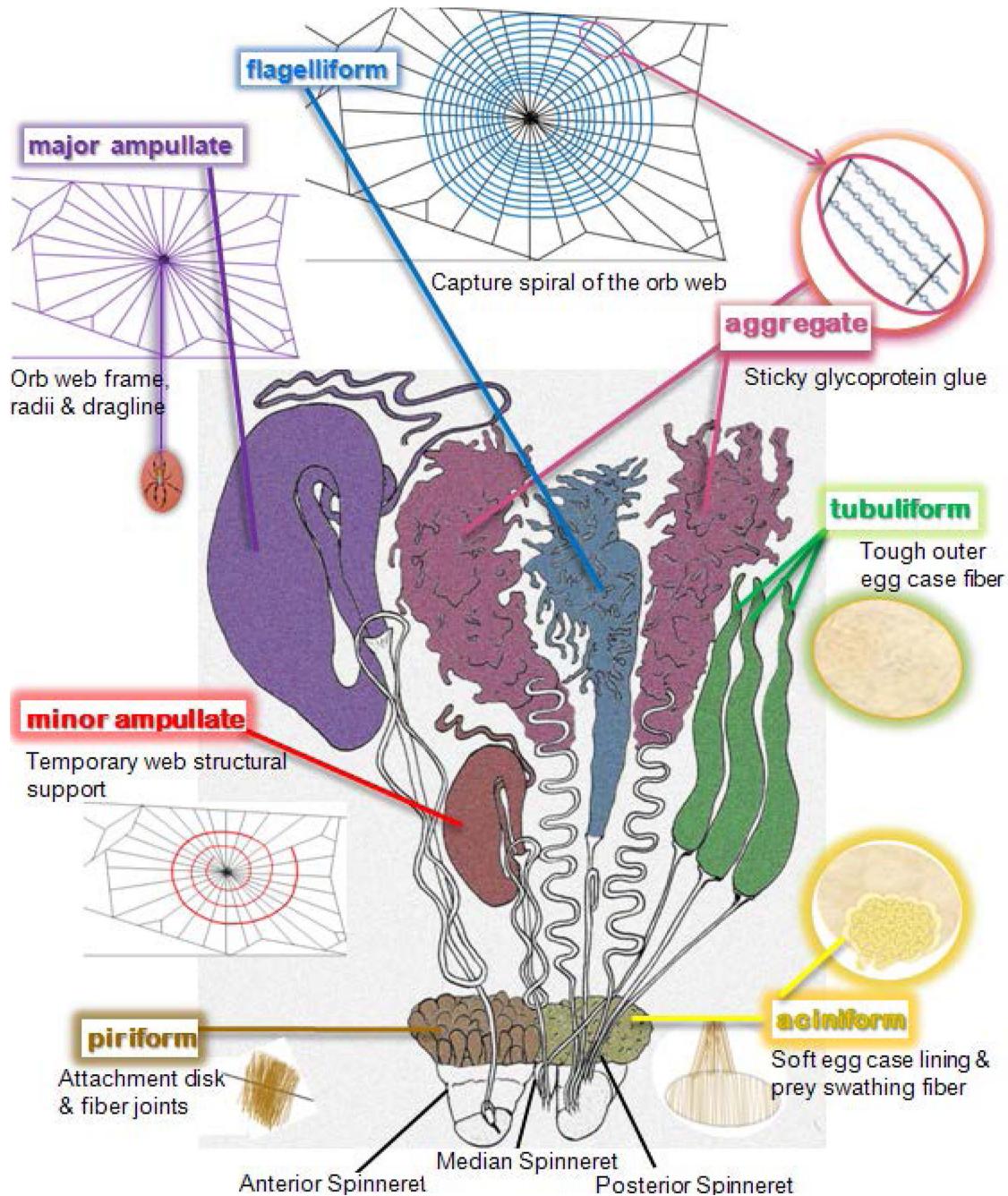
## Works Cited

1. Denny MW. Symp. Soc. Exp. Biol. 1980; 34:247–272. [PubMed: 7256555]
2. Foelix RF. Biology of Spiders. 2nd ed. New York, NY: Oxford University Press, Inc; 1996. p. 330
3. Gosline JM, DeMont ME, Denny MW. Endeavour. 1986; 10(1):37–43.
4. Gosline JM, Guerette PA, Ortlepp CS, Savage KN. J. Exp. Biol. 1999; 202:3295–3303. [PubMed: 10562512]
5. Stauffer S, Coguill S, Lewis RV. J. Arachnol. 1994; 22(1):5–11.
6. Vollrath F. Scientific American. 1992; 266:70–76.
7. Hayashi CY, Lewis RV. J. Mol. Biol. 1998; 275:773–784. [PubMed: 9480768]

8. Hayashi CY, Lewis RV. *Bioessays*. 2001; 23(8):750–756. [PubMed: 11494324]
9. Gatesy J, Hayashi C, Motriuk D, Woods J, Lewis R. *Science*. 2001; 291(5513):2603–2605. [PubMed: 11283372]
10. Vollrath F, Porter F, Holland C. *Soft Matter*. 2011; 7(20):9595–9600.
11. Hinman M, Lewis RV. *J. Biol. Chem.* 1992; 267(27):19320–19324. [PubMed: 1527052]
12. Simmons AH, Michal CA, Jelinski LW. *Science*. 1996; 271(5245):84–87. [PubMed: 8539605]
13. van Beek JD, Hess S, Vollrath F, Meier BH. *Proc. Natl. Acad. Sci.* 2002; 99(16):10266–10271. [PubMed: 12149440]
14. Holland GP, Creager MS, Jenkins JE, Lewis RV, Yarger JL. *J. Am. Chem. Soc.* 2008; 130(30): 9871–9877. [PubMed: 18593157]
15. Holland GP, Jenkins JE, Creager MS, Lewis RV, Yarger JL. *Chem. Commun. (Camb)*. 2008; 5568–5570. [PubMed: 18997954]
16. Savage KN, Gosline JM. *J. Exp. Biol.* 2008; 211(Part 12):1937–1947. [PubMed: 18515724]
17. Savage KN, Gosline JM. *J. Exp. Biol.* 2008; 211(Part 12):1948–1957. [PubMed: 18515725]
18. Jenkins JE, Creager MS, Butler EB, Lewis RV, Yarger JL, Holland GP. *Chem. Commun. (Camb)*. 2010; 46(36):6714–6716. [PubMed: 20733981]
19. Creager MS, Jenkins JE, Thagard-Yeaman LA, Brooks AE, Jones JA, Lewis RV, Holland GP, Yarger JL. *Biomacromolecules*. 2010; 11(8):2039–2043. [PubMed: 20593757]
20. Vollrath F, Fairbrother W, Williams R, Tillinghast E, Bernstein D, Gallager K, Townley M. *Nature*. 1990; 345(6275):526–528.
21. Choresh O, Bayarmagnai B, Lewis RV. *Biomacromolecules*. 2009; 10(10):2852–2856. [PubMed: 19731928]
22. Bonthrone KM, Vollrath F, Hunter BK, Sanders JKM. *Proc. R. Soc. B*. 1992; 248(1322):141–144.
23. Vollrath F, Edmonds DT. *Nature*. 1989; 340(6231):305–307.
24. Liu Y, Shao Z, Vollrath F. *Nat. Mater.* 2005; 4(12):901–905. [PubMed: 16299506]
25. Seidel A, Liivak O, Jelinski LW. *Macromolecules*. 1998; 31(19):6733–6736.
26. Seidel A, Liivak O, Calve S, Adaska J, Ji GD, Yang ZT, Grubb D, Zax DB, Jelinski LW. *Macromolecules*. 2000; 33(3):775–780.
27. Teulé F, Addison B, Cooper AR, Ayon J, Henning RW, Benmore CJ, Holland GP, Yarger JL, Lewis RV. *Biopolymers*. 2012; 97(6):418–431. [PubMed: 22012252]
28. Peters HM. *Z. Naturforsch.* 1955; 10:395–404.
29. Urry DW, Luan CH, Peng SQ. *Ciba Found. Symp.* 1995; 192:4–22. [PubMed: 8575267]
30. Van Dijk AA, Van Wijk LL, Van Vliet A, Haris P, Van Swieten E, Tesser GI, Robillard GT. *Protein Sci.* 1997; 6(3):637–648. [PubMed: 9070446]
31. Zhou Y, Wu S, Conticello VP. *Biomacromolecules*. 2001; 2(1):111–125. [PubMed: 11749162]
32. Becker N, Oroudjev E, Mutz S, Cleveland JP, Hansma PK, Hayashi CY, Makarov DE, Hansma HG. *Nat. Mater.* 2003; 2(4):278–283. [PubMed: 12690403]
33. Hayashi CY, Shipley NH, Lewis RV. *Int. J. Biol. Macromol.* 1999; 24(2–3):271–275. [PubMed: 10342774]
34. Debelle L, Alix AJ, Jacob MP, Huvenne JP, Berjot M, Sombret B, Legrand P. *J. Biol. Chem.* 1995; 270(44):26099–26103. [PubMed: 7592811]
35. Sponner A, Vater W, Rommerskirch W, Vollrath F, Unger E, Grosse F, Weisshart K. *Biochem. Biophys. Res. Co.* 2005; 338(2):897–902.
36. Rammensee S, Slotta U, Scheibel T, Bausch AR. *Proc. Natl. Acad. Sci. U.S.A.* 2008; 105(18): 6590–6595. [PubMed: 18445655]
37. Heim M, Ackerschott CB, Scheibel T. *J. Struct. Biol.* 2010; 170(2):420–425. [PubMed: 20045468]
38. Askarieh G, Hedhammar M, Nordling K, Saenz A, Casals C, Rising A, Johansson J, Knight SD. *Nature*. 2010; 465(7295):236–238. [PubMed: 20463740]
39. Gnesa E, Hsia Y, Yarger JL, Weber W, Lin-Cereghino J, Lin-Cereghino G, Tang S, Agari K, Vierra C. *Biomacromolecules*. 2011; 13:304–312. [PubMed: 22176138]
40. Lazaris A, Arcidiacono S, Huang Y, Zhou JF, Duguay F, Chretien N, Welsh EA, Soares JW, Karatzas CN. *Science*. 2002; 295(5554):472–476. [PubMed: 11799236]

41. Teulé F, Furin WA, Cooper AR, Duncan JR, Lewis RV. *J. Mater. Sci.* 2007; 42:8974–8985.
42. Brooks AE, Lewis RV. *Biomed. Sci. Instrum.* 2004; 40:232–237. [PubMed: 15133963]
43. An B, Hinman MB, Holland GP, Yarger JL, Lewis RV. *Biomacromolecules.* 2011; 12(6):2375–2381. [PubMed: 21574576]
44. Xia X, Xu Q, Hu X, Qin G, Kaplan D. *Biomacromolecules.* 2011; 12:3844–3850. [PubMed: 21955178]
45. Parkhe AD, Seeley SK, Gardner K, Thompson L, Lewis RV. *J. Mol. Recognit.* 1997; 10(1):1–6. [PubMed: 9179774]
46. Dong Z, Lewis RV, Middaugh CR. *Arch. Biochem. Biophys.* 1991; 284(1):53–57. [PubMed: 1989503]
47. Thiel B, Kunkel D, Viney C. *Biopolymers.* 1994; 34(8):1089–1097.
48. Thiel BL, Viney C, Jelinski LW. *Science.* 1996; 273(5281):1480–1481. [PubMed: 8801632]
49. Rikel C, Branden C, Craig C, Ferrero C, Heidelbach F, Muller M. *Int. J. Biol. Macromol.* 1999; 24(2–3):179–186. [PubMed: 10342763]
50. Lefèvre T, Pézolet M. *Soft Matter.* 2012; 8:6350–6357.
51. Qin G, Rivkin A, Lapidot S, Hu X, Preis I, Arinu S, Dgany O, Shoseyov O, Kaplan D. *Biomaterials.* 2011; 32:9231–9243. [PubMed: 21963157]
52. Teulé F, Cooper AR, Furin WA, Bittencourt D, Rech EL, Brooks A, Lewis RV. *Nat. Protoc.* 2009; 4(3):341–355. [PubMed: 19229199]
53. Subramanian S. *Proteins.* 1998; 32:1–2. [PubMed: 9672036]
54. Mangalam, H. TAGC: Analyze a NS for Restriction Enzyme Sites. UC Irvine: SDSC Biology Workbench; 1996. <http://seqtool.sdsc.edu>
55. Prince JT, McGrath KP, DiGirolamo CM, Kaplan DL. *Biochemistry.* 1995; 34(34):10879–10885. [PubMed: 7662669]
56. Lewis RV, Hinman M, Kothakota S, Fournier MJ. *Protein Expr. Purif.* 1996; 7(4):400–406. [PubMed: 8776759]
57. Teulé F, Miao YG, Sohn BH, Kim YS, Hull JJ, Fraser MJ Jr, Lewis RV, Jarvis DL. *Proc. Natl. Acad. Sci. U.S.A.* 2012; 109(3):923–928. [PubMed: 22215590]
58. Scheller J, Guhrs KH, Grosse F, Conrad U. *Nat. Biotechnol.* 2001; 19(6):573–577. [PubMed: 11385464]
59. Sambrook, J.; Fritsch, E.; Maniatis, T.; Russel, D. *Molecular Cloning: A Laboratory Manual.* 3rd ed. ed. New York, NY: Cold Spring Harbor Laboratory Press; 2001. p. 2344
60. Brooks AE, Creager MS, Lewis RV. *Biomed Sci Instrum.* 2005; 41:1–6. [PubMed: 15850073]
61. Brooks AE, Nelson SR, Jones JA, Koenig C, Hinman M, Stricker S, Lewis RV. *Nanotechnol. Sci. Appl.* 2008; 1:9–16. [PubMed: 20657704]
62. Elices M, Guinea G, Plaza G, Karatzas C, Rikel C, Agullo-Reueda F, Daza R, Perez-Rigueiro J. *Macromolecules.* 2011; 44:1166–1176.
63. Brooks AE, Stricker SM, Joshi SB, Kamerzell TJ, Middaugh CR, Lewis RV. *Biomacromolecules.* 2008; 9(6):1506–1510. [PubMed: 18457450]
64. Gruber T, Anderson S, Brewer H, Chen Y-S, Cho H, Dashdorj N, Henning R, Kosheleva I, Macha G, Meron M, Pahl R, Ren Z, Ruan S, Schotte F, Srager V, Viccaro P, Westferro F, Anfinrud P, Moffat K. *J. Synchr. rad.* 2011; 18:658–670.
65. Dicko C, Vollrath F, Kenney JM. *Biomacromolecules.* 2004; 5(3):704–710. [PubMed: 15132650]
66. Gaines WA, Sehorn MG, Marcotte J, William R. *J. Biol. Chem.* 2010; 285(52):40745–40753. [PubMed: 20959449]
67. Brooks AE, Steinkraus HB, Nelson SR, Lewis RV. *Biomacromolecules.* 2005; 6(6):3095–3099. [PubMed: 16283732]
68. Blackledge TA, Hayashi CY. *J. Exp. Biol.* 2006; 209(Pt 13):2452–2461. [PubMed: 16788028]
69. Blamires S, Wu C, Blackledge TA, Tso I. *J. R. Soc. Interface.* 2012; 9(75):2479–2487. [PubMed: 22628213]
70. Qin G, Hu X, Cebe P, Kaplan D. *Nat. Commun.* 2012; 3(1003):1–9.

71. Gosline J, Pollack C, Guerette P, Cheng A, DeMont M, Denny M. ACS Symp. Ser. 1994; 544:328–341.
72. Fukushima Y. Biopolymers. 1998; 45(4):269–279. [PubMed: 9491757]
73. Maji S, Wang L, Greenwald J, Riek R. FEBS Letters. 2009; 583(16):2610–2617. [PubMed: 19596006]
74. Hauser C, Deng R, Mishra A, Loo Y, Khoe U, Zhuang F, Cheong D, Accardo A, Sullivan M, Riekel C, Ying J, Hauser U. Proc. Natl. Acad. Sci. U.S.A. 2011; 108(4):1361–1366. [PubMed: 21205900]
75. Kummerlen J, van Beek JD, Vollrath F, Meier BH. Macromolecules. 1996; 29(8):2920–2928.



**Figure 1. Abdominal silk glands and silks of *Nephila clavipes***

One half of the bilateral pair of the abdominal spider silk glands of *Nephila clavipes* are shown attached via detailed ductwork to their corresponding spinnerets<sup>2, 28</sup>. Pictorial representations of the gland products are illustrated in the margins.

*N. clavipes* Flag: [GPGGX]<sub>43-63</sub>[GGX]<sub>7</sub>TIIEDLDITIDGADGPITISEELTISGA GGS

Amino acid substitutions (X):

[GGX]<sub>7</sub>, X = A, S, T or V;

[GPGGX]<sub>43-63</sub>, X = A, S, Y or V (alternating X/X residues: A/A, S/Y, or V/S)

**Figure 2. Flag protein consensus sequence**

The *N. clavipes* flagelliform silk Flag protein consensus sequence is presented with the three core motifs: GPGGX in blue, GGX in green and the Flag spacer in purple. The corresponding X amino acids of the glycine-rich motifs are noted<sup>7, 9, 33</sup>.

Structural Motif	Module Name	Amino Acid Sequences
GGX	G	GGAGGSAGGSGGVGGSGGT
GPGGX	Y	(GPGGSGPGGY) <sub>4</sub>
Flag spacer	F	TIIEDLDITIDGADGPITISEELTISGAGGS

**Figure 3. *Nephila clavipes* motif sequences used for protein module design**

The *Nephila clavipes* based flagelliform amino acid sequences are shown for the structural motifs and their module designations for this research.

a)

SpeI	BamHI	NdeI	XmaI																	
\	\	\	\																	
actagtggatccgcatatgggccgggtggcgctggcgtagcggtggcgcggtggttc																				
tgatcacctaggcgtaatcccggccaccgcgaccgcacgcaccgcgcaccacaag																				
L	V	D	P	H	M	G	P	G	G	A	G	G	S	G	G	A	G	G	S	
										BspEI		BamHI								
										\	\									
tggtgtgttggcggttctggcgacccgggtccggagatatctggatccaaactg																				
accaccacaaccgccaagaccgcgtggccaggcctctatagacctagggtttgac																				
G	G	V	G	G	S	G	T	G	P	E	I	S	G	S	K	L				

b)

SpeI	BamHI	NdeI	XmaI																
\	\	\	\																
actagtggatccccatatgggccggaggtgccgggggaagtggggagcaggagggtc																			
tgatcacctagggtatacccgggccctcacggcccccctcaccctcgtccctccag																			
L	V	D	P	H	M	G	P	G	G	A	G	G	S	G	G	A	G	G	S
										aggtgtgggtcggcggcagcggggcaccaccatattgaggatctggacatcacaatcga									
										tccaccccaagccgcgtcgccccgtggtagtaactcctagaccctgttagttagct									
G	G	V	G	G	S	G	T	T	I	I	E	D	L	D	I	T	I	D	
										cggcgcgcacggccctatcacaattagcgaagaactcactattagcggagctggcggtc									
										gccgcggctgcgggatagtgttaatcgcttctgagtgataatcgccctgaccgcagg									
G	A	D	G	P	I	T	I	S	E	E	L	T	I	S	G	A	G	G	S
										BspEI		BamHI		HindIII					
										\	\		\						
cggtccggagatatctggatccaagctt																			
gccaggcctctatagacctagggtcgaa																			
G	P	E	I	S	G	S	K	L											

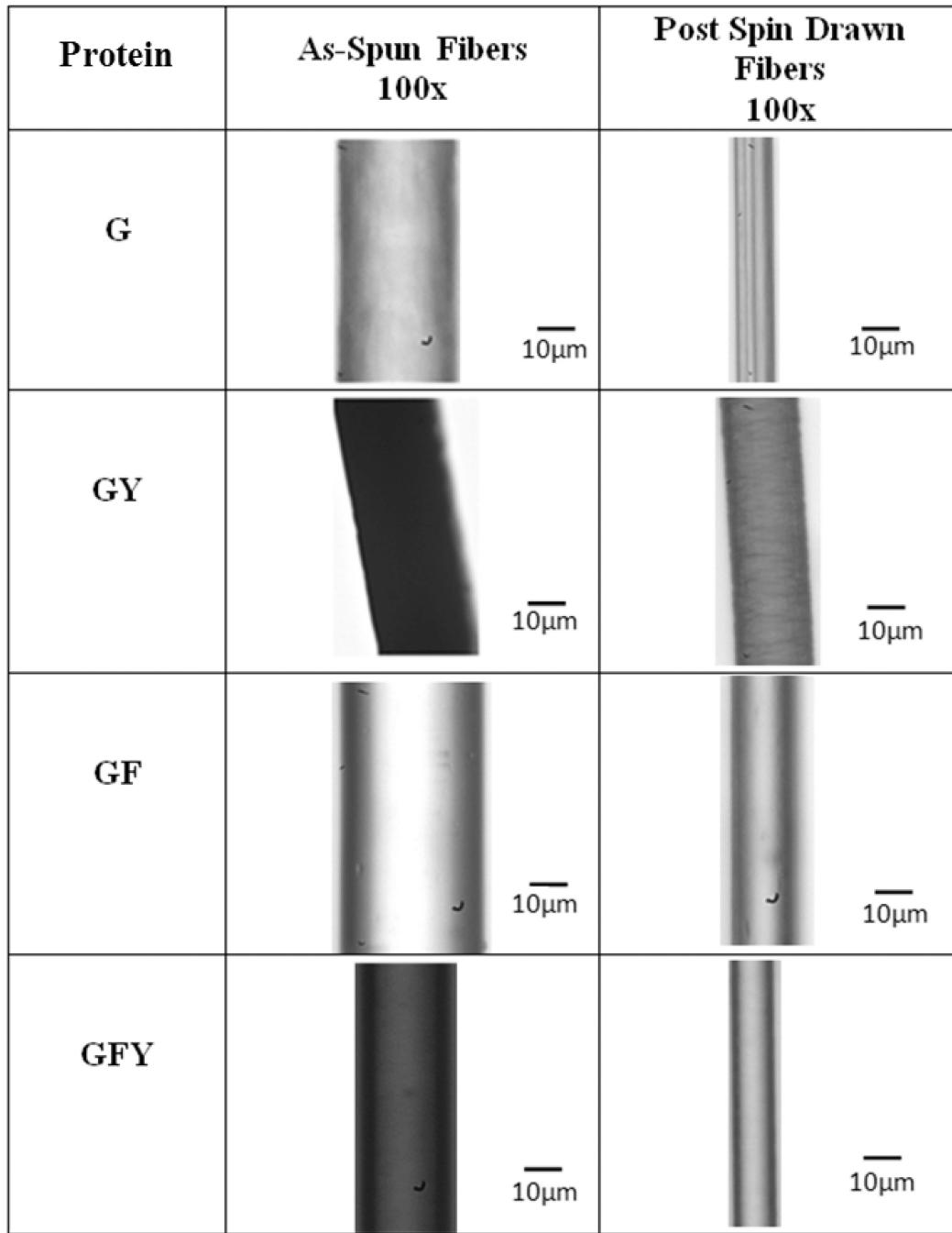
**Figure 4. Restriction sites and amino acids of designed dsDNA sequences**

The restriction enzyme sites that were utilized or designed into the vector sequences are shown above the dsDNA sequences. The corresponding amino acid sequences are shown below the designed dsDNA sequences using the TAGC program from Biology Workbench<sup>53,54</sup>. a) The G module and b) the GF module sequence designs.

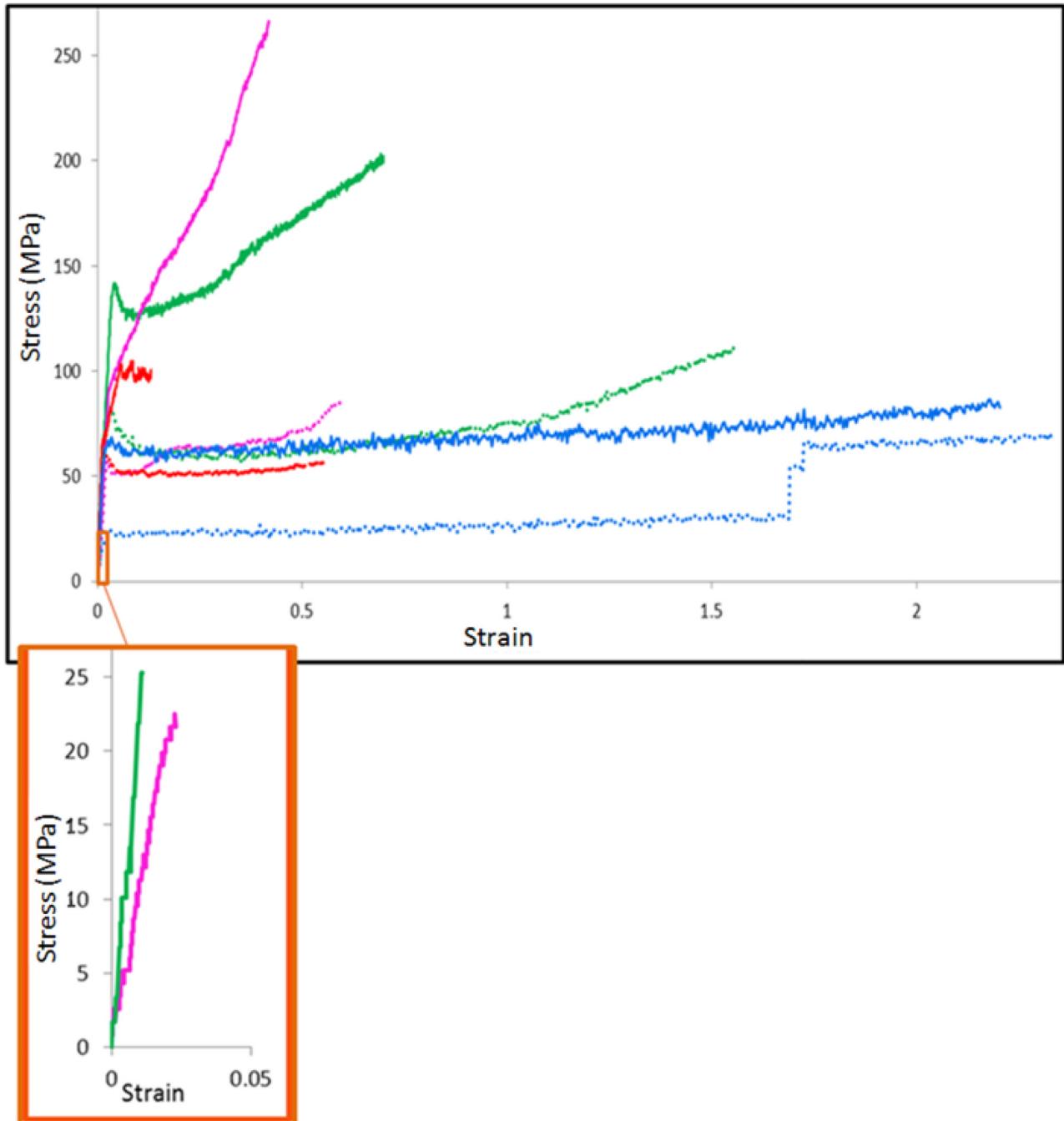
Protein Designation	Primary sequence
G	[GGAGGSGGAGGSGGVGGSGGT] <sub>32</sub>
GY	[GGAGGSGGAGGSGGVGGSGGT(GPGGSGPGGY) <sub>4</sub> ] <sub>12</sub>
GF	[GGAGGSGGAGGSGGVGGSGGTIIIEDLDITIDGADGPITISEELTISGAGGS] <sub>12</sub>
GFY	[GGAGGSGGAGGSGGVGGSGGTIIIEDLDITIDGADGPITISEELTISGAGGS(GPGGSGPGGY) <sub>4</sub> ] <sub>8</sub>

**Figure 5. Primary sequences of the four recombinant protein variants**

The protein designation for the G, GY, GF, and GFY proteins associated with their primary sequence and motif module content.

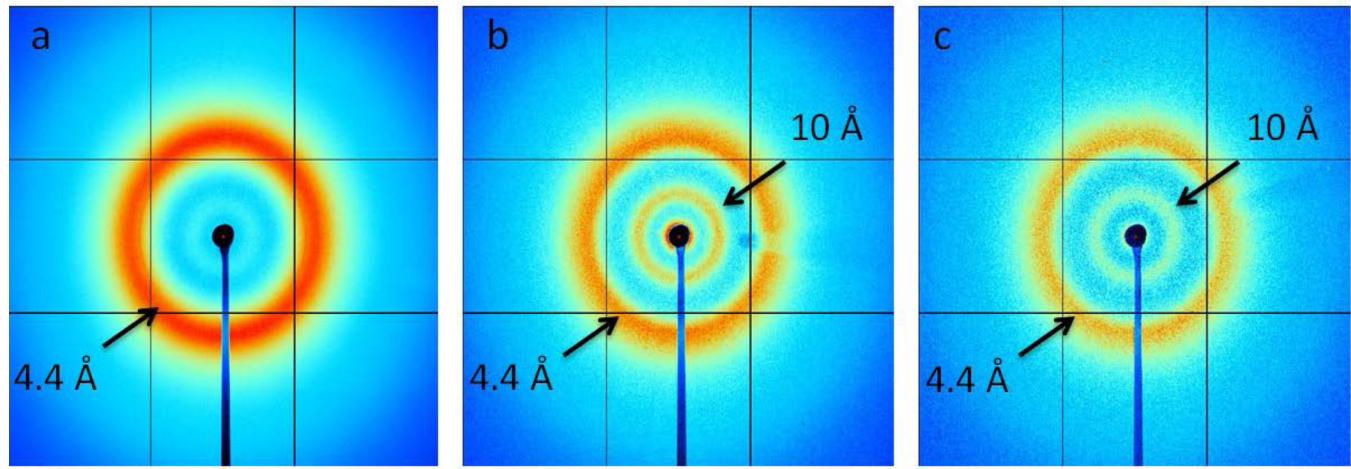


**Figure 6. Fibers spun from the four protein variants**  
 As spun (AS) fibers and post-spin drawn (PSD) fibers observed at 100 $\times$  light microscopic magnification.



**Figure 7. Comparative stress/strain analysis of top performing fibers**

The stress/strain data of the post-spin drawn (PSD) fibers that showed the highest tensile strength (solid lines) or greatest % extensibility (dotted lines) are presented. Fiber Color Key: G = blue; GY = red; GF = pink, and GFY = green in both graphs. The small orange boxed region is enlarged and horizontally stretched to show the stress/strain analyses of the top performing as-spun (AS) fibers for GF (1 of 5) and GFY (1 of 3). These data from the single fibers shown, possessed both the highest tensile strength and greatest % extensibility of these AS fiber types. The G and GY AS fibers were too brittle to test.



**Figure 8. Wide Angle X-Ray Scattering analysis of the as-spun fibers**

Wide angle X-ray scattering (WAXS) patterns of GY (a), GF (b), and GFY (c) as-spun synthetic fibers. The beam stop and fiber axis were both aligned vertically. These data were collected with a detector distance of 300 mm, an exposure time of 60 seconds, and an X-ray wavelength of 0.9787 Å. Black arrows indicate rings centered at 4.4 Å or 10 Å that are visible in the different WAXS patterns.

**Table 1**  
**Module ensembles, iterations, encoding DNA data & protein size**

The designed constructs were generated to produce recombinant proteins. The various encoding dsDNA and protein module names are designated based on the motif content.

Flag Motif Module Ensembles	Module Ensemble Iterations	Size of DNA Coding Sequence kbp	Molecular Wt. of the Recombinant Protein kDa	DNA or Protein Construct Designation
(GGX) <sub>7</sub>	32×	2.31	54	G
(GGX) <sub>7</sub> + (GPGGX) <sub>8</sub>	12×	2.37	60	GY
(GGX) <sub>7</sub> +(spacer)	12×	2.05	59	GF
(GGX) <sub>7</sub> +(spacer)+ (GPGGX) <sub>8</sub>	8×	2.37	66	GFY

**Table 2**  
**Mean mechanical data from spun protein fibers**

The average mechanical data collected for both as-spun (AS) and post-spin drawn (PSD) fiber groups are presented with their respective standard deviations. The mechanical data for the few AS fibers that were included are only included for observation as most of the AS fibers broke.

Fiber	Diameter microns	Extension % (strain × 100)	Tensile Strength MPa (stress)	Toughness MJ.m <sup>-3</sup>
G-AS	*			
G-PSD	<b>13.31</b> ± 2.29	<b>132.8</b> ± 76.28 <sup>1,2,3</sup>	<b>55.73</b> ± 16.05 <sup>2,3</sup>	<b>61.55</b> ± 47.96
GY-AS	*			
GY - PSD	<b>21.46</b> ± 4.51	<b>45.38</b> ± 43.47 <sup>1,5</sup>	<b>47.07</b> ± 25.11 <sup>4,5</sup>	<b>17.77</b> ± 23.24
GF-AS	<b>36.81</b> ± 1.36	<b>1.09</b> ± 0.94	<b>18.73</b> ± 4.90	<b>0.12</b> ± 0.11
GF - PSD	<b>20.43</b> ± 3.89	<b>36.61</b> ± 12.45 <sup>2,6</sup>	<b>136.42</b> ± 60.27 <sup>2,4</sup>	<b>35.68</b> ± 14.71
GFY - AS	<b>27.24</b> ± 0.77	<b>0.66</b> ± 0.41	<b>26.31</b> ± 17.69	<b>0.06</b> ± 0.06
GFY - PSD	<b>15.06</b> ± 1.29	<b>84.5</b> ± 37.82 <sup>3,5,6</sup>	<b>150.58</b> ± 31.32 <sup>3,5</sup>	<b>89.05</b> ± 23.93

The complete absence of data due to fiber brittleness is indicated by an asterisk (\*).

The PSD fiber superscripts for % Extension and Tensile Strength indicate significant differences between fiber types where p<0.05 (Supplemental Table S1), as follows: 1 = G and GY, 2 = G and GF, 3 = G and GFY, 4 = GY and GF, 5 = GY and GFY, and 6 = GF and GFY.