To The University of Wyoming:

The members of the Committee approve the dissertation of Sherry L. Adrianos, presented on April 18, 2012.

Randolph V. Lewis, Chairperson

Stephen K. Herbert, External Department Member

Pamela J. Langer, Co-Chairperson

Kurt W. Miller

Anne W. Sylvester

APPROVED:

Dr. Mark Stayton, Department Chair, Department of Molecular Biology

Dr. Donna M. Brown, Associate Dean, College of Agriculture and Natural Resources

Adrianos, Sherry L., Protein Fiber Extensibility and Strength: GGX and Spacer Motif Contributions to Mechanical and Structural Properties of Flagelliform Spider Silk Determined through Recombinant Protein Characterization, Ph.D.,
Department of Molecular Biology, December 2012.

Abstract: Flagelliform spider silk is a highly extensible protein polymer produced by orb weaver spiders with remarkable mechanical properties combining strength with incredible extensibility. The flagelliform fibers are composed of a single protein, Flag, and can stretch up to six times their original length. The core of the large 400 kDa Flag protein is composed of three simple amino acid motifs. The GPGGX motif has been previously shown to contribute the elasticity of the silk fibers but the source of the strength of these fibers has been elusive.

The GGX and "spacer" core motifs were selected for the purpose of this study to determine their specific contributions to the fiber mechanical properties and structure. Four recombinant proteins were designed and genetically engineered to contain repeats of the GGX; the GGX and GPGGX, the GGX and spacer; or the GGX, spacer and GPGGX modules. The recombinant proteins, ranging in size from 55-65 kDa, were utilized as molecular tools to investigate the mechanical property contributions of the individual motifs to the silk fibers.

Structural studies of the recombinant silk fibers provided new insight into motif structural contributions of the native flagelliform Flag protein. Raman spectroscopic analyses correlate with an increase in β-turn structure with increased spacer motifs in the recombinant proteins. Wide Angle X-ray diffraction studies provide evidence of increased molecular structural orientation correlating with the inclusion of the spacer motif. The GGX and GPGGX fibers produce amorphous X-ray diffraction data. These results suggest that the spacer motif is a major contributor to the strength of native flagelliform silk. These recombinant proteins provide a platform for additional studies that will contribute to the understanding of the structure/function correlation that produce the material mechanical properties of flagelliform silk.

PROTEIN FIBER EXTENSIBILITY AND STRENGTH: GGX AND SPACER MOTIF CONTRIBUTIONS TO MECHANICAL AND STRUCTURAL PROPERTIES OF FLAGELLIFORM SPIDER SILK DETERMINED THROUGH RECOMBINANT PROTEIN CHARACTERIZATION

by

Sherry Lee Adrianos

A dissertation submitted to the Department of Molecular Biology

And to the University of Wyoming

In partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY
in
MOLECULAR BIOLOGY

Laramie, Wyoming

December 2012

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Acknowledgements

I would like to express my deep appreciation to my advisor, Randy Lewis, for his support and for the opportunity to work on this project. I would like to thank my co-advisor Pam Langer for the extra efforts she has made, the use of her lab, and for her encouragement, support, and advice throughout the this project. I would like to thank my committee members, Kurt Miller, Anne Sylvester, Pamela Langer, and Steve Herbert for their constructive comments and suggestions on the project and dissertation.

I would like to thank all my fellow lab members and coworkers for the discussions that we have had and the science that we have shared. Special thanks go to Shane Nelson, Mike Hinman, Florence Teulé, Justin Jones, Holly Steinkraus, Heather Rothfuss, and David Perry for their technical and writing advice. I would like to thank Ted John for his assistance and expertise in the lab. I would like to thank my fellow graduate students, Kelley Moss, Amy Albertson, and Leah Kyle for their camaraderie and support in addition to previous graduate students, Bo An, Melinda Creager, Amanda Brooks, Mara Motriuk-Smith, and Maozhen Tian. I would also like to thank Karen White and Jackie Keele for their friendship and assistance, always ready to help out with Science Posse. Special thanks go to Christoph Geisler and Kelley Moss for being there, especially for making the lab a safer place to work late at night.

I would like to thank Terry McClean for DNA sequencing and related assistance, Virginia Schmit for assistance with Raman spectroscopy, Zackie Salmon and Susan Stoddard from the McNair Program. I would also like to thank Jeff Yarger of Arizona State University, Joel Ayon, Warner Webber, and Anthony Karrick for their assistance with Raman and XRD analyses and methods and Chris Benmore at the Argonne National Laboratory for assistance with XRD. I would like give additional thanks to Florence Teulé for collecting XRD data on the fibers, discussions, and comments toward paper writing.

I would like to thank the Molecular Biology Department staff: Vicki, Lexi, Debbie, Matt, and Meghan for their assistance. Special thanks to Mark Stayton and to Don Jarvis for their encouragement and counsel. Thanks go to Professors, Lee Belden, Nancy Peterson, Jordanka Zlatanova, Peter Thorsness, David Fay, and David Liberles. I would like to thank Don Roth and the members of the Science Posse for the camaraderie and educational experience. Thank you all, I would not be here without your input.

This work is dedicated

To my very best friend and loving husband, James,

Whose love, patience, encouragement, and support helped make this work possible

and to our children: Sarah, Joshua, and Rebekah

who have grown up to become incredible human beings

with appreciation for their love and encouragement

And in loving memory of:

Donald J. Carneal

Robert B. & Leona D. Carneal

William & Blanche Schell

Shirley (Hill) Webb

James J. and Martha Adrianos

Lynn Chenot &

Belinda Schofield

Table of Contents

List of Tables

List of Figures

Abbreviations and Definitions

AS – As-spun

As-spun fibers **–** *Ex vivo* synthetically produced silk fibers as they are spun in the lab without additional processing

Coagulation bath – Isopropanol solution used to solidify *ex vivo* synthesized silk fibers

Flagelliform silk – The spider silk fiber used by orb weavers to make the capture spiral of the orb web (no abbreviation)

Flag – The major protein of flagelliform silk fibers (not the abbreviation for flagelliform)

MaSp – Major ampullate spidroin protein from the **major** ampullate silk gland

MiSp – Minor ampullate spidroin protein from the **minor** ampullate silk gland

Motif – An amino acid sequence that forms a molecular structure

Motif module – A double-stranded DNA sequence that codes for a single or series of iterated motifs that are designed for the production of synthetic proteins

PSD – Post-spin drawn

Post-spin drawn fibers **–** *Ex vivo* synthetically produced silk fibers that are further processed from the "as-spun" state by stretching in aqueous isopropanol

Post-spin draw – The method of stretching as-spun fibers submerged in aqueous isopropanol or other alcohol solution to improve the fiber mechanical properties.

Spidroin – Spider silk-specific protein fibers

Spin dope – Liquid protein solution used to spin protein fibers

This page is intentionally left blank.

Chapter 1: Introduction

Flagelliform Spider Silk

GGX and Spacer Motifs

Introduction

The specific aim of this research was to investigate the mechanical property contributions of the Flag GGX and spacer motifs found in the flagelliform silk of the orb weaver *Nephila clavipes*. The mechanical properties of recombinantly produced Flag-like silk protein fibers were studied to understand the protein motif-specific structure/function correlation of distinct protein motif modules found in flagelliform spider silk. The sequence and number of each module was specifically tailored to clarify how protein structure motif characteristics alter protein mechanical properties. A genetic engineering method exploiting single, paired or triple motif combination arrays was utilized to help define the motif contributions. The sequence/property correlation specifically helps associate data results to the native Flag protein. The recombinant protein characterization permits the investigation of specific structural motif combinations in addition to individual components.

The flagelliform silk fiber from orb weaver spiders (Araneae) is an intriguingly tough protein polymer. This silk is one of several distinct types of silk fibers that orb weavers produce. The different spider silks possess an impressive array of strength and extensibility, which are indicators of the mechanical performance of the fibers (Gosline *et al*., 1999). These properties vary depending on the particular silk and are comparable to other materials (Table 1.1; Gosline *et al.,* 1986; Stauffer *et al.*, 1994; Vollrath *et al*., 1992). The mechanical properties of each type of silk fiber directly correlate with the

Material	**Strength** GPa	**Elasticity** %	**Energy to Break** $KJ.kg^{-1}$
Major ampullate silk	4	35	400
Minor ampullate silk	1	5	30
Flagelliform silk	1	>200	400
Kevlar	4	5	30
Rubber	0.001	600	80
Tendon	1	5	5
Silkworm silk B. mori	0.5	15[1]	60

Table 1.1. The mechanical properties of various material fibers.

The mechanical properties of the three orb web silks are displayed with manmade and natural materials for comparison. Strength is generated from the greatest stress to failure. Elasticity % is the percentage extension stretch of a material without deformation and its return to the original length/shape after the force is removed. The exception to elasticity % is noted with a superscript 1, where this value represents the percent extension until the fiber breaks in this one case. Energy to Break is the area under the curve of the stress/strain data and is also called toughness (data compiled from: Gosline *et al.*, 1986; Stauffer *et al.*, 1994; Shao & Vollrath 2002).

primary sequences of the specific silk protein(s) and the molecular structures that are produced. Despite previous spider silk research and advances in molecular biology, our understanding of spider silk proteins (spidroins) is limited, yet there is much more to be learned from studying these proteins (Vollrath *et al*., 2011).

Flagelliform silk is the specific silk of interest in this project, yet the orb weaver *Nephila clavipes* produces up to six different proteinaceous silk fibers and one glycoprotein-glue. Each type of silk protein has different structural and mechanical properties that are utilized by the spider for different purposes (Table 1.1; Fig. 1.1, Foelix 1996). The proteins are secreted by columnar epithelial cells lining each abdominal silk gland (Bell and Peakall, 1969) and stored in a liquid state (Willcox *et al.,* 1996; Vollrath & Knight, 2001). The silk protein transitions from an aqueous liquid protein spin-dope, to insoluble fibers as they are pulled from the spinnerets by the spider (Gosline *et al*., 1986). This *in-vivo* liquid-to-solid transition mechanism is not fully understood and adds additional challenges to *ex-vivo* recombinant fiber production.

Whether native spider silk protein or recombinant protein, the transition from a liquid state to a solid fiber is a process that is typically referred to as "spinning" and is often associated with the construction of the orb web by the spider. The study of the different spider silk primary sequences provides a platform for a better understanding of the protein folding that occurs as the spin-dope transitions from a liquid to a solid. Studies of recombinant flagelliform proteins built from defined subunits will further the understanding of the proteins and their intramolecular interactions.

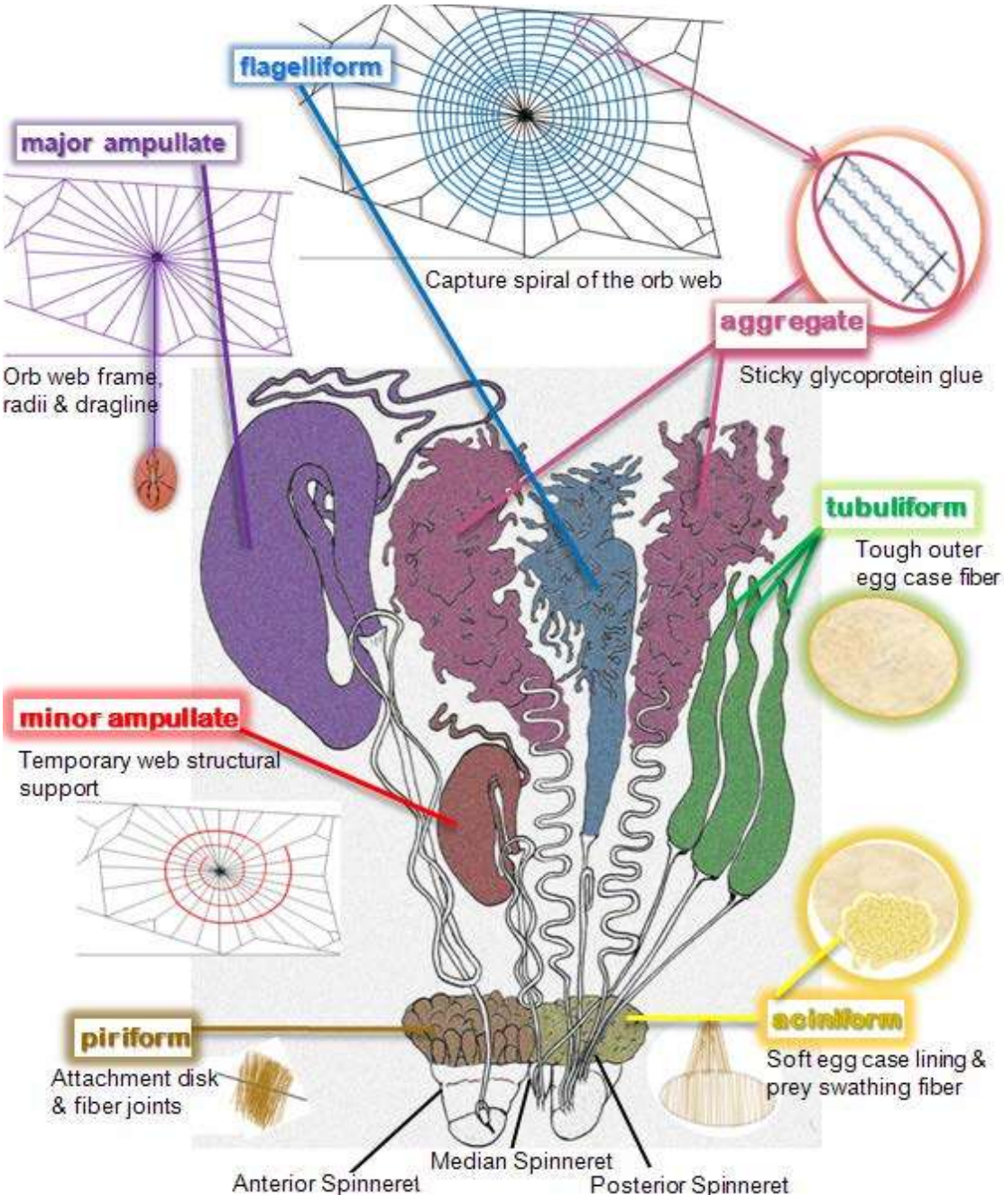


Figure 1.1. Abdominal silk glands and silks of *Nephila clavipes*.

One half of the bilateral pair of the abdominal spider silk glands of *Nephila clavipes* are shown attached via the ductwork to their corresponding spinnerets (Modified by Sherry Adrianos from Foelix 1996). Pictorial representations of the gland products are illustrated in the margins.

The flagelliform, major ampullate (dragline) and minor ampullate spider silks are the three primary orb web components (Fig. 1.1). The silk fibers are produced from the proteins Flag, major ampullate (MaSp) 1 and 2, and minor ampullate (MiSp) 1 and 2, respectively. The mechanical properties of these silks are compared to other materials in Table 1.1. The incredible extensibility of flagelliform silk (>200%) and the high tensile strength of dragline silk (4 GPa) make these two silks the toughest of all the spider silks though their strengths and extensibilities differ (Table 1.1). Flagelliform silk has elastic properties similar to rubber. Both dragline and flagelliform silk are much tougher than Kevlar (Table 1.1). Minor ampullate silk is a strong silk that is not very extensible and comparable to tendon in strength and elasticity (Table 1.1.). These are the three primary web silks though researchers are investigating other types of spider silk as well.

Spider Silk Motifs

A motif pattern emerged as the sequences were discovered for various orb weaver spider silks that suggested a modularity of the silk proteins (Hayashi *et al.,* 1998 & 2001; Gatsey *et al.,* 2001). The silk protein sequences of several orb weavers were aligned and five motifs common to spider silk were reported: 1) GPGXX, 2) a "linker", 3) GGX, 4) $A_n/(GA)_n$, and 5) a "spacer" (Hayashi *et al.,* 1998, 1999, 2000, & 2001; Gatesy *et al*., 2001). The amino acid motifs found in spider silk have been conserved for 136 million years (Garb *et al.,* 2006) and provide a basis for protein study.

Three of the five common spider silk motifs are present in the Flag protein of flagelliform silk: the elastic β-spiral GPGGX, the polyglycine (PG) II helix GGX, and the spacer, so called because it interrupts the glycine rich regions of the two prior motifs (Hayashi *et al.,* 1998). The linker and $A_n/(GA)_n$ motifs are both absent from the Flag protein. The $A_n/(GA)_n$ motif forms antiparallel crystalline β-sheets common to the ampullate proteins (Simmons *et al*., 1996; Parkhe *et al*., 1997; Jelinski *et al*., 1999; Holland *et al*., 2008). Four common spider silk motifs are depicted in Fig. 1.2 from the orb web silks.

By correlating the mechanical properties of the toughest two fibers, flagelliform and dragline silks, with the motifs that make up their respective amino acid sequences, a clear modular source for both strength and extensibility can be determined (Hayashi *et al.*, 1999). Flagelliform silk has a high concentration of the GPGXX motif. The 41-67 contiguous iterations of the GPGXX motif generate 75% of the sequence of the Flag protein. A comparison of the GPGXX motif content of Flag with that of the MaSp2 protein of dragline silk which contains 4-6 iterations of the GPGXX motif followed by one of the A_8 strength motifs, reveals three major characteristics. First, the Flag protein has no A_n motif to generate interlocked β-sheets for strength. Second, the Flag GPGXX motif is continuous for hundreds of amino acids while the GPGXX repeats of the MaSp proteins are interrupted at least twice every fifty amino acids by the $A_n/(GA)_n$ motifs, likely suppressing in MaSp2, the longer β-springs from iterated GPGXX motifs that can form in the Flag protein. Third, the MaSp 1 and 2 proteins have no spacer region. These

points are clear when observing the silk protein sequences in Fig. 1.2 (Hayashi *et al*., 1998 & 2001; Gatsey *et al*., 2001) and then comparing relative extensions and strengths in Table 1.1.

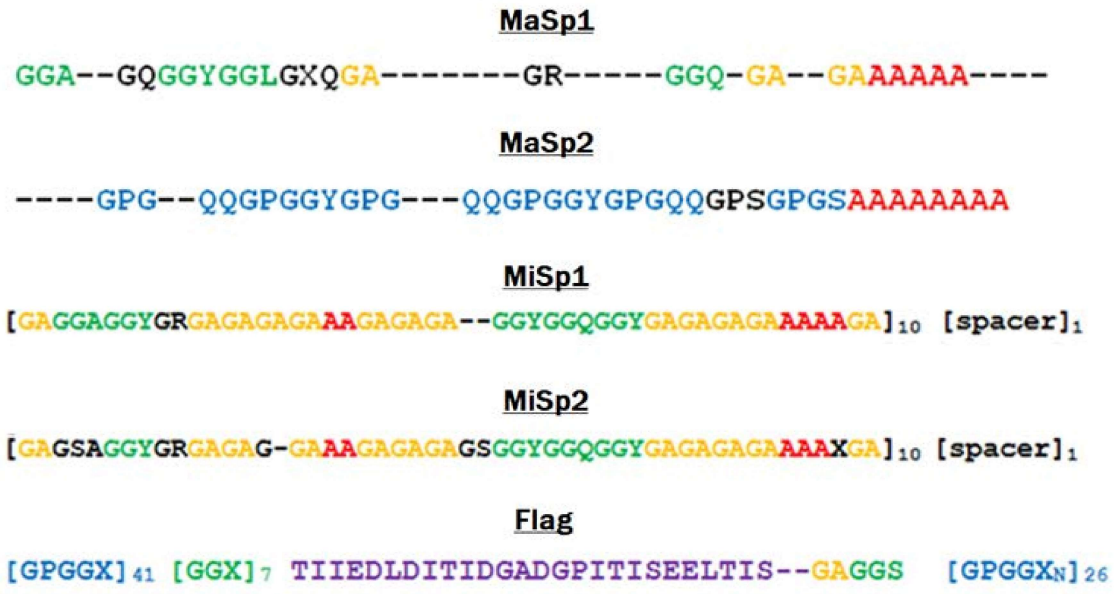


Figure 1.2. The spider silk protein sequence repeats of the three primary orb web silks of *Nephila clavipes*.

The consensus amino acid sequences of orb weaver spider silk proteins (spidroins) of the three primary orb web silks from *Nephila clavipes* are shown. Motifs are common between flagelliform (Flag), dragline (MaSp) 1 and 2, and minor ampullate (MiSp) 1 and 2 silk proteins. The blue GPGXX motif forms β-spirals and confers elasticity (Hayashi *et al*., 1998; Jenkins *et al.,* 2010; Teulé *et al*., 2012). The β-sheet forming, polyalanine red A_n and yellow $(GA)_n$ motifs confer strength (Simmons *et al*., 1994; Parkhe *et al*., 1997; An *et a*l., 2011). The purple Flag spacer function is not known. The black MiSp "spacer" structure/function is also unknown. The green GGX motif is predicted to form polyglycine II helices (Gatesy *et al*., 2001; Hayashi *et al,* 2001; Holland *et al.,* 2008).

Flagelliform strength must be conferred by regions other than polyalanine sequences, which are absent in the Flag protein. Previous research has contributed to what is known about the three Flag motifs, the GGX, the spacer and the GPGXX. While the GPGGX motif undoubtedly plays a role in the overall mechanical properties of the flagelliform silk protein, its primary contribution is elasticity (Hayashi *et al.,* 1999; Jenkins *et al*., 2010; Teulé *et al.*, 2012). The lesser studied flagelliform-specific silk protein structural motifs GGX and Flag-spacer are not well understood, yet it seems apparent that the GGX or the Flag spacer must contribute strength to the fiber.

The GGX motif was hypothesized to produce polyglycine II turns that stack or pack in such a way to add strength to the fiber through the formation of inter-helix hydrogen bonds cross-linking the helices (Dong *et al.,* 1991; Kümmerlen *et al.,* 1996; Simmons *et al.,* 1996) or β-sheets (Thiel *et al.,* 1994; Gosline *et al.*, 1994; Bykov *et al.,* 2010). The alanine contributes hydrophobicity and the serines help with solubility (Zhou *et al.,* 2009). In previous studies, the peptide GLGGQGGGAGQGGTG was shown to form β-sheets in a thin film, however this GGX containing peptide is more MaSp1-like than Flag-like (Fukushima *et al.,* 1998). Studies have shown that small GGAGGAGGS peptides produce nanorods suggesting interactions occur between GGX motifs of neighboring proteins facilitating fiber formation (Zhou *et al.,* 2009). NMR studies of the dragline GGX motif determined a 3_1 helix (Holland *et al.,* 2008) although differences in the X residues in the third position between the Flag and MaSp proteins may facilitate the formation of multiple structures. In spite of what has been learned about the GGX motif,

the structure remains unclear (Vendrely *et al*., 2007). Determining the Flag structural motif combinations and the compound structure/function relationship of variations of Flag motifs will clarify the role of specific motifs in protein biomechanics.

The spacer region was hypothesized to contribute strength to flagelliform fibers (Hayashi *et al*., 1998). There is a lack of Flag biophysical data relative to the structure and function of the spacer motif, therefore the spacer structure and function are not yet fully understood. Recently, the spacer region was reported to be responsible for the strength of flagelliform silk (Lefévre *et al.,* 2012) by a new Raman spectroscopy method that is similar to methods used to determine elastin structure (Debelle *et al*., 1995). While no one structural conformation absolutely explains elasticity (Cheng *et al*., 2010), a dynamic structure concept is likely for other fiber properties such as strength.

In chapter two, the details of this research project are presented in a manuscript submitted for publication describing a set of recombinant proteins specifically designed for experimental study. The goal of the study was to investigate the structure/function and mechanical properties of the Flag-like GGX and spacer motifs in synthetic silk fibers. The results of the research and a brief discussion are included in chapter 2. The final chapter presents a discussion of the research of this dissertation project and a few suggestions for future Flag protein research. The appendices follow with the collected stress/strain data on each of the fiber groups (Appendices A.1. to A.6), mechanical property methods (Table A.1) and cloning doubling strategy (Fig. A.7). The work

depicted within this dissertation indicates that protein motif content affects the characteristics of the synthetic fibers. The flagelliform GGX motif contributes extensibility and the spacer motif contributes strength to the polymeric proteins in synthetic silk fibers, adding to the understanding of the Flag protein.

This page is intentionally left blank.

Chapter 2

Flagelliform GGX Motifs Contribute Extensibility and Spacer Motifs Contribute Strength to Synthetic Spider Silk Fibers

Chapter two is a manuscript prepared for submission to the Journal of Biomacromolecules. The complete paper is included on the following pages of this chapter. The paper encapsulates the main body of work undertaken through the years working on the dissertation. The different recombinant proteins produced for this project proved to be valuable tools, facilitating increased understanding of the mechanical properties and protein structure of flagelliform silk fibers. The analyses of the recombinant protein fibers produced useful data in the study of the Flag protein. The stress/strain data for each different fiber group discussed in chapter two are found in the appendix section for additional reference (Appendices A.1-A.6). Chapter three follows with additional comments and direction for future work.

Flagelliform GGX Motifs Contribute Extensibility and Spacer Motifs Contribute Strength to Synthetic Spider Silk Fibers

Sherry L. Adrianos,[1] Florence Teulé,[2] Michael B. Hinman,[2] Justin A. Jones,[2] Virginia Schmit,[1] Joel Ayon,[3] Jeffery L. Yarger,[3,4] Randolph V. Lewis,[2]

[1]Department of Molecular Biology, University of Wyoming, Laramie, Wyoming 82071

[2]Department of Biology, Synthetic Biomanufacturing Center, Utah State University, Logan, Utah 84422-5305

[3]Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604

[4]X-ray Science Division, Advanced Photon Source, Argonne National Laboratories, Argonne, Illinois, 60439

Correspondence to: Sherry L. Adrianos; e-mail: 7SherryA@gmail.com

Keywords:

Flagelliform, spacer, Flag, GGX, Raman, XRD

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. To investigate the mechanical contributions and structure of the flagelliform GGX and spacer motifs, four recombinant proteins were produced containing variations of the three core motifs of the flagelliform Flag protein: GGX, spacer, and GPGGX. 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. Raman spectroscopy and X-ray diffraction analysis of the recombinant fibers indicate that the spacer motif orchestrates a structural architecture not clearly defined but responsible for increased mechanical strength. These results indicate that the spacer is a primary contributor of strength in the native Flag protein fibers of flagelliform silk.

INTRODUCTION

Native flagelliform spider silk is a tough proteinaceous fiber produced by orb weaver spiders to make the capture spiral of the orb web (Denny 1980; Foelix 1996). The strength and remarkable extensibility of this fiber produces mechanical properties tougher than most natural and manmade materials, including Kevlar (Stauffer *et al.,* 1994; Gosline *et al.,* 1999). Much of what is known about spider silk comes from the studies of major ampullate (dragline) silk, the strongest spider silk (Hinman *et al*., 1992; Simmons *et al.,* 1996; Gosline *et al*., 1999; van Beek *et al.,* 2002; Holland *et al*., 2008; Savage *et al*., 2008; Jenkins *et al*., 2010; Creager *et al.,* 2010). However, unlike dragline silk, flagelliform silk is difficult to obtain in appreciable quantities for fiber structural analysis because it cannot be silked from the spider. Therefore recombinant protein techniques provide tools that will further our understanding of the structure/function relationships of the native flagelliform silk protein.

The prey-capturing adhesiveness of flagelliform silk originates from the application of an aqueous glycoprotein "glue" coating. Hygroscopic components in the glycoprotein attract water from the environment. The water keeps the glycoprotein and silk fibers of the web hydrated (Vollrath *et al.,* 1990; Choresh *et al.,* 2009). Hydration of native flagelliform silk by the aqueous glue has been shown to supplement extensibility (Bonthrone *et al.,* 1992; Vollrath & Edmonds, 1989). Water naturally plasticizes silk proteins and thus plays a critical role in the mechanical properties of native silk fibers

(Gosline *et al.,* 1986; Liu *et al*., 2005). 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 (Seidel *et al.*, 1998 & 2000; Teulé *et al*., 2012). 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 (Hayashi *et al.,* 1998).

Flagelliform silk is composed of a single very large repetitive protein, Flag. The Flag protein is approximately 360 kDa in size with a glycine rich protein core that forms 90% of the protein (Hayashi *et al*., 1998 & 2001). Three motifs dominate the Flag protein: GPGGX, GGX (X=A, S, V, and Y or T), and a glycine poor "spacer" (Fig 2.1). Iterations of the GPGGX motif, $(GPGGX)_n$, provide elasticity and extensibility produced from type II β-turns (Urry *et al.,* 1995; Van Dijk *et al*., 1997; Zhou *et al.,* 2001) generating β-turn nano-spring structures (Hayashi *et al.,* 1999; Becker *et al.,* 2003; Creager *et al.,* 2010; Jenkins *et al.*, 2010). The β-turn generated structures are similar to those found in the elastic proteins elastin and gluten (Urry *et al.,* 1995; Debelle *et al*., 1995; Van Dijk *et al.,* 1997). The $(GPGGX)_{43\text{-}62}$ 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 $(GGX)_{7\text{-}12}$ (Fig. 2.1). The negatively charged spacer motif, composing collectively 7.5% of the protein core in near equal proportion to the $(GGX)_{7\text{-}12}$ motifs, interrupts the long glycine rich arrays. Each Flag protein contains up

N. clavipes Flag: $[GPGGX]_{43\text{-}63}[GGX]_7$TIIEDLDITIDGADGPITISEELTISGA GGS

Amino acid substitutions (X):

$[GGX]_7$, X = A, S, T or V;

$[GPGGX]_{43\text{-}63}$, X = A, S, Y or V (alternating X/X residues: A/A, S/Y, or V/S)

Figure 2.1. Flag protein consensus sequence.

The *N. clavipes* flagelliform silk Flag protein consensus sequence is presented with the three core motifs. The corresponding X amino acids of the glycine-rich motifs are noted. (Hayashi *et al.,* 1998 & 1999; Gatesy *et al*., 2001).

to fourteen repeats of the $[(GPGGX)_n\ (GGX)_n$ spacer$]_n$ ensemble (Hayashi *et al*. 1998 & 2001). 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 (Sponner *et al.,* 2005) and also in recombinant silk proteins in solution (Rammensee *et al.,* 2008; Heim *et al.,* 2010; Askarih *et al.,* 2010). However, several recombinant spider silk-like proteins have been shown to form fibers without the inclusion of the terminal sequences (Lazaris *et al*., 2002; Teulé *et al.,* 2007 & 2012; Brooks *et al*., 2005; An *et al.,* 2011).

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, $(A)_n/(GA)_n$, that form stacked crystalline β-sheet structures that align parallel to the axis of the fiber (Parke *et al.,* 1997; Dong *et al.,* 1991; Thiel *et al.,* 1994 & 1996; Riekel *et al.,* 1999; Holland, *et al.,* 2008). However, the Flag protein does not contain this typical polyalanine strength motif. Understanding the extensibility of the Flag protein is important, yet the source of its strength has remained a mystery.

Raman spectroscopy has been used to probe the structure of dragline spider silk (Shao *et al*., 1999; Vollrath *et al*., 2001, Stephens *et al*, 2005; Rousseau *et al*., 2004). This is achieved by fitting the amide I region of the spectrum to various protein structural components, *i.e.,* random coil, helical, β-sheet, and β-turn (Shao *et al*., 1999). Raman spectroscopy has the advantage over other structural analysis methods as it can easily be

performed on a single silk fiber without isotope labeling (Vollrath & Knight, 2001). Recently, Raman spectroscopy has been used to characterize gland contents and natural silk fibers from the common forms of silk produced by orb-weaver spiders (Rousseau *et al.,* 2009).

Raman analysis has shown that one of the unique characteristics of flagelliform silk is that the native spinning process does not induce a detectable structural transition in the Flag protein compared to the other spider silks that do experience a structural transition between their liquid protein dope and solid silk fiber states (Lefèvre *et al.,* 2011). Initially, this was best explained by the observed lower tensile strength of flagelliform silk that seemed to result from the lack of classical silk-like β-sheet forming sequences or crystalline forming domains in the Flag protein. However, amide I decomposition of the Raman spectra of flagelliform silk fibers from different orb weaver spiders revealed β-sheet structures that align parallel to the fiber axis in the native fiber(Lefèvre *et al.,* 2012). This may provide a mechanism for fiber tensile strength that would be similar to that of other silks. However, these data have not yet been confirmed by other biophysical analyses.

Structural studies utilizing X-ray diffraction (XRD) methods such as small-angle X-ray scattering (SAXS) and wide-angle X-ray scattering (WAXS) have been used to investigate spider silk protein structure (Grubb and Jelinski, 1997; Martel *et al.*, 2008; Benmore *et al.,* 2012). Dragline spider silk fibers from different species of orb-weavers

were found to be similar in structure (Benmore *et al*., 2012). Both Raman and XRD analyses contribute to the limited structure understanding of spider silk. Further structural studies will provide additional information as new methods are developed for the flagelliform silk.

In this study, four Flag-like protein variants were designed to investigate both the motif contributions to mechanical properties and the roles of the structural domains of the GGX and spacer motifs 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 $(GPGGX)_8$ motif module were designed into three of the encoding $(GGX)_7$ module containing constructs. Recombinant silk proteins were produced in *E. coli* by fermentation, purified, and analyzed. The recombinant silk proteins were solubilized in 1,1,1,3,3,3-hexafluro-2-propanol (HFIP) to generate liquid spin dopes for fiber formation by extrusion into an isopropanol coagulation bath. The dry fibers were processed by a 3x stretching in an 80% aqueous isopropanol bath and air dried. The fibers were tested mechanically and further analyzed using Raman spectroscopy and Wide Angle X-ray Scattering (WAXS) techniques to investigate the secondary structures of both the Flag spacer and $(GGX)_7$ motifs in their solid fiber state. Using a standard spinning method for all fibers, the resulting fiber data was compared to each other. Based on the Flag-like motif content and function in the recombinant proteins, the data was correlated with the similar motif

content of the native Flag protein. These studies help elucidate 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 $(GGX)_7$ structural motif module as the constant motif component designated as "G" in the new proteins (Fig.. 2.2). The Flag-spacer is designated as "F" and the $(GPGGX)_8$ module as "Y". Three dsDNA sequences were engineered encoding separately the $(GGX)_7$ – 5'-actagtggatccgcatatgggcccgggtggcgctggcggtagcggtggcgcgggtggttctggtggtgttggcggttctggcggcaccggtccggagatatctggatccaaactg 3', the [(GGX)7 spacer] – 5'-actagtggatccccatatgggcccgggaggtgccgggggaagtgggggagcaggagggtcaggtggggtcggcggcagcgggggcaccaccatcattgaggatctggacatcacaatcgacggcgccgacggccctatcacaattagcgaagaactcactattagcggagctggcgggtccggtccggagatatctggatccaagctt 3', and the $(GPGGX)_8$ (Teulé *et al.*, 2007) 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 *Spe*I and *Hind*III restriction sites, with the introduction of new flanking *BamH*I and *Nde*I, restriction sites by Bio S&T Inc. (Fig. 2.3; Montreal, Quebec). The new G and GF plasmid clones encoded the motif sequences: GGAGGSGGAGGSGGVGGSGGT for 'G' and

TIIEDLDITIDGADGPITISEELTISGAGGS for ‘F’ (Fig. 2.3). 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 $(GPGGSGPGGY)_4$ sequence was produced from the same core sequence that was designed and supplied as the “Y2” clone by Dr. M. B. Hinman. The Y2 dsDNA core sequence was released from the pBluescript SK+ plasmid utilizing the 5’ *Xma*I and the 3’ *BspE*I 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.

Structural Motif	Module Name	Amino Acid Sequence
GGX	G	GGAGGSGGAGGSGGVGGSGGT
GPGGX	Y	$(GPGGSGPGGY)_4$
spacer	F	TIIEDLDITIDGADGPITISEELTISGAGGS

Figure 2.2. Motif sequences used for protein module design.

Based on the *N. clavipes* Flag protein, the amino acid sequences are shown for the structural motifs and their module designations for this research.

a)

```
SpeI  BamHI   NdeI   XmaI
\     \       \      \
actagtggatccgcatatgggcccgggtggcgctggcggtagcggtggcgcgggtggttc
tgatcacctaggcgtatacccgggcccaccgcgaccgccatcgccaccgcgcccaccaag
 L  V  D  P  H  M  G  P  G  G  A  G  G  S  G  G  A  G  G  S
                              BspEI        BamHI
                              \            \
tggtggtgttggcggttctggcggcaccggtccggagatatctggatccaaactg
accaccacaaccgccaagaccgccgtggccaggcctctatagacctaggtttgac
 G  G  V  G  G  S  G  G  T  G  P  E  I  S  G  S  K  L
```

b)

```
SpeI  BamHI   NdeI   XmaI
\     \       \      \
actagtggatccccatatgggcccgggaggtgccgggggaagtgggggagcaggagggtc
tgatcacctaggggtatacccgggccctccacggccccccttcacccccctcgtcctcccag
 L  V  D  P  H  M  G  P  G  G  A  G  G  S  G  G  A  G  G  S

aggtggggtcggcggcagcgggggcaccaccatcattgaggatctggacatcacaatcga
tccaccccagccgccgtcgccccccgtggtggtagtaactcctagacctgtagtgttagct
 G  G  V  G  G  S  G  G  T  T  I  I  E  D  L  D  I  T  I  D

cggcgccgacggccctatcacaattagcgaagaactcactattagcggagctggcgggtc
gccgcggctgccgggatagtgttaatcgcttcttgagtgataatcgcctcgaccgcccag
 G  A  D  G  P  I  T  I  S  E  E  L  T  I  S  G  A  G  G  S

   BspEI        BamHI HindIII
   \            \     \
cggtccggagatatctggatccaagctt
gccaggcctctatagacctaggttcgaa
 G  P  E  I  S  G  S  K  L
```

Figure 2.3. 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 Biology Workbench (Mangalam, 1996; Subramaniam, 1998). a) The G module and b) the GF module sequence designs.

The structural motif combinations that form the module basis for each of the new proteins (Fig. 2.4) were constructed through the manipulation of their encoding dsDNA. The restriction enzymes *BspE*I or *Xma*I, in conjunction with *AlwN*I, were used utilizing an established compatible non-regenerable restriction site method (Prince *et al*., 1995; Lewis & Hinman, 1996; Teulé *et al*., 2007 & 2009; An *et al.,* 2011). 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 32x, 12x, 12x, and 8x to get final DNA constructs with sizes of 2.3, 2.4, 2.1, and 2.4 kbp respectively, using the compatible non-regenerable method (Prince *et al*., 1995; Lewis & Hinman, 1996). 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 $(GGX)_7$ 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 (Teulé *et al.,* 2012). Thus, all of these newly engineered Flag-like protein variants were specifically designed to address the functional role of both the $(GGX)_7$ and the spacer structural domains (*i.e.* the G and F modules) found in the native silk protein.

Protein Designation	Primary sequence
G	$[GGAGGSGGAGGSGGVGGSGGT]_{32}$
GY	$[GGAGGSGGAGGSGGVGGSGGT(GPGGSGPGGY)_4]_{12}$
GF	$[GGAGGSGGAGGSGGVGGSGGTTIIEDLDITIDGADGPITISEELTISGAGGS]_{12}$
GFY	$[(GGAGGS)_2GGVGGSGGTTIIEDLDITIDGADGPITISEELTISGAGGS(GPGGSGPGGY)_4]_8$

Figure 2.4. 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.

Recombinant Silk Gene Expression and Protein Purification

The four newly constructed sequences were separately inserted into the plasmid vector pET19b-k (kan^R) 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 modified by Dr. M. Hinman to provide kanamycin resistance (see Teulé *et al.*, 2007 for details on the construction of this vector).

The proteins were expressed, harvested, and purified using methods previously reported (Teulé *et al*., 2009; An *et al.*, 2011). All recovered cell extracts were heat treated to remove most native *E. coli* proteins (Scheller *et al*., 2001; Teulé *et al*., 2007). The His-tagged proteins were then purified using automated 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 (Sambrook *et al.,* 1989), frozen, and lyophilized. The dry recombinant silk (rSilk) proteins were utilized for spin dopes and characterized by SDS-PAGE and Western blot analyses using a 6x His mAb-HRP conjugate (Clontech) anti-His tag antibody as described previously (Teulé *et al*., 2007).

Preparation of the Spin Dopes

Fifteen % (w/v) rSilk-HFIP protein spin dopes were made by dissolving each of the purified lyophilized proteins into 1,1,1,3,3,3-hexafluro-2-propanol (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 (Nexia Biotechnologies Inc., Montreal, Canada, unpublished data 2005; Brooks *et al,* 2005). 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 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 fibers were spun as previously reported (Teulé *et al.,* 2007 & 2009; Brooks *et al*., 2008; An *et al*., 2011) 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 (Teulé *et al*., 2007, 2009, & 2012) were investigated. 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 (Brooks *et al.,* 2008; Elices *et al.*, 2011; An *et al*., 2011).

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) isopropanol bath with 20% (v/v) water 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 into 3cm lengths and mounted on testing cards with an initial fiber length (L_0) of 19.05 mm. The fibers were visualized under a Nikon Eclipse E200 microscope (Nikon Metrology Inc., Brighton, MI) at a 40x 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 then used 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 (25º 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 exported into a Microsoft Excel (Microsoft Corporation, Redmond, WA) spreadsheet and used to calculate engineering stress ($\boldsymbol{\sigma_e}$) and strain ($\boldsymbol{\varepsilon_e}$) using the following formulas: $\boldsymbol{\sigma_e}=F/A$ and $\boldsymbol{\varepsilon_e}=L\text{-}L_0/L_0$; 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 ($\boldsymbol{\varepsilon_e}$ x 100) also called % 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 strength and extensibility were calculated in Excel using a confidence level of $\alpha = 0.05$.

Structural Analysis: Raman Spectromicroscopy

Full spectrum Raman analysis: Fibers were affixed to quartz slides and subjected to Raman spectroscopic analysis using a New Dimension Raman microscope (Snowy Range Instruments). Spectra were collected with a 1064 nm laser focused through a long working distance 50 mm objective lens. A polystyrene control was used to calibrate the instrument prior to each data collection. The following five peaks were used for calibration from the polystyrene control: 620, 795.8, 1001.4, 1450.5, and 1602.3 cm^{-1}. Raman spectra were acquired for 10 seconds at ambient room temperatures of 25°C with 12-22% relative humidity. Spectra acquisition was analyzed for fluorescence interference and sample degradation. The background vibrational noise was subtracted using a Savitsky-Golay smoothing method (polynomial 3 points = 17). The peak intensities were adjusted for best-fit of collected spectra. The spectra were manipulated utilizing GRAMS IQ software (Adept Scientific) for the spectra overlays.

Amide I region Raman analysis: Raman spectra of the proteins were collected from as-spun fibers using a lab-built confocal Raman microscopy instrument for the analysis of the amide I region (1630-1700 cm^{-1}). Raman spectra were acquired in 180-degree geometry using a PSB-40 Raman system with a laser wavelength of 532 nm. The power was kept under 1 mW using neutral density filters. The laser was focused onto the sample using a 100x long working distance Mitutoyo objective lens. The signal was

discriminated from the laser excitation light using a Kaiser Supernotch filter. Cyclohexane was used for Raman system instrument calibration. Raman spectra were acquired, averaged and the background vibrational noise was subtracted. The data was collected using a Shamrock 303 Spectrograph and an Andor CCD detector. The peak intensities were adjusted for best-fit of collected spectra and analyzed for structural content with the following 5 component correlations: 1640 cm^{-1} = random coil (rc); 1655 cm^{-1} = 3_1- helix; 1670 cm^{-1} = β-sheet; 1785 cm^{-1} = β-turn (Lefèvre *et al*., 2007).

Structural Analysis: X-ray Diffraction.

Wide angle X-ray scattering (WAXS) diffraction experiments were performed in sector 14 BM-C/BioCARS of the Advanced Photon Source (APS) at Argonne National Laboratory (Argonne, IL, USA; Graber *et al.,* 2011). The WAXS data were recorded with a large area 9-chip ccd detector (ADSC Quatum-315) placed 300 nm behind the sample. The monochromatic X-ray beam was focused to 150 x 200 microns at an incident energy of 12.67 keV (0.978 Angstrom). 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 sample-beam stop 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. Although the nature of the elasticity of this silk is attributed to the presence and abundance of the $(GPGGX)_n$ 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 Flag-like repetitive DNA sequences were successfully cloned into *E. coli* producing the corresponding His-tagged proteins with expected molecular weights ranging from 54-66 kDa (Fig. 2.5, Table 2.1). The protein production levels varied between the various clone types: G 50 ug/L; GY 500 µg/L; GF 5 µg/L; and GFY 50 µg/L (Fig. 2.5). 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 (data not shown) for the G protein but it primarily produced dimers (Fig. 2.5).

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. 2.2): the G module or $(GGX)_7$, the F module or 'spacer', and the Y module which is one of the putative Flag elastic $(GPGGX)_8$ motifs, $(GPGGY\ GPGGS)_4$ (Hayashi *et al.*, 1998). 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 (Hayashi *et al*., 1998 & 2001). Structural work on the non-repetitive flanking regions has suggested an *in vivo* aggregation role in fiber formation (Sponner *et al.*, 2005). Recently, it was reported that the flanking regions of recombinant proteins in solution drive fiber formation (Rammensee *et al.,* 2008; Heim *et al*., 2010). However, many studies have shown that these flanking regions are not necessary in the recombinant silk proteins to allow fiber formation (Lazaris *et al.,* 2002; Brooks *et al*., 2005; Teulé *et al*., 2007, 2009, 2012; An *et al*., 2011). For the sake of targeted structural motif study, these 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 1; Lazaris *et al.*, 2002). 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 that 20% (v/v) water was the best processing treatment for all four different fiber types (Fig. 2.6). 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. 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 (Seidel *et al.,* 2000; Teulé *et al*., 2012). 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.

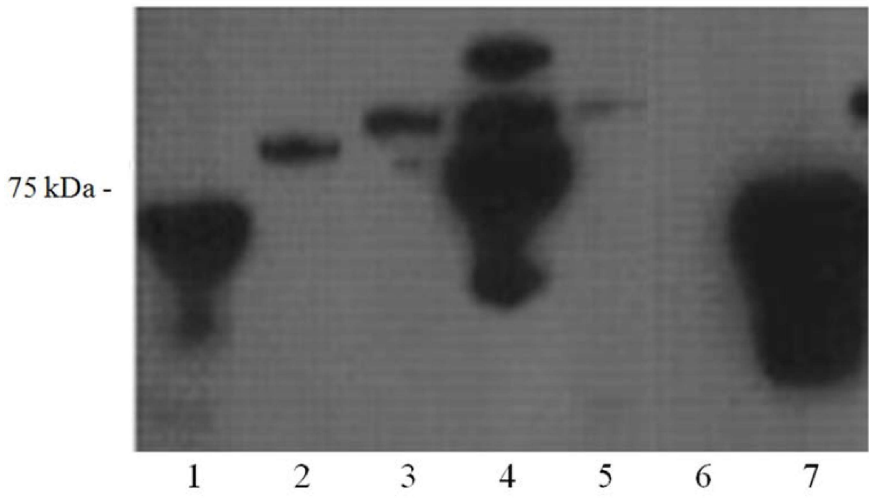


Figure 2.5. Recombinant protein analysis by Western blot.

Western blot analysis of the four soluble proteins confirms recombinant protein production. A 6x His mAb-HRP conjugate was used to identify recombinant proteins. Lanes 1-3 and 5 show protein extracts from clones GY, GFY, G, and GF. Lanes 6 and 7 contain the negative and positive controls respectively. The Plus Protein™ Dual Color Standards, (Bio-Rad Life Science Research, Hercules, CA, USA) molecular weight marker is in lane 4 with the thickest band at 75 kDa.

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
$(GGX)_7$	32x	2.31	54	G
$(GGX)_7$ + $(GPGGX)_8$	12x	2.37	60	GY
$(GGX)_7$ +(spacer)	12x	2.05	59	GF
$(GGX)_7$ +(spacer)+ $(GPGGX)_8$	8x	2.37	66	GFY

Table 2.1. 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.

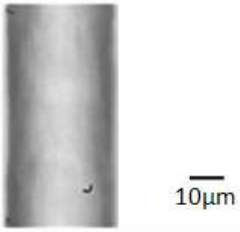
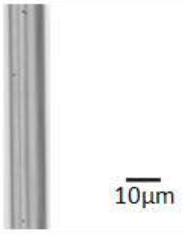
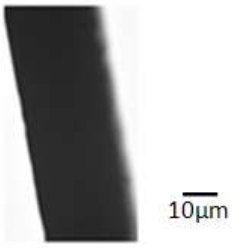
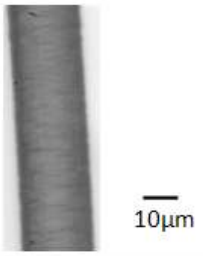
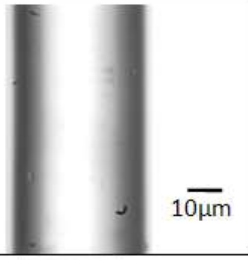
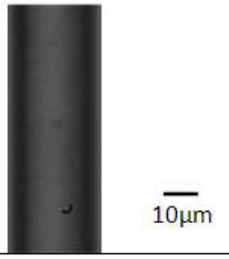
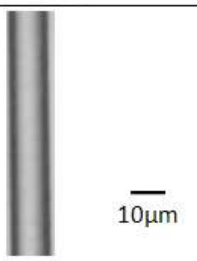
Protein	As-Spun Fibers 100x	Post Spin Drawn Fibers 100x
G		
GY		
GF		
GFY		

Figure 2.6. Fibers spun from the four protein variants.

As spun (AS) fibers and Post-spin drawn (PSD) fibers observed at100x light microscopic magnification.

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 (Rammensee *et al*., 2008; Heim *et al*., 2010; Askarieh *et al.,* 2010). 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 (Teulé *et al*., 2007, & 2012; Brooks *et al*., 2008; An *et al.,* 2011).

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 2.2). This is not unexpected as similar observations were reported for the mechanical performance of native (Blackledge *et al.*, 2005) and synthetic (An *et al*., 2011; Teulé *et al.,* 2012) silk fibers. 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 (Teulé *et al.,* 2012). This variation generated by processing is further supported by recent work showing that protein processing influences native spider silk performance (Blamires *et al.,* 2012).

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). 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. 2.7). 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 2.2 & Fig. 2.7) thus corresponding with variations in the silk protein primary structure.

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

Table 2.2. 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 absence of data due to fiber brittleness is indicated by an asterisk (*). The PSD fiber superscripts for % Extension and Strength indicate significant differences between fiber types where $p<0.05$, 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.

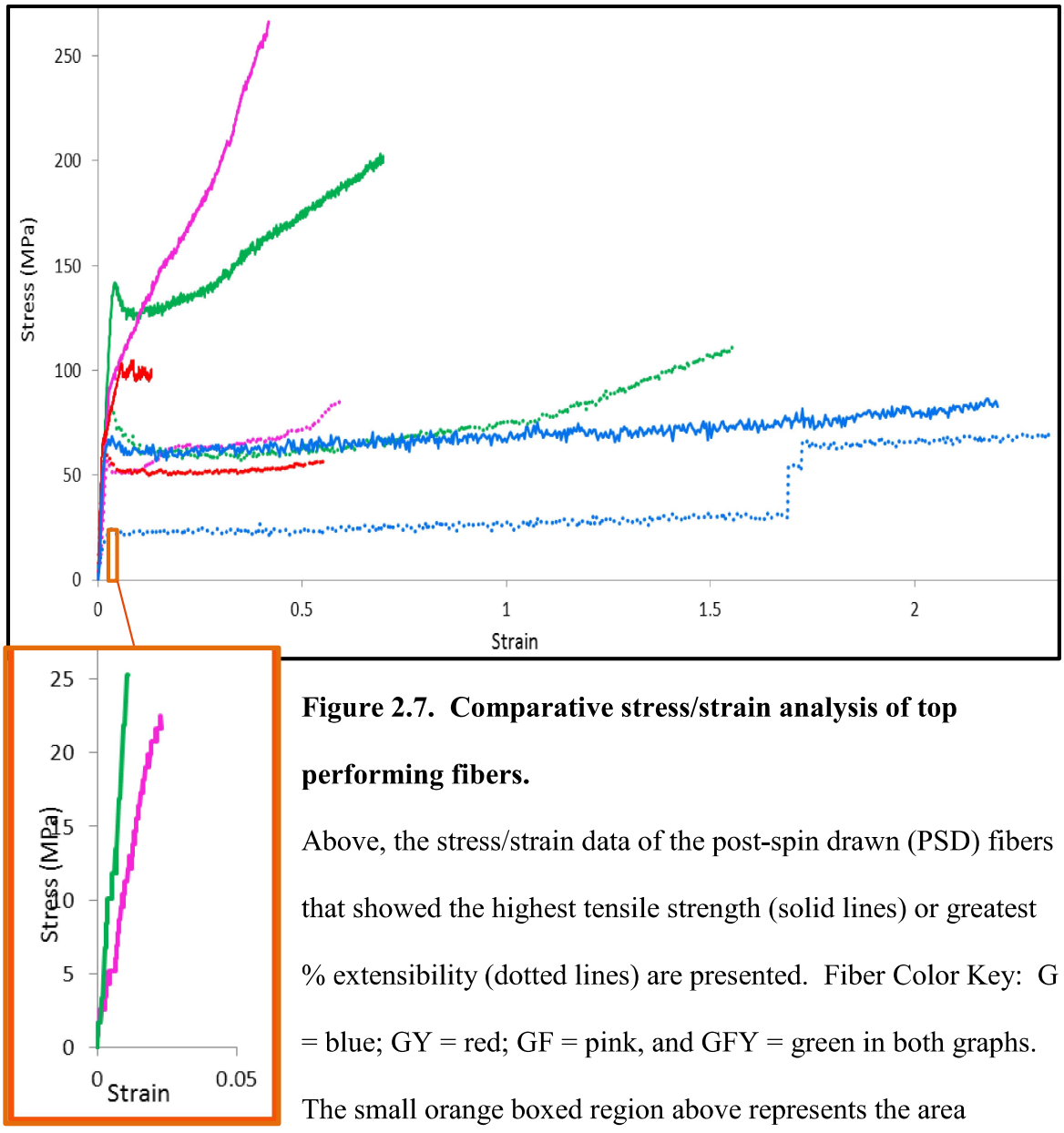


Figure 2.7. Comparative stress/strain analysis of top performing fibers.

Above, 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 above represents the area occupied by the data to the left. This graph shows an enlarged and horizontally stretched representation to show the stress/strain analyses of the top performing as spun (AS) fibers for GF (1 of 5) and GFY (1 of 3). The 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.

Interestingly, the presence of the $(GGX)_7$ motif 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. 2.7). 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 2.2). Furthermore, the addition of the $(GPGGX)_8$ 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. 2.7, Table 2.2). The GY protein, with its β-turn/β-spiral forming $(GPGGX)_8$ 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 $(GPGGX)_8$ 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 2.2 & Fig. 2.7). 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 orchestrate 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.

Structural Analysis of the Synthetic Fibers.

Raman spectroscopy and wide angle X-ray diffraction scattering (WAXS) structural analyses were performed on the fibers to investigate the structural origin of the mechanical properties displayed by the different fiber types. These methods are well adapted for the study of smaller silk samples (Shao *et al*., 1999; Rousseau *et al.,* 2004 & 2009; Lefèvre *et al.,* 2009 & 2012; Benmore *et al*., 2012).

The Raman spectra collected from the proteins of the G, GY, GF, and GFY solid fibers (Fig. 2.8) display similarities to that of published native flagelliform silk fibers (Rousseau, *et al.,* 2009; Lefèvre *et al.,* 2011).

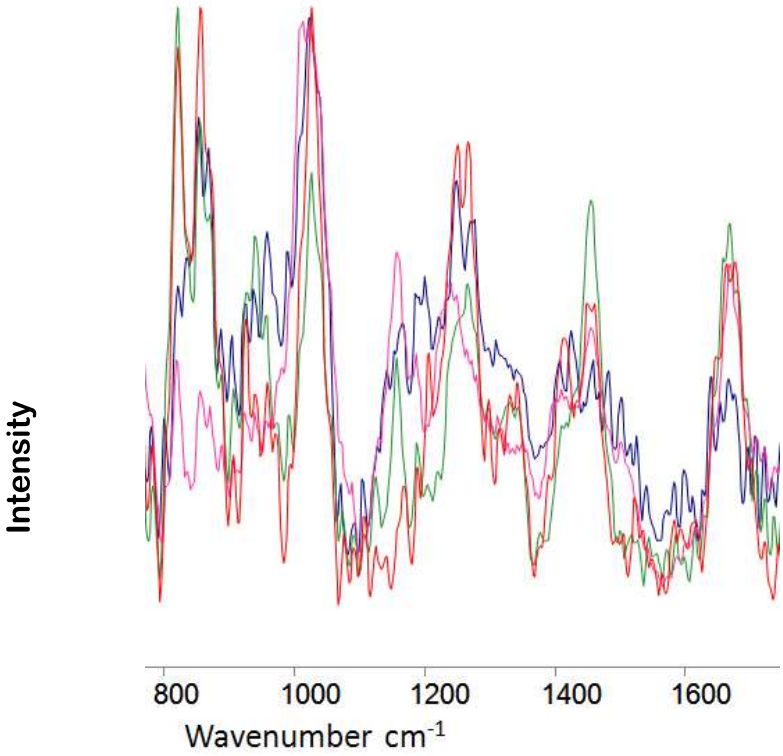


Figure 2.8. Raman spectra of the G, GY, GF, & GFY protein fibers.

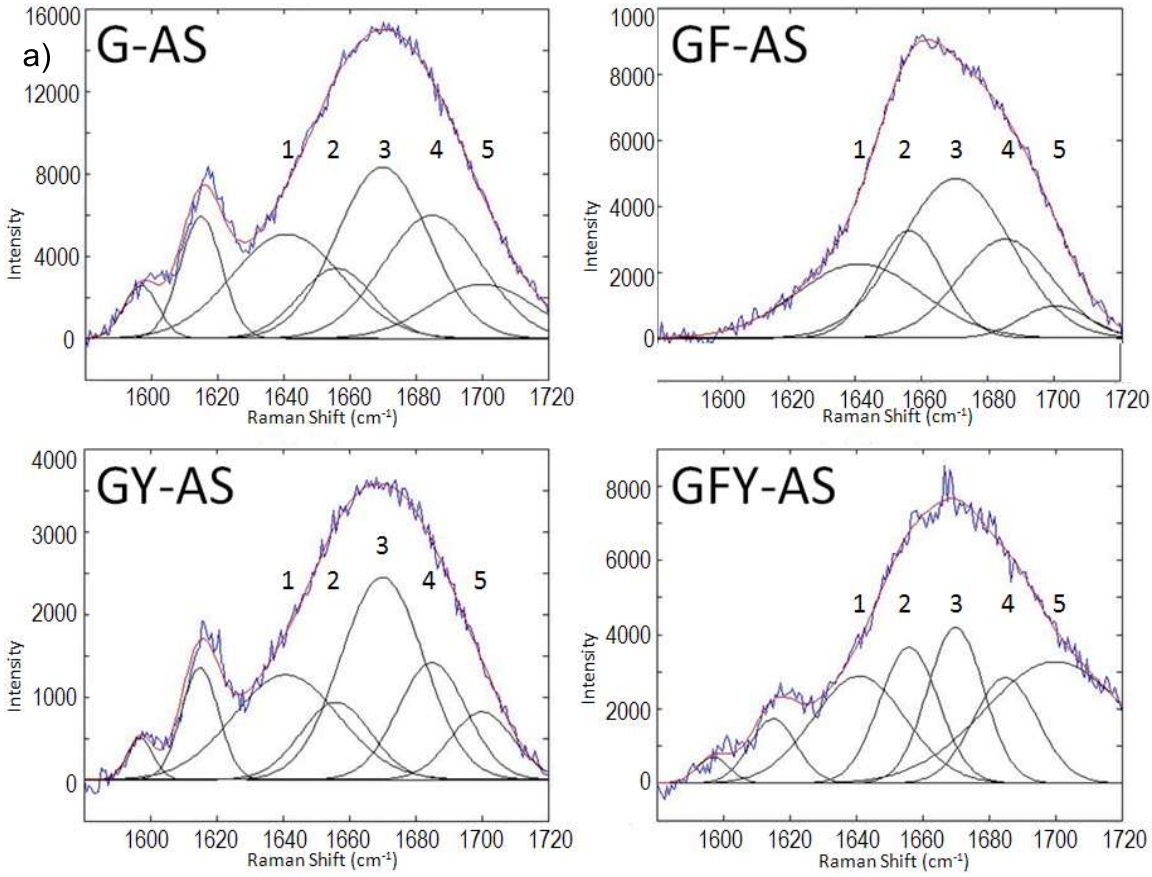
These Raman spectra were collected using a 1064 nm laser from as spun (AS) or post spin drawn (PSD) fibers spun from the Flag-like variant proteins. The fiber spectra collected are shown as: the G PSD fiber, blue; the GY AS fiber, red; the GF PSD fiber, pink; and the GFY AS fiber, green. The intensity of the various spectra differ, therefore the overlay analysis was normalized to provide general spectra comparisons thus is not quantitative.

The Raman spectra for each of the AS synthetic fibers, except for the AS GF fibers, showed that the synthetic proteins had a broad amide I peak maximum centered at 1669-1670 cm^{-1} (data not shown). This result agrees with the reported typical amide I band positions of 1670 cm^{-1} for native dragline silk proteins after spinning and with the unchanged 1668-1669 cm^{-1} position for native Flag proteins (Lefèvre *et al.,* 2011). The amide I signal for the GF protein shifted downward to 1662cm^{-1}, potentially suggesting a higher propensity for helical structures in this recombinant protein (Tuzikov *et al.,* 2003; Zhou *et al*., 2006). These latter data support the proposed polyglycine helical (PGII) structure of the $(GGX)_n$ motif in the protein fibers (Holland *et al.,* 2008). The GF fiber analysis also suggests a preferred helical conformation for the spacer motif when juxtaposed to the $(GGX)_n$ motif, supporting the functional role of the spacer as a module organizer as previously hypothesized (Hayashi *et al.,* 2001).

In the Raman spectra of the protein, the conformation-sensitive amide I region (C=0 stretch) located at 1630-1700 cm^{-1}, was further analyzed (Fig. 2.9a) specifically to investigate the structural differences between the proteins within the AS fibers using established amide I peak deconvolution methods (Lefèvre *et al.,* 2007). Differences in silk fiber protein structure can be observed in this region of the spectra using a five-component decomposition that correlates each component to four secondary structures in the solid silk proteins (Fig. 2.9b). Subtle protein structural differences are observed as the components from the G, GY, GF, and GFY fibers indicate variation in the proportional combinations of the four different secondary structures (Fig. 2.9a). As seen

for the G, GY and GF amide I analysis (Fig. 2.9a), the curves of the components corresponding to the four secondary structures identified are exceptionally broad thus indicating that these structures are likely not well defined.

a)



b)

Component	1	2	3	4	5
Location (cm^{-1})	1640	1655	1670	1685	1700
Structure	Random coil	3_1-helix	β-sheet	β-turn	β-turn

Figure 2.9. Decomposition of the amide I region of as-spun fiber proteins.

Raman spectra were collected on as-spun (AS) fibers of each variant protein type: G, GY, GF, and GFY. a) A five component deconvolution of the amide I region (1630 – 1700 cm^{-1}) using the Raman spectra of the protein of each fiber was performed. b) The five structural components are listed with their location in the spectra.

The existence of the β-sheet structure in the Flag protein is controversial and yet, we see evidence of this structure in the data generated using these methods in our structural studies of the variant proteins. Using the five component decomposition, all four of the AS synthetic fibers were dominated by component 3, at 1670 cm^{-1}, indicative of β-sheet type structures (Zhao *et al*., 2005). However, other structures not yet identified are believed to contribute to this specific component manifestation (Lefèvre *et al.,* 2007) as the Flag protein does not contain the known β-sheet forming polyalanine regions that are so dominant in dragline silk proteins.

The WAXS patterns of the GY, GF and GFY fibers are all characterized by the presence of a diffuse ring centered at 4.4Å (Fig. 2.10). These isotropic patterns are consistent with the presence of an amorphous phase composed of largely disoriented structures such as helixes or of largely disoriented β-sheet nanocrystals formed upon extrusion as indicated previously in the case of regenerated dragline silk and synthetic AS silk-like fibers (Seidel *et al*., 2000; Teulé *et al*., 2012). For both the G and the GY AS sample, the formation of imperfect polyglycine-type nanosheets (PGI) may contribute to this broader 1670 cm^{-1} component as none of these proteins possess any other classical β-sheet motifs in their sequences. The formation of β-sheet in flagelliform silk was reported in quantities high enough to crosslink the molecular network but not reinforce the network (Gosline *et al*., 1994). Other glycine-rich spider silk peptides formed β-sheets when cast into films (Fukushima *et al.*, 1998). 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. 2.10). The reflection must result from the presence of the spacer in these proteins. Raman data collected on these two types of fibers does suggest that these proteins accommodate β-sheet structures with a higher proportion of 3_1 helices compared to the composition of the G and the GY fiber samples. These specific double reflection WAXS patterns are characteristic of stacked β-sheet type structures (4.7 Å and 8-11 Å) observed in amyloid fibers (Maji *et al.*, 2009; Hauser *et al.,* 2011). 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).

This specific helical enhancement in these AS GF and GFY fibers (component 2; Fig. 2.9) not only made them the least brittle of all but also potentially set the stage for the superior tensile strength performances that they achieved. Similarly, when comparing the GY and the GFY fiber decomposition, a proportional increase of component 2 is visible in the GFY fiber analysis, in addition to an increase in component 5 (Fig. 2.9a). This suggests that, in addition to the facilitation of the 3_1-helix formation, the spacer motif may also function to enhance the folding of the $(GPGGX)_n$ motif. As a result of the proper formation of β-turn-like structures in the GFY protein, the GFY fibers have increased extensibility. The apparent structural coordination by the spacer motif is best seen in the GFY fiber spectra as the inclusion of this motif in the silk protein produces

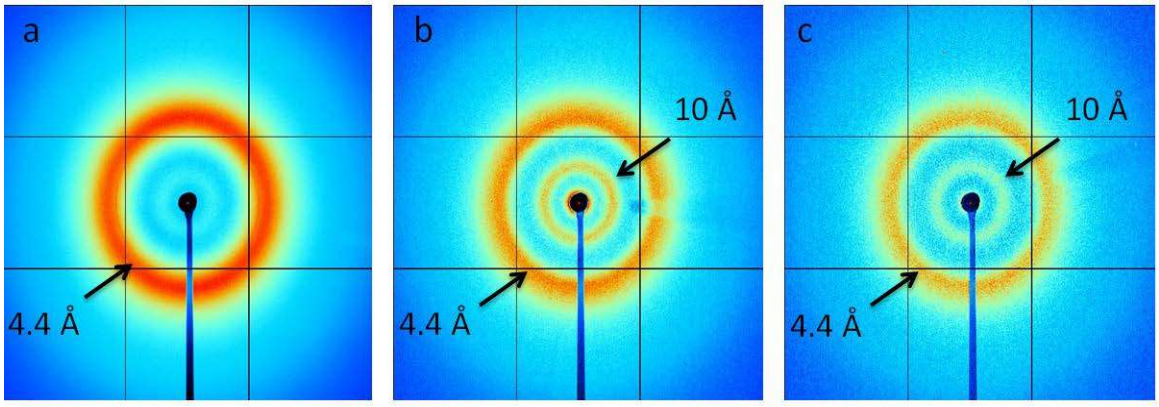


Figure 2.10. Wide Angle X-Ray Scattering analysis of the as-spun fibers.

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. The data were collected with a detector distance of 300 nm and an exposure time of 60 seconds. Black arrows indicate rings centered at 4.4 Å or 10 Å that are visible in the different WAXS patterns.

spectra with narrower peaks, which indicate a more defined structure compared to all other fiber spectra in this group.

Interestingly, in the amide I decomposition of GY AS fibers, component 3 (Fig. 2.9a) is observed but seems to diminish when processed in 80/20% IPA and water (Fig. 2.11). This possible β-sheet structure seems to be getting pulled apart with stretching and increases the PGII type structure, yet this processing contributes to an increase in the mechanical performance of the fiber (Table 2.2). This is the opposite of what is reported in dragline silk, where the β-sheet structures were reported to increase with processing that actually formed more crystals and strengthened the fiber (An *et al.,* 2011).

In our study we determined that the amide I five-component decomposition to be a more informative method that the previous methods (Shao *et al*., 1999; Zhou *et al.,* 2005) used due to the many spider silks that have been analyzed using the more refined method (Lefèvre *et al*., 2007 & 2011; Rousseau *et al*., 2009). It has just recently been reported that indeed, small quantities of β-sheet were observed by Raman spectra in native flagelliform spider silk fiber (Lefèvre *et al*., 2012). However, in that same study, the five-component deconvolution method that has been so widely used was abandoned by Lefèvre (2012) and a different decomposition pattern containing only four components was developed specifically for flagelliform silk. The amide I deconvolution 1670 cm^{-1} component observed in these fibers cannot be solely attributed to β-sheets and must originate from some other stacked or packed structure perhaps stacked PGII helices

that also emits at signal ~ 1670 cm^{-1} 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 β-sheet. The difficulty in obtaining samples of native flagelliform silk inhibits supportive structural analysis such as NMR to clearly determine which method is more accurate.

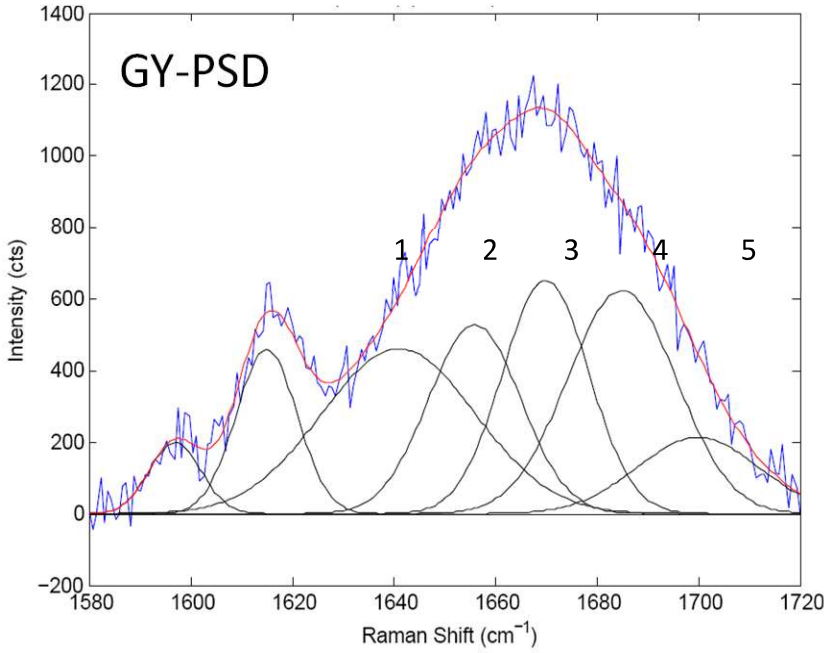


Figure 2.11. Amide I region decomposition of the post-spin drawn GY protein. The processing of the GY fiber significantly increases the proportion of component 2 helices (1655 cm^{-1}) and component 4 β-turns (1685 cm^{-1}) while diminishing the proportion of the component 3 peak that is known primarily for β-sheet structure (1670 cm^{-1}) .

The discussion of whether or not the Flag protein produces β-sheets or the quantification of the β-sheet structures that it does produce, will likely continue until new methods become available to biophysically confirm the data that is being generated. A structural protein model of the various spider silk proteins is needed to better understand the Flag protein and how the β-sheet region form and function. The absence of the polyalanine β-sheet forming regions, common to dragline silk, in the native Flag protein and now also in these Flag-like variant protein fibers, leaves a conundrum for the production of β-sheet structures. Formation of β-sheets in the absence of the polyalanine β-sheet forming regions, would be unusual but may indicate a previously unknown mechanism. The structural analysis data generates information but the structure cannot be inconclusively determined from this data.

CONCLUSION

The 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 different mechanical properties and structure. 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. The toughest, GFY, with the greatest strength and moderate extensibility was produced from the one protein that contained 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. 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 (Lefèvre *et al*., 2012). This data provides evidence for the mechanical contribution of both the Flag GGX and spacer motifs and indicates a structural association attributed to the spacer that could be β-sheet or helical. Further studies will be necessary to elucidate the specific structures as technology and methods advance.

REFERENCES

An, B.; Hinman, M.; Holland, G.; Yarger, J.; Lewis, R. *Biomacromolecules.* **2011**. 12:2375-81.

Askarieh, G.; Hedhammar, M.; Nordling, K.; Saenz, A.; Casals, C.; Rising, A.; Johansson, J.; Knight, S. *Nature.* **2010**. 465:236-239.

Becker, N.; Oroudjev, E.; Mutz, S.; Cleveland, J.; Hansma, P.; Hayashi, C.; Makarov, D.; Hansma, H. *Nature Materials*. **2003**. 2:278-283.

Benmore, C.; Izdebski, T.; Yarger, J. *Physical Review Letters*. **2012**. 108(17):8102-1-4.

Blackledge, T.; Cardullo, R.; Hayashi, C. *Invertebrate Biology.* **2005**. 124(2)165-173.

Blamires, S.; Wu, C.; Blackledge, T.; Tso, I. *J. Roy. Soc. Int.* in press. Epub ahead of print retrieved May 24, **2012**. 9(75) 2479-2487.

Bonthrone, K.; Vollrath, F.; Hunter, B.; Sanders, J. *Proc. Roy. Soc. London*. **1992**. 248:141-144.

Brooks, A.; Steinkraus, H.; Nelson, S.; Lewis, R. *Biomacromolecules*. **2005**. 6(6):3095-99.

Brooks, A.; Stricker, S.; Joshi, S.; Kamerzell, T.; Middaugh, C.; Lewis, R. *Biomacromolecules*. **2008**. 9 (6):1506-1510.

Choresh, O.; Bayamagnai, B.; Lewis, R. *Biomacromolecules.* **2009**. 10(10):2852-2856.

Creager, M.; Jenkins, J.; Yeaman, L.; Brooks, A.; Jones, J.; Lewis, R.; Holland, G.; Yarger, J. *Biomacromolecules.* **2010**. 11(8):2039-2043.

Debelle, L.; Alix, A.; Baco, M.; Huvenne, J.; Berjot, M; Sombret, B.; Legrand, P. *J. Biol. Chem.* **1995**. 270(44):26099-26103.

Denny, M.W. *Cambridge University Press*, Cambridge. **1980**. 247-272.

Dong, Z.; Lewis, R.; Middaugh, C. *Arch. Biochem. Biophys.* **1991**. 284:53-57.

Elices, M.; Guinea, G.; Plaza, G.; Karatzas C.; Riekel, C.; Agulló-Reueda, F.; Daza, R.; Pérez-Rigueiro, J. *Macromolecules.* **2011**. 44:1166-1176.

Foelix, R. *Biology of Spiders.* 2nd edition. New York: Oxford University Press, Inc. **1996**. 110-149.

Fukushima, Y. *Biopolymers.* **1998**. 45:269-279.

Gatesy, J.; Hayashi, C.; Motriuk, D.; Woods, J.; Lewis, R. *Science.* **2001**. 291:2603-2605.

Gosline, J.; DeMont, M.; Denny, M. *Endeavour.* **1986**. 10:37-43.

Gosline, J.; Pollak, C.; Guerette, P.; Cheng, A.; DeMont, M.; Denney, M. *ACS Symp. Ser.* 544 (Silk Polymers). **1994**. 328-341.

Gosline, J.; Guerette, P.; Ortlepp, C.; Savage, K. *J. Exp. Biol.* **1999**. 202:3895-3303.

Graber, 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.; Srajer, V.; Viccaro, P.; Westferro, F.; Anfinrud, P.; Moffat, K. *J. Synchrotron Radiat.*. **2011.** 18:658-670.

Grubb, D.; and Jelinski, L. *Macromolecules.* **1997**. 30:2860-2867.

Hauser, C.; Deng, R.; Mishra, A.; Loo, Y.; Khoe, U.; Zhuang, F.; Cheong, D.; Accardo, A.; Sullivan, M.; Riekel, C.; Ying, J.; Hauser, U. *PNAS*. **2011**. 108(4):1361-1366.

Hayashi, C.; Lewis, R. *J. Mol. Biol.* **1998**. 275:773-784.

Hayashi, C.; Shipley, N.; Lewis, R. *Int. J. Biol. Macromol.* **1999**. 24(2-3):271-275.

Hayashi, C.; Lewis, R. *Bioassays*. **2001.** 23:750-756.

Heim, M.; Ackerschott, C.; Scheibel, T. *J. Struct. Biol.* **2010.** 170:420-425.

Hinman, M. & Lewis, R. *J. Biol. Chem.*. **1992**. 267(27):19320-19324.

Holland, G.; Creager, M.; Jenkins, J.; Lewis, R.; Yarger, J. *Am. Chem. Soc*. **2008.** 130(30):9871-9877.

Holland, G.; Jenkins, J.; Creager, M.; Lewis, R.; Yarger, J. *R. Soc. Chem*. **2008**. 5568-5570.

Jenkins, J.; Creager, M.; Butler, E.; Lewis, R.; Yarger, J.; Holland, G.; *Chem. Commun.* **2010**. 46:6714-6716.

Lazaris, A.; Arcidiacono, S.; Huang, Y.; Zhou, J.; Duguay, F.; Chretien, N.; Welsh, E.; Soares, J.; Karatza, C. *Science.* **2002**. 295:472-476.

Lefèvre, T.; Rousseau, M.; Pézolet, M. *Biophys. J.* **2007**. 92:2885-2895.

Lefèvre, T.; Paquet-Mercier, F.; Lesage, S.; Rousseau, M.; Bedard, S.; Pézolet, M. *Vib. Spectrosc.* **2009**. 51:136-141.

Lefèvre, T.; Boudreault, S.; Cloutier, C.; Pézolet, M. *J. Mol. Biol*. **2011**. 405:238-253.

Lefèvre, T.; Paquet-Mercier, F.; Rioux-Dubé, J.; Pézolet, M. *Biopolymers.* **2011**. 97(6):322-336.

Lefèvre, T. & Pézolet, M. *Soft Matter*. **2012**. 8:6350-6357.

Lewis, R.; Hinman, M.; Kothakota, S.; Fournier, M. *Protein Expres. Purif.* **1996**. 7(4):400-406.

Liu, Y.; Shao, Z.; Vollrath, F. *R. Soc.Chem. Comm.* **2005**. 2489-2491.

Maji, S.; Wang, L.; Greenwald, J.; Riek, R. *FEBS Letters.* **2009**. 2610-2617.

Mangalam, H. UC Irvine. *SDSC Biology Workbench TACG*. **1996**. Online. <http://workbench.sdsc.edu>

Nexia Biotechnologies Inc., Montreal, Canada. Unpublished data. **2005**.

Parkhe, A.; Seeley, S.; Gardener, K.; Thompson, L.; Lewis, R. *J. Mol Recongnit*. **1997**. 10:1-6.

Prince, J.; McGrath, K.; DiGirolamo, C.; Kaplan, D. *Biochemistry*. **1995**. 34:10879-10885.

Rammensee, S.; Slotta, U.; Scheibel, T.; Bausch, A. *PNAS*. **2008**. 105:6590-6595.

Riekel, C.; Branden, C.; Craig, C.; Ferrero, C.; Heidelbach, F.; Muller, M. *Int. J. Biol. Macromol.* **1999**. 24:179-186.

Rousseau, M.; Lefèvre, T.; Beaulieu, L.; Asakura, T.; Pézolet, M. *Biomacromolecules.* **2004**. 5:2247-2257.

Rousseau, M.; Lefèvre, T.; Pézolet, M. *Biomacromolecules.* **2009**. 10:1945-2953.

Sambrook, J.; Fritsch, E.; Maniatis, T. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. **1989**.

Savage, K. & Gosline, J. *J. Exp. Biol.* **2008**. 211:1937-1947.

Scheller, J.; Guhrs, K-H.; Grosse, F.; Conrad, U. *Nature Biotech.* **2001**. 19:573.577

Seidel, A.; Liivak, O.; Jelinski, L. *Macromolecules.* **1998**. 31:6733-6736.

Seidel, A.; Liivak, O.; Calve, S.; Adaska, J.; Ji, G.; Yang, Z.; Grubb, D.; Zax, D.; Jelinski, L. *Macromolecule*s. **2000**. 33(3):775–780.

Shao, Z.; Vollrath, F.; Sirichaisit, J.; Young, R. *Polymer.* **1999**. 40:2493-2500.

Simmons, A.; Michal, C.; Jelinski, L. *Science*. **1996**. 271(5245):84-87.

Sponner, A.; Vater, W.; Rommerskirch, W.; Vollrath, F.; Unger, E.; Grosse, F.; Weisshart, K. *Biochem. Biophys. Res. Commun*. **2005**. 338(2):897-902.

Stauffer, S.; Coguill, S.; Lewis R. *J. Arachnol.* **1994.** 22:5-11.

Stephens, J.; Fahnestock, S.; Farmer, R.; Kiick, K.; Chase D.; Rabolt.;J. *Biomacromolecules*. **2005**. 6(3)1405-1413.

Subramanian, S. *Proteins.* **1998**. 32:1-2. <http://workbench.sdsc.edu>

Teulé, F.; Furin, W.; Cooper, A.; Duncan, J.; Lewis, R. *J. Mater. Sci*. **2007**. 42:8974-8985.

Teulé, F.; Cooper A.; Furin, W.; Bittencourt, D.; Rech, E.; Brooks, A.; Lewis, R. *Nat. Protoc.* **2009**. 4 (2):341-355.

Teulé, F.; Addison, B.; Cooper, A.; Ayon, J.; Henning, R.; Benmore, C.; Holland, G., Yarger, J.; Lewis, R. *Biopolymers*. **2012**. 97(6):418-431.

Thiel, B.; Kunkel, D.; Viney, C. *Biopolymers.* **1994**. 34:1089-1097.

Thiel, B. & Viney, C. *Science.* **1996**. 273:1480-1481.

Tuzikov, A.; Chinarev, A.; Gambaryan, A.; Oleinikov, V.; Klinov, D.; Matsko, N.; Kadykov, V.; Ermishov, M.; Demin, I.; Demin, V.; Rye, P.; Bovin, N. *ChemBioChem.* **2003**. 4:147-154.

Urry, D.; Luan, C.; Peng, S. *Ciba. Found. Symp.* **1995**. 192:4-30.

van Beek, J.; Hessc S.; Vollrath, F.; Meier, B. *PNAS.* **2002**. 99(16):10266-71.

Van Dijk, A.; De Beof, E.; Bekkers, A.; Van Wijk, L.; Van Swieten, E.; Hamer, R. Robillard, G. *Protein Sci.* **1997.** 6:649-656.

Vollrath, F. & Edmonds, D. *Nature*. **1989**. 340:305-307.

Vollrath, F.; Fairbrother, W.; Williams, R.; Tillinghast, E.; Bernstein, D.; Gallager, K.; Townley, M. *Nature.* **1990**. 345:526–528.

Vollrath, F.; Knight, D. *Nature*. **2001**. 410:541-548.

Zhou, Y.; Wu, X.; Conticello, V. *Biomacromolecules.* **2001**. 2:111-125.

Zhou, L.; Chen, X.; Shao, Z.; Huang, Y.; Knight, D. *J. Phys. Chem. B*. **2005**. 109(35):16937-16945.

Zhou, L.; Chen, X., Dai, W.; Shao, Z. *Biopolymers.* **2006**. 82:144-151.

This page is intentionally left blank.

Chapter 3

Summary and Concluding Comments

Further Discussion and Recommendations

The specific aims of this study were to investigate the mechanical properties of the flagelliform GGX and spacer motifs and to analyze the relationship between motif structure and function in fiber proteins. Recombinant proteins were generated consisting of repeating ensembles of different combinations of the GGX, GPGGX, and the spacer motifs of the Flag protein that produce flagelliform spider silk. The G, GY, GF and GFY, proteins were spun into fibers containing the corresponding ensembles of the motif modules $(GGX)_7$, $(GGX)_7$+$(GPGGX)_8$, $(GGX)_7$+spacer, and $(GGX)_7$+spacer+$(GPGGX)_8$. The fibers were mechanically tested and subjected to Raman spectroscopy and X-ray diffraction analysis for structural determination. Conclusions from these studies show the following:

Mechanical testing:

- The GGX motif contributes to extensibility or molecular mobility.
- The spacer motif contributes to tensile strength.

Structural Analysis:

- Raman analysis of the fibers suggests increased PGII helixes and β-sheet formation increases proportionately as spacer motif proportion increases.
- Fiber XRD analysis suggests a proportional increase in molecular order with an increase in spacer motif proportion.

Spacer motif

The mechanical testing results provide the first evidence that the spacer motif is indeed a contributor of tensile strength in the flagelliform protein fibers. This mechanical strength is supported structurally by the occurrence of an observable ring in the XRD analysis suggesting the existence of helix or β-sheet type structures found only in the spacer containing protein fibers GF and GFY. Unfortunately, the limitations of current technology do not allow the elucidation of fine structural differences for proteins within thin fibers (Lefèvre *et al.,* 2012). The Raman spectroscopic analysis of the proteins in the fibers showed greater ratios of PGII helices and the formation of β-sheets in greater proportion in the spacer containing fibers. The collective structural data suggests increased molecular organization in spacer containing fibers but does not define the structure of the spacer.

GGX motif

The GGX motif likely facilitates 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 β-sheets (Kümmerlen *et al*., 1996; van Beek *et al*., 2002) and facilitate structural alignment of other motifs that may form β-sheets, nanosprings, or other structures, tethering the structures together while acting as a flexible connector. 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.

β-sheets

Previous reports indicate a lack of β-sheet structure observed with low molecular orientation in native flagelliform silk (Rousseau *et al*., 2009). However, our Raman data of the Flag-like proteins indicates β-sheets and helical structures. While some previous studies suggested the possibility of β-sheet structures in flagelliform silk, Lefèvre and co-workers (2012) showed the presence of β-sheet in flagelliform silk of multiple species of orb weaver spiders. The actual structures will need to be confirmed by additional definitive biophysical studies in the future.

It could be argued that there are no organized structures in native flagelliform silk fibers. This is due in part to previous XRD data of native flagelliform silk that found no observable structure signal at 10Å (Craig *et al*., 2002). However, we would expect the signal for the structure observed in the native fiber protein to be very weak if detectable at all. The lack of an observable ring at 10Å in native flagelliform could be explained by the low content of the spacer motif, which comprises only 7.5% of the native Flag protein core. The spacer motif content in the proteins for the GF and GFY fibers is approximately 60% and 34% respectively, of total amino acids present in the protein. Not only does this help explain the observed structural signal in the recombinant protein

XRD data, it also explains why previous X-ray diffraction studies have not detected the structures in native Flag. It may be that the 7.5 % content of the spacer motif of native Flag produces structures that are quantitatively below the limit of detection in native flagelliform fibers.

Alternative Approach

One alternative approach to this project would be to focus on only a single motif per dissertation project. This may help bring more depth to the work and allow for two studies that could be complementary to each other. The two projects could be designed such they would be compatible with future similar investigations of the protein X amino acid variations of the different motifs. The cloning strategy would remain the same for all projects, maintaining compatible cloning strategies for project expansion. Additional recombinant proteins could be developed and added to the primary group, one at a time, fully characterizing them and comparing them to the previously characterized proteins.

Future work

The Flag protein of flagelliform silk provides a base for further protein studies. More may be learned through investigations of the mechanical and structural characteristics of new synthetic fibers generated by the variation of the amino acids in the

X position found in native Flag. In addition, the number of motif repeats without interruption could be investigated. Questions could be addressed such as: How strong can a fiber become by increasing the polyalanine region? How extensible can a fiber become by increasing the GPGXX motif repeat region? At what point do you lose extensibility by truncating the GPGXX motif repeat array? Are the spacer motifs of MiSp1 and Flag interchangeable? As understanding develops, more questions will be generated.

Lefèvre recently developed a new method specifically for flagelliform silk that reports 8% β-sheet in flagelliform fibers proteins (Lefèvre *et al.,* 2012). Their results are consistent with our mechanical and structural findings. The future comparison of this new method with that used to analyze the G, GY, GF, and GFY proteins would produce valuable data and a new perspective of the structure analysis.

Future advances in technology could present the greatest enhancement to silk research. Protein modeling can provide significant information in the future. Additional information may be learned from silk protein models that would save both time and research dollars of the wet bench science. Technological breakthroughs that provide the means to collect biophysical data from very small samples would also be greatly beneficial as this is currently restricting progress toward clearly understanding Flag protein structure/function relationships.

Artificially constructed proteins were directly compared in this research. One future project could be the development of optimized methods of the fiber spinning process for each of the proteins in this study. While this information would be interesting in further characterizing the individual proteins, it is time consuming and adds multiple variables. Any added variables can greatly affect the mechanical properties and the structure of the proteins. The focus of this dissertation was to determine what can be learned from the data comparisons of the recombinant proteins that can be correlated with the native Flag protein.

Further study of the G, GY, GF, and GFY recombinant proteins should include studies of the interaction of the proteins with water. The development of methodologies for spinning techniques that incorporate water into the fiber formation method may lead to interesting discoveries. Fiber exposure to water and water additions to the coagulation bath as the fibers were forming, as reported successful for Flag-MaSp2 chimeric proteins (Teulé *et al*., 2009), were detrimental to all four types of recombinant protein fibers in this project, G, GY, GF, and GFY. Although not published or presented here, some of the recombinant proteins showed promise in the development of methods to produce an aqueous spin dope. The Flag-like fibers are likely plasticized by water and may form more optimal molecular structures within the fibers, as described for native and recombinant silks (Teulé *et al.,* 2012). The effects of heat, extreme cold, UV, impact, or tension on the proteins and fiber performance would add more information about the proteins and their mechanical properties.

Designing and constructing additional new proteins would produce new molecular tools and protein models for analyses. For example, adding the A_n motif to the GFY protein ($GFYA_n$) would generate a Flag/MaSp2-like chimeric protein. Previous XRD analysis of a protein fiber that was constructed with the Y motif combined with the A_n motif generated proteins that formed β-sheet regions (Teulé *et al.,* 2012). Changing the concentration of the GPGGX motif amino acids relative to the total amino acids in the protein correlated with changes in extensibility of the fiber (Teulé *et al.,* 2012). The expected results of the XRD analysis of the proposed $GFYA_n$ fiber should indicate that β-sheet crystalline structures were produced proportionally to the A_n region. This would be due to the addition of the crystalline forming A_n region to the GFY motifs of the protein. Experiments along this line could improve characterization data that is presently inhibited for some large proteins by the lack of methods and technology to collect fine structural differences.

Flagelliform and other spider silks are also of interest for the development of biomaterials (Lewis 2006; Venderly *et al.,* 2007; Kluge *et al*., 2008; Heim *et al*., 2009; Spiess *et al*., 2010; Rising *et al*., 2011; Vollrath *et al.,* 2011). The flagelliform-like recombinants, or some variation of the Flag protein, may address biomaterial venues that require materials that are tough but more extensible than other materials. These Flag-like proteins may be very well suited for pharmaceutical delivery systems as they display the propensity to dissolve when exposed to moisture; a requirement of the drug delivery

system. Both the delivery system and these proteins will need further study before the Flag-like silk proteins will be ready for implementation into such a system.

This dissertation research supports utilizing specifically designed recombinant proteins as a tool to tackle the investigation of larger more complicated proteins such as Flag. The mechanical data and structural analyses contribute to the knowledge of flagelliform Flag protein determining the spacer motif as a contributor of tensile strength of the fibers. Additional technological advances and additional research will further clarify the structure and function of the flagelliform Flag GGX and spacer motifs.

Works Cited

Altman, G. H; Diaz, F; Jakuba, C.; Calabro, T; Horan, R L; Chen, J; Lu, H; Richmond, J; Kaplan, D. L. Silk-Based Biomaterials. *Biomaterials.* **2003**. *24:*401-416.

An, B.; Hinman, M.; Holland, G.; Yarger, J.; Lewis, R. Inducing Beta-sheet Formation in Fibers by Aqueous Post-spin Stretching. *Biomacromolecules.* **2011**. 12:2375-81.

Askarieh, G.; Hedhammar, M.; Nordling, K.; Saenz, A.; Casals, C.; Rising, A.; Johansson, J.; Knight, S. Self-assembly of Spider Silk Proteins is Controlled by a pH-Sensitive Relay. *Nature.* **2010**. 465:236-239.

Becker, N.; Oroudjev, E.; Mutz, S.; Cleveland, J.; Hansma, P.; Hayashi, C.; Makarov, D.; Hansma, H. Molecular Nanosprings in Spider Capture Silk Threads. *Nature Materials*. **2003**. 2:278-283.

Bell, A.; Peakall, D. Changes in Fine Structure During Silk Protein Production in the Ampullate Gland of the Spider *Araneus sericatus*. *Cell Biology*. **1969**. 42:284.

Benmore, C.; Izdebski, T.; Yarger, J. Total X-Ray Scattering of Spider Dragline Silk. *Physical Review Letters.* **2012**. 108(17):8102-1-4.

Blackledge, T.; Hayashi, C. Silken Toolkits: Biomechanics of Silk Fibers Spun by the Orb Web Spider *Argiope argentata* (Fabricius 1775). *Journal of Experimental Biology*. **2006**. 209:2452-2461.

Blamires, S.; Wu, C.; Blackledge, T.; Tso, I. Post-secretion Processing Influences Spider Silk Performance. Journal of the Royal Society. **2012**. Online.

Blamires, S.; Wu, C.; Blackledge, T.; Tso, I. Post-secretion Processing Influences Spider Silk Performance. *Journal of the Royal Society International.* **2012.** 1-9. Online.

Bonthrone, K.; Vollrath, F.; Hunter, B.; Sanders, J. The elasticity of spider's webs is due to water-induced mobility at a molecular level. *Proceedings of the Royal Society of London*. **1992**. 248:141-144.

Brooks, A.; Steinkraus, H.; Nelson, S.; Lewis, R. Probing the Elastic Nature of Spider Silk in Pursuit of the Next Designer Fiber. *Biomacromolecules*. **2005**. 6(6):3095-99.

Brooks, A.; Stricker, S.; Joshi, S.; Kamerzell, T.; Middaugh, C.; Lewis, R. Properties of Synthetic Spider Silk Fibers Based on *Argiope aurantia* MaSp2. *Biomacromolecules*. **2008**. 9 (6):1506-1510.

Bykov, S.; Asher, S. Raman Studies of Solution Polyglycine Conformations. *Journal of Physical Chemistry. B.* **2010**. 114:6636-6641.

Cheng, S.; Cetinkaya, M.; Gräter, F. How Sequence Determines Elasticity of Disordered Proteins. *Biophysical Journal*. **2010**. 99(12)386-3869.

Choresh, O.; Bayamagnai, B.; Lewis, R. Spider Web Glue: Two Proteins Expressed from Opposite Strands of the Same DNA Sequence. *Biomacromolecules.* **2009**. 10(10):2852-2856.

Craig, C. & Riekel, C. Comparative Architecture of Silks, Fibrous Proteins and Their Encoding Genes in Insects and Spiders. *Comparative Biochemistry and Physiology B. Biochemistry and Molecular Biology.* **2002**. 133(4):493-507.

Creager, M.; Jenkins, J.; Yeaman, L.; Brooks, A.; Jones, J.; Lewis, R.; Holland, G.; Yarger, J. Solid-State NMR Comparison of Various Spider's Dragline Silk Fiber. *Biomacromolecules.* **2010**. 11(8):2039-2043.

Debelle, L.; Alix, A.; Jacob, M.; Huvenne, J.; Berjot, M.; Sombret, B.; Legrand, P. Bovine Elastin and K-Elastin Secondary Structure Determination by Optical Spectroscopies. *Journal of Biological Chemistry.* **1995**. 270(44):26099-26103.

Denny, M.W. Silks – Their Properties and Functions. In: Vincent, J.F.V., Currey, J.D (Eds.), Mechanical Properties of Biological Materials. *Cambridge University Press*, Cambridge. *Cambridge University Press*, Cambridge. **1980**. 247-272.

Dong, Z.; Lewis, R.; Middaugh, C. Molecular Mechanisms of Spider Silk Elasticity. *Archives of Biochemistry and Biophysics.* **1991**. 284:53-57.

Elices, M.; Guinea, G.; Plaza, G.; Karatzas C.; Riekel, C.; Agulló-Reueda, F.; Daza, R.; Pérez-Rigueiro, J. Bioinspired Fibers Follow the Track of Natural Spider Silk. *Macromolecules.* **2011**. 44:1166-1176.

Foelix, R. *Biology of Spiders.* 2nd edition. New York: Oxford University Press, Inc. **1996**. 110-149.

Fukushima, Y. Genetically Engineered Synthesis of Tandem Repetitive Polypeptides Consisting of Glycine-Rich Sequence of Spider Dragline Silk. *Biopolymers.* **1998**. 45:269-279.

Garb, J.; DiMauro, T.; Vo, V.; Hayashi, C. Silk Genes Support the Single Origin of Orb-Webs. *Science*. **2006**. 312:1762.

Gatesy, J.; Hayashi, C.; Motriuk, D.; Woods, J.; Lewis, R. Extreme Diversity, Conservation and Convergence Spider Silk Fibroin Sequences. *Science.* **2001**. 291:2603-2605.

Gosline, J.; DeMont, M.; Denny, M. The Structure and Properties of Spider Silk. *Endeavour.* **1986**. 10:37-43.

Gosline, J.; Pollak, C.; Guerette, P.; Cheng, A.; DeMont, M.; Denney, M. Elastomeric Network Models for the Frame and Viscid Silks from the Orb Web of the Spider *Araneus diadematus*. *American Chemical Society Symposium Series*. 544 (Silk Polymers). **1994**. 328-341.

Gosline, J.; Guerette, P.; Ortlepp, C.; Savage, K. The Mechanical Design of Spider Silks; from Fibroin Sequence to Mechanical Function. *Journal of Experimental Biology.* **1999**. 202:3895-3303.

Graber, 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.; Srajer, V.; Viccaro, P.; Westferro, F.; Anfinrud, P.; Moffat, K. BioCARS: a Synchrotron Resource for Time-Resolved X-ray Science. *Journal of Synchrotron Radiation.* **2011.** 18:658-670.

Grubb, D.; Jelinski, L. Fiber Morphology of spider Silk: The Effects of Tensile Deformation. *Macromolecules.* **1997**. 30:2860-2867.

Hauser, C.; Deng, R.; Mishra, A.; Loo, Y.; Khoe, U.; Zhuang, F.; Cheong, D.; Accardo, A.; Sullivan, M.; Riekel, C.; Ying, J.; Hauser, U. Natural Tri- to Hexapeptides Self-Assemble in Water to Amyloid β-type Fiber Aggregates by Unexpected α-helical Intermediate Structures *Proceedings of the National Academies of Sciences.* **2011**. 108(4):1361-1366.

Hayashi, C.; Lewis, R. Evidence from Flagelliform Silk cDNA for the Structural Basis of Elasticity and Modular Nature of Spider Silks. *Journal of Molecular Biology*. **1998**. 275:773-784.

Hayashi, C.; Shipley, N.; Lewis, R. Hypotheses that Correlate the Sequence, Structure, and Mechanical Properties of Spider Silks. *International Journal of Biological Macromolecules*. **1999**. 24(2-3):271-275.

Hayashi, C.; Lewis, R. Molecular Architecture and Evolution of a Modular Spider Silk Protein Gene. *Science.* **2000**. 287:1477-1479.

Hayashi, C.; Lewis, R. Spider Flagelliform Silk: Lessons in Protein Design, Gene Structure, and Molecular Evolution. *Bioassays*. **2001.** 23:750-756.

Heim, M.; Keerl, D.; Scheibel, T. Spider Silk: From Soluble Protein to Extraordinary Fiber. *Angewandte Chemie International Edition: English.* **2009**. 45(20):3584-3596.

Heim, M.; Ackerschott, C.; Scheibel, T. Characterization of Recombinantly Produced Spider Flagelliform Silk Domains. *Journal of Structural Biology.* **2010.** 170:420-425.

Hinman, M. & Lewis, R. Isolation of a Clone Encoding a Second Dragline Silk Fibroin. *Journal of Biological Chemistry*. **1992**. 267(27):19320-19324.

Holland, G.; Creager, M.; Jenkins, J.; Lewis, R.; Yarger, J. Determining Secondary Structure in Spider Dragline Silk by Carbon-Carbon Correlation Solid-State NMR Spectroscopy. *American Chemical Society*. **2008.** 130(30):9871-9877.

Holland, G.; Jenkins, J.; Creager, M.; Lewis, R.; Yarger, J. Quantifying the Fraction of Glycine and Alanine in β-Sheet and Helical Conformations in Spider Dragline Silk using Solid State NMR. *Royal Society of Chemistry*. **2008**. 5568-5570.

Jelinski, L.; Bly, A.; Liivak. O.; Michal. C.; LaVerde, G.; Seidel, A.; Shan, N.; Yang, Z. Orientation, Structure, Wet-spinning, and Molecular Basis for Supercontraction of Spider Dragline Silk. *International Journal of Biological Macromolecules*. **1999**. 24:197-201.

Jenkins, J.; Creager, M.; Butler, E.; Lewis, R.; Yarger, J.; Holland, G.; Solid-State NMR Evidence for Elastin-like β-turn structure in Spider Dragline Silk. *Chemistry Communications.* **2010**. 46:6714-6716.

Kluge, J.; Rabotyogova, O.; Leisk, G.; Kaplan, D. Spider Silks and Their Applications. *Trends Biotechnology.* **2008.** 26(5):244-51.

Kümmerlen, J.; van Beek, J.; Vollrath, F.; Meier, B. Local Structure in Spider Dragline Silk Investigated by Two-Dimensional Spin-Diffusion Nuclear Magnetic Resonance. *Macromolecules.* **1996**. 20:2920-2928.

Lazaris, A.; Arcidiacono, S.; Huang, Y.; Zhou, J.; Duguay, F.; Chretien, N.; Welsh, E.; Soares, J.; Karatza, C. Spider Silk Fibers Spun from Soluble Recombinant Silk Produced in Mammalian Cells. *Science.* **2002**. 295:472-476.

Lefèvre, T.; Rousseau, M.; Pézolet, M. Protein Secondary Structure and Orientation in Silk as Revealed by Raman Spectromicroscopy. *Biophysical Journal.* **2007**. 92:2885-2895.

Lefèvre, T.; Boudreault, S.; Cloutier, C.; Pézolet, M. Diversity of Molecular Transformations Involved in the Formation of Spider Silks. *Journal of Molecular Biology.* **2011**. 405:238-253.

Lefèvre, T.; Paquet-Mercier, F.; Rioux-Dubé, J.; Pézolet, M. Review Structure of Silk by Raman Spectromicroscopy: From the Spinning Glands to the Fibers. *Biopolymers.* **2011**. 97(6):322-336.

Lefèvre, T. & Pézolet, M. Unexpected β-sheets and Molecular Orientation in Flagelliform Spider Silk as Revealed by Raman Spectromicroscopy. *Soft Matter*. **2012**. 8:6350-6357.

Lewis, R.; Hinman, M.; Kothakota, S.; Fournier, M. Expression and Purification of a Spider Silk Protein: a New Strategy for Producing Repetitive Proteins. *Protein Expression and Purification*. **1996**. 7(4):400-406.

Lewis, Randolph V. Spider Silk: Ancient Ideas for New Biomaterials. *Chemistry Review*. **2006**. 106(9):3762-3774.

Liu, Y.; Shao, Z.; Vollrath, F. Relationships Between Supercontraction and Mechanical Properties of Spider Silk. *Royal Society of Chemistry Communication.* **2005**. 2489-2491.

Lotz, B.; Colonna, C. Chemical Structure and the Crystalline Structures of *Bombyx mori* Silk Fibroin. *Biochimie.* **1979**. 61(2):205-214.

Maji, S.; Wang, L.; Greenwald, J.; Riek, R. Chemical Structure and the Crystalline Structures of *Bombyx mori* Silk Fibroin. *FEBS Letters.* **2009**. 2610-2617.

Mangalam, H. UC Irvine. SDSC Biology Workbench TACG. **1996**. <http://workbench.sdsc.edu>.

Omenetto, F. & Kaplan, D. New Opportunities for an Ancient Material. *Science*. **2010**. 329(5991):528-531.

Parkhe, A.; Seeley, S.; Gardener, K.; Thompson, L.; Lewis, R. Structural Studies of Spider Silk Proteins in the Fiber. *Journal of Molecular Recognition.* **1997**. 10:1-6.

Prince, J.; McGrath, K.; DiGirolamo, C.; Kaplan, D. Construction, Cloning, and Expression of Synthetic Genes Encoding Spider Dragline Silk. *Biochemistry.* **1995**. 34:10879-10885.

Rammensee, S.; Slotta, U.; Scheibel, T.; Bausch, A. Assembly Mechanism of Recombinant Spider Silk Proteins. *Proceedings of the National Academy of Science.* **2008**. 105:6590-6595.

Riekel, C.; Branden, C.; Craig, C.; Ferrero, C.; Heidelbach, F.; Muller, M. Aspects of X-ray Diffraction on Single Spider Fibers. *International Journal of Biological Macromolecules.* **1999**. 24(2-3):179-186.

Rising, A.; Widhe, M.; Johansson, J.; Hedhammar, M. Spider Silk Proteins: Recent Advances in Recombinant Production, Structure-function Relationships and Biomedical Applications. *Cellular and Molecular Life Science.* **2011**. 68(2):169-184.

Rousseau, M.; Lefèvre, T.; Beaulieu, L.; Asakura, T.; Pézolet, M. Study of Protein Conformation and Orientation in Silkworm and Spider Silk Fibers using Raman Microspectroscopy. *Biomacromolecules.* **2004.** 5:2247-2257.

Rousseau, M.; Lefèvre, T.; Pézolet, M. Conformation and Orientation of Proteins in Various Types of Silk Fibers Produced by *Nephila clavipes* Spiders. *Biomacromolecules.* **2009**. 10:2945-2953.

Sambrook, J.; Fritsch, E.; Maniatis, T. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. **1989**.

Savage, K. & Gosline, J. The Effect of Proline on the Network Structure of Major Ampullate Silks as Inferred from their Mechanical and Optical Properties *Journal of Experimental Biology.* **2008a**. 211:1937-1947.

Savage, K.& Gosline, J. The Role of Proline in the Elastic Mechanism of Hydrated Spider Silks. *Journal of Experimental Biology.* **2008b**. 211:1948-1957.

Scheller, J.; Guhrs, K-H.; Grosse, F.; Conrad, U. Production of Spider Silk Proteins in Tobacco and Potato. *Nature Biotech.* **2001**. 19:573-577.

Seidel, A.; Liivak, O.; Jelinski, L. Artificial Spinning of Spider Silk. *Macromolecules.* **1998**. 31:6733-6736.

Seidel, A.; Liivak, O.; Calve, S.; Adaska, J.; Ji, G.; Yang, Z.; Grubb, D.; Zax, D.; Jelinski, L. Regenerated Spider Silk: Processing, Properties, and Structure. *Macromolecule*s. **2000**. 33(3):775–780.

Shao, Z.; Vollrath, F.; Sirichaisit, J.; Young, R. Analysis of Spider Silk in Native and Supercontracted States using Raman Spectroscopy. *Polymer.* **1999**. 40:2493-2500.

Shao, Z.; Vollrath, F. Surprising Strength of Silkworm Silk. *Nature.* **2002**. 418(6899):741.

Simmons, A.; Michal, C.; Jelinski, L. Molecular Orientation and Two-Component Nature of the Crystalline Fraction of Spider Dragline Silk. *Science*. **1996**. 271(5245):84-87.

Spiess, K.; Lammel, A.; Scheibel, T. Recombinant Spider Silk Proteins for Applications in Biomaterials. *Macromolecular Bioscience*. **2010**. 10:998-1007.

Sponner, A.; Vater, W.; Rommerskirch, W.; Vollrath, F.; Unger, E.; Grosse, F.; Weisshart, K. The Conserved C-termini Contribute to the Properties of Spider Silk Fibroins. *Biochemistry and Biophysical Research Communication*. **2005**. 338(2):897-902.

Stauffer, S.; Coguill, S.; Lewis R. Comparison of the Physical Properties of Three Silks from *Nephila clavipes* and *Araneus gemmoides*. *Journal of Arachnology.* **1994.** 22:5-11.

Stephens, J.; Fahnestock, S.; Farmer, R.; Kiick, K.; Chase D.; Rabolt.;J. Effects of Electrospinning and Solution Casting Protocols on the Secondary Structure of a

Genetically Engineered Dragline Spider Silk Analogue Investigated via Fourier Transform Raman Spectroscopy. *Biomacromolecules*. **2005**. 6(3)1405-1413.

Teulé, F.; Furin, W.; Cooper, A.; Duncan, J.; Lewis, R. Modifications of Spider Silk Sequences in an Attempt to Control the Mechanical Properties of the Synthetic Fibers. *Journal of Material Science*. **2007**. 42:8974-8985.

Teulé, F.; Cooper A.; Furin, W.; Bittencourt, D.; Rech, E.; Brooks, A.; Lewis, R. Protocol for the Production of Recombinant Spider Silk-Like Proteins for Artificial Fiber Spinning. *Nature Protocols*. **2009**. 4 (2):341-355.

Teulé, F.; Addison, B.; Cooper, A.; Ayon, J.; Henning, R.; Benmore, C.; Holland, G., Yarger, J.; Lewis, R. Combining Flagelliform and Dragline Spider Silk Motifs to Produce Tunable Synthetic Biopolymer Fibers. *Biopolymers*. **2012**. 97(6):418-431.

Thiel, B.; Kunkel, D.; Viney, C. Physical and Chemical Microstructure of Spider Dragline: A Study by Analytical Transmission Electron Microscopy. *Biopolymers*. **1994**. 34:1089-1097.

Thiel, B. & Viney, C. Beta Sheets and Spider Silk. *Science*. **1996**. 273:1480-1481.

Tuzikov, A.; Chinarev, A.; Gambaryan, A.; Oleinikov, V.; Klinov, D.; Matsko, N.; Kadykov, V.; Ermishov, M.; Demin, I.; Demin, V.; Rye, P.; Bovin, N. Polyglycine II Nanosheets: Supramolecular Antivirals. *ChemBioChem*. **2003**. 4:147-154.

Urry, D.; Luan, C.; Peng, S. Molecular Biophysics of Elastin Structure, Function and Pathology. *Ciba. Foundation Symposium.* **1995**. 192:4-30.

van Beek, J.; Hessc S.; Vollrath, F.; Meier, B. The Molecular Structure of Spider Dragline Silk: Folding and Orientation of the Protein Backbone. *Proceedings of the National Academy of Science.* **2002**. 99(16):10266-71.

Van Dijk, A.; De Beof, E.; Bekkers, A.; Van Wijk, L.; Van Swieten, E.; Hamer, R. Robillard, G. Structure Characterization of the Central Repetitive Domain of High Molecular Weight Gluten Proteins. *Protein Science.* **1997**. 6:649-656.

Vendrely, C.; Ackerschott, C.; Römer, L.; Scheibe. Molecular Design of Performance Proteins with Repetitive Sequences. *Methods in Molecular Biology.* **2008**. 474:3-14.

Vendrely, C. & Scheibel T. Biotechnological Production of Spider-Silk Proteins Enables New Applications. *Macromolecular. Bioscience.* **2007.** 10;7(4):401-409.

Vollrath, F. & Edmonds, D. Modulation of the Mechanical Properties of Spider Silk Coating with Water. *Nature*. **1989**. 340:305-307.

Vollrath F. Spider Webs and Silk. *Scientific American.* **1992**. 266:70–76.

Vollrath, F. & Knight, D. Liquid Crystalline Spinning of Spider Silk. *Nature.* **2001**. 410:541-548.

Vollrath, F.; Porter, D.; Holland, C. There are Many More Lessons Still to be Learned from Spider Silks. *Soft Matters.* **2011**. 7:9595-9600.

Willcox, P.; Gido, S.; Muller, W.; Kaplan, D. Evidence of a Cholesteric Liquid Crystalline Phase in Natural Silk Spinning Processes. *Macromolecules.* **1996**. 29:5106-5110.

Zhou, Y.; Wu, X.; Conticello, V. Genetically Directed Synthesis and Spectroscopic Analysis of a Protein Polymer Derived from a Flagelliform Silk Sequence. *Biomacromolecules.* **2001**. 2:111-125.

Zhou, L.; Chen, X.; Shao, Z.; Huang, Y.; Knight, D. Effect of Metallic Ions on Silk Formation in the Mulberry Silkworm, *Bombyx mori*. *Journal of. Physical Chemistry B*. **2005**. 109(35):16937-16945.

Zhou, L.; Chen, X., Dai, W.; Shao, Z. X-ray Photoelectron Spectroscopic and Raman Analysis of Silk Fibroin–Cu(II) Films. *Biopolymers.* **2006**. 82:144-151.

Zhou, Q.; Lin, J.; Yan, F.; Ye, Z.; Qiu, F.; Tan, C.; Chen, Y.; Zhao, X. Self-assembly from Low Dimension to Higher Conformation of GGX Motif in Spider Silk Protein. *Current Nanoscience.* **2009**. 5:457-464.

This page is intentionally left blank.

Appendix

Tables

Figures

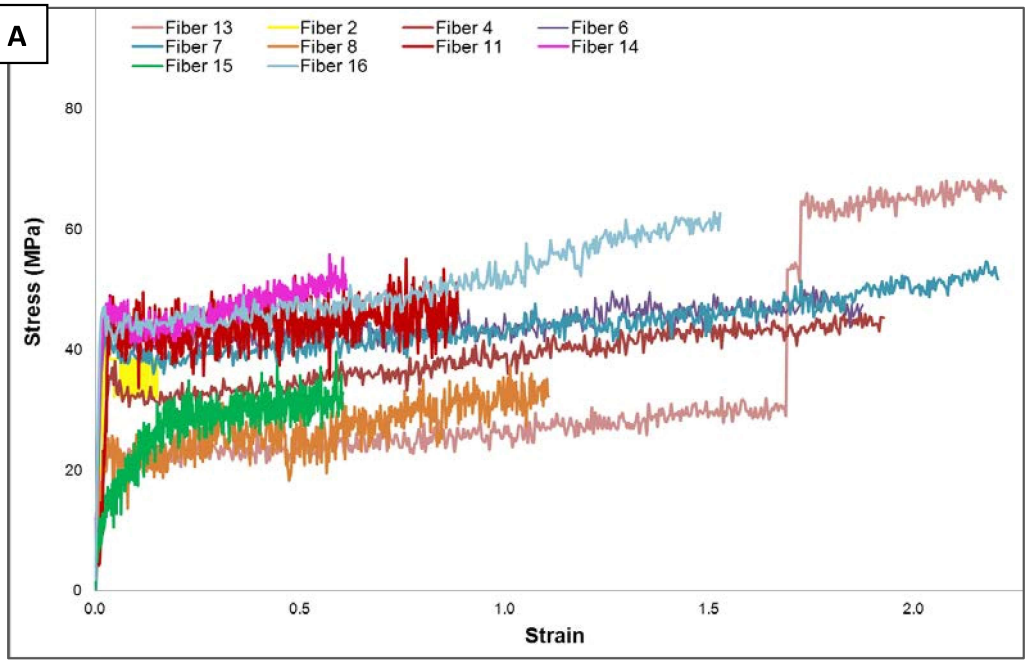
Mechanical Property	Method	Description	Units
Strength	Maximum Stress = σ_{max} Stress is obtained directly from the stress strain curve and is representative of strength/tensile strength	Greatest stress endured by the fiber throughout fiber testing prior to breaking	Pascals
Toughness	Energy to Break Area under the stress/strain curve	Total energy expensed to the breaking point of the fiber	MJ/m^3
Extensibility	Strain is the extensibility of the fiber $\frac{\text{change in fiber length (mm)}}{\text{initial fiber length (mm)}}$ $\Delta L/L_0$	Maximum tolerated strain at fiber failure	unitless
Extension %	Strain $(\Delta L/L_0) \times 100$	Presentation of the percent of extension of a fiber	Percent (%)
Initial Stiffness	Young's Modulus Initial Slope of the stress/strain curve	One elastic property of a material as it resists deformation	Pascals

Stress = Force (N)/Area (cross sectional area of the fiber (m2)

Table A.1. Mechanical property methods and descriptions.

The mechanical properties of materials are determined through calculations and various fiber analyses. The table summarizes the method or source used to determine the specific property. A brief description of the method and the output results of the mechanical property unit are listed. (Hearle and Morton, 2008)

A



B

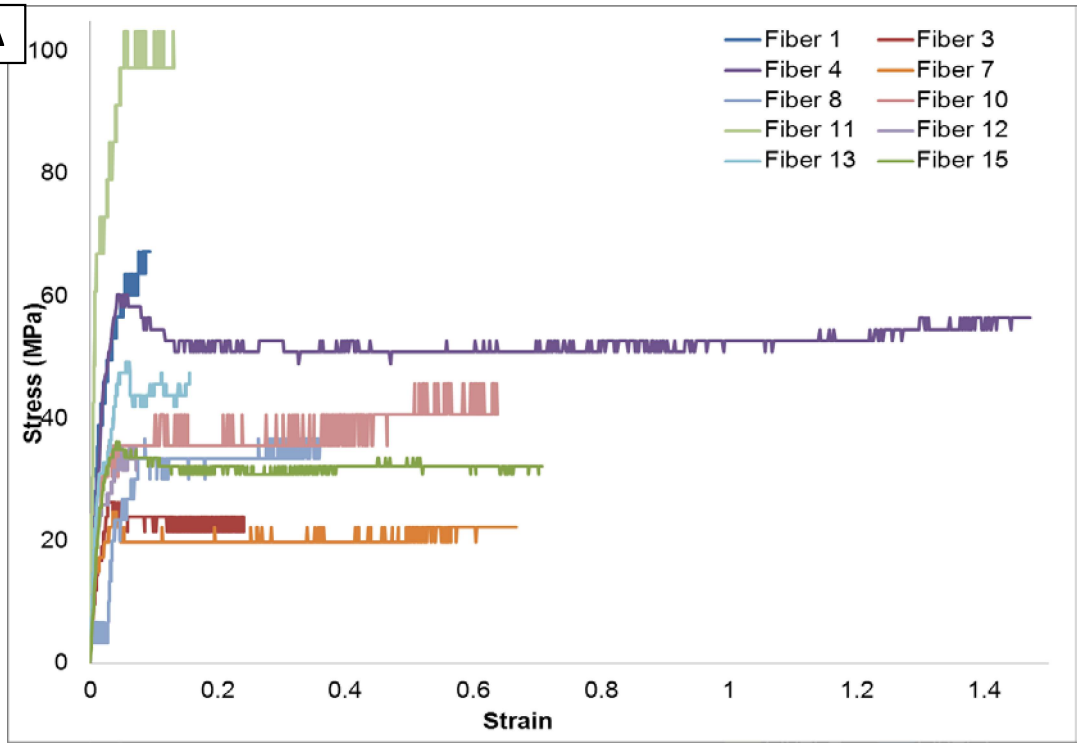
Fiber	Diameter microns	Maximum % Extension	Maximum Stress Mpa	Young's Modulus	Toughness $MJ.m^{-3}$
2	14.09	15.16	38.5	1481.8	5.32
4	15.25	192.59	43.79	1258.9	73.83
6	12.72	187.38	70.85	3787.2	119.62
7	12.12	220.58	86.66	4660.1	153.44
8	10.97	110.67	44.92	***2112.1	27.87
11	9.53	88.62	56.11	***959.05	37.55
13	15.16	238.12	72	2972.1	87.79
14	15.26	61.32	54.67	2136.3	28.43
15	11.46	60.58	38.77	***155.4	16.15
16	16.57	153.03	50.99	2860.8	65.51
Mean	13.31	132.81	55.73	2736.74	61.55
Std. Dev.	2.29	76.28	16.05	1657.36	47.96

***= $r^2 < 0.9$ excluded

Figure A.1. Post-spin drawn fiber mechanical properties.

Processing as-spun fibers by stretching in 80% aqueous alcohol refined the mechanical properties of the fibers. A) Stress/strain analyses of G post-spin drawn (PSD) fibers. B) Compiled averaged mechanical parameters for G PSD fibers.

A



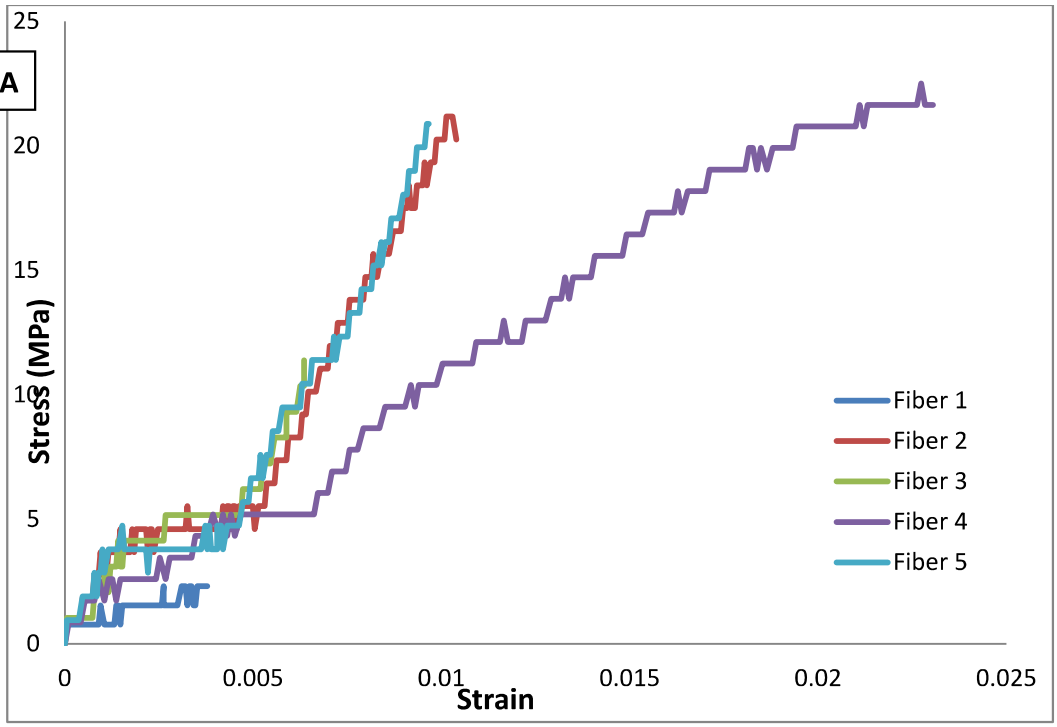
B

Fiber	Diameter microns	Maximum % Extensibility	Maximum Stress Mpa	Young's Modulus Mpa	Toughness $MJ.m^{-3}$
1	18.97	9.35	67.25	2568.5	4.84
3	23.08	24.04	23.89	775.6	5.40
4	26.02	147.08	56.44	2195.8	76.79
7	22.67	66.79	22.29	1532.3	13.56
8	19.51	35.99	36.78	**1720.9	10.86
10	15.83	63.77	45.75	2319.6	23.88
11	14.47	13.13	103.4	6688.3	11.59
12	26.22	7.24	35.18	2514.4	1.99
13	26.40	15.61	47.49	2418.9	6.35
15	30.79	70.82	32.23	1783.55	22.46
Mean	21.46	45.38	47.07	2532.99	17.77
Std. Dev.	4.51	43.47	25.11	1857.19	23.24

Figure A.2. Post-spin drawn GY fiber mechanical properties.

Processing as-spun fibers by stretching in 80% aqueous alcohol refined the mechanical properties of the fibers. A) Stress/strain analyses of GY post-spin drawn (PSD) fibers. B) Compiled averaged mechanical parameters for GY PSD fibers.

A



B

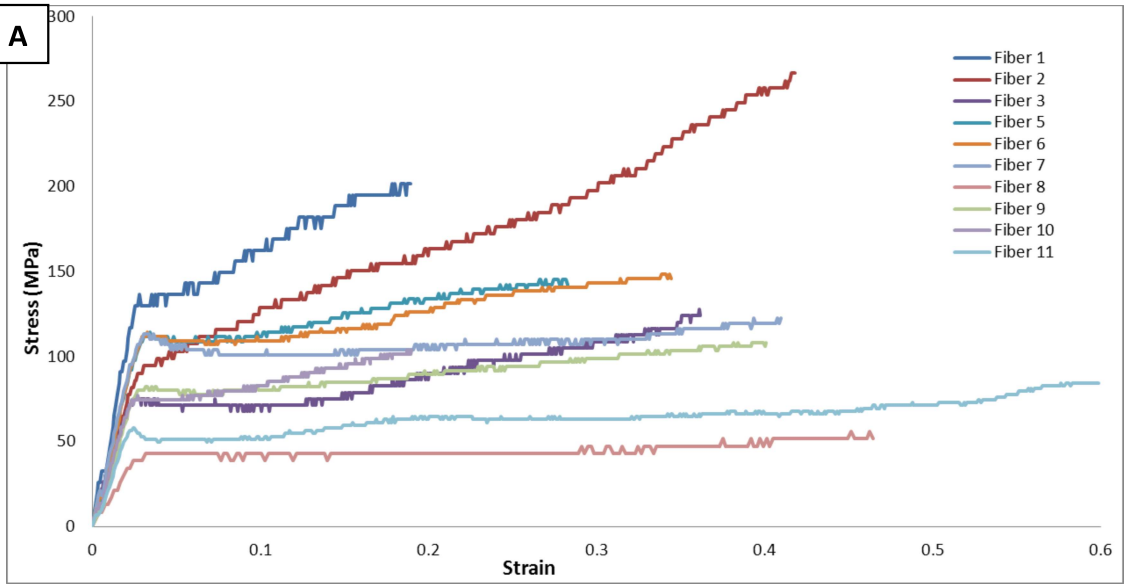
Fiber	Diameter microns	Maximum % Extensibility	Maximum Stress Mpa	Young's Modulus Mpa	Toughness $MJ.m^{-3}$
1	40.67	0.38	2.31	***	0.005
2	37.19	1.04	21.18	2706.8	0.09
3	35.08	0.01	11.38	2163.4	0.03
4	38.35	2.31	21.46	921.43	0.28
5	36.62	0.98	20.89	2773	0.08
Mean	36.81	1.09	18.73	2141.16	0.12
Std. Dev.	1.36	0.94	4.90	857.79	0.11

*** = $r^2 < 0.9$ excluded

Figure A.3. As-spun GF fiber mechanical properties.

A) The Stress/strain analyses of GF as-spun fibers. B) Compiled averaged mechanical parameters for GF as-spun fibers.

A



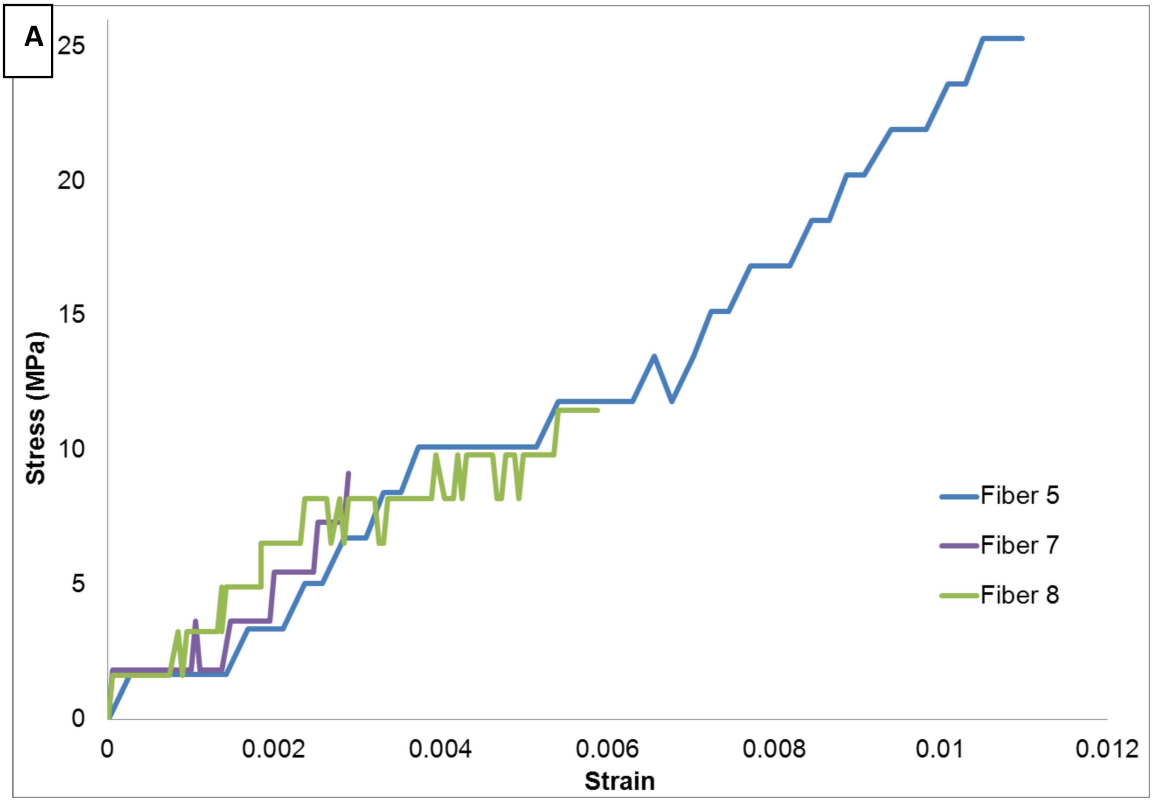
B

Fiber	Diameter microns	Maximum % Extensibility	Maximum Stress Mpa	Young's Modulus Mpa	Toughness $MJ.m^{-3}$
1	13.99	18.94	201.54	4981.5	28.97
2	17.21	41.90	266.55	3496.0	69.24
3	18.42	36.15	127.65	3231.5	31.20
5	21.35	28.27	145.19	4084.3	33.47
6	22.89	34.35	148.22	4180.1	41.25
7	20.39	40.96	122.54	4162.0	42.90
8	17.22	46.47	55.84	1542.9	20.29
9	23.26	40.12	108.28	3298.0	35.28
10	21.86	19.03	103.91	3334.3	15.20
11	27.72	59.90	84.51	2698.4	39.03
Mean	20.43	36.61	136.42	3500.90	35.68
Std. Dev.	3.89	12.45	60.27	945.76	14.71

Figure A.4. Post-spin drawn GF fiber mechanical properties.

A) Stress/strain analyses of GF processed (PSD) fibers. B) Compiled averaged mechanical parameters for GF PSD fibers.

A



B

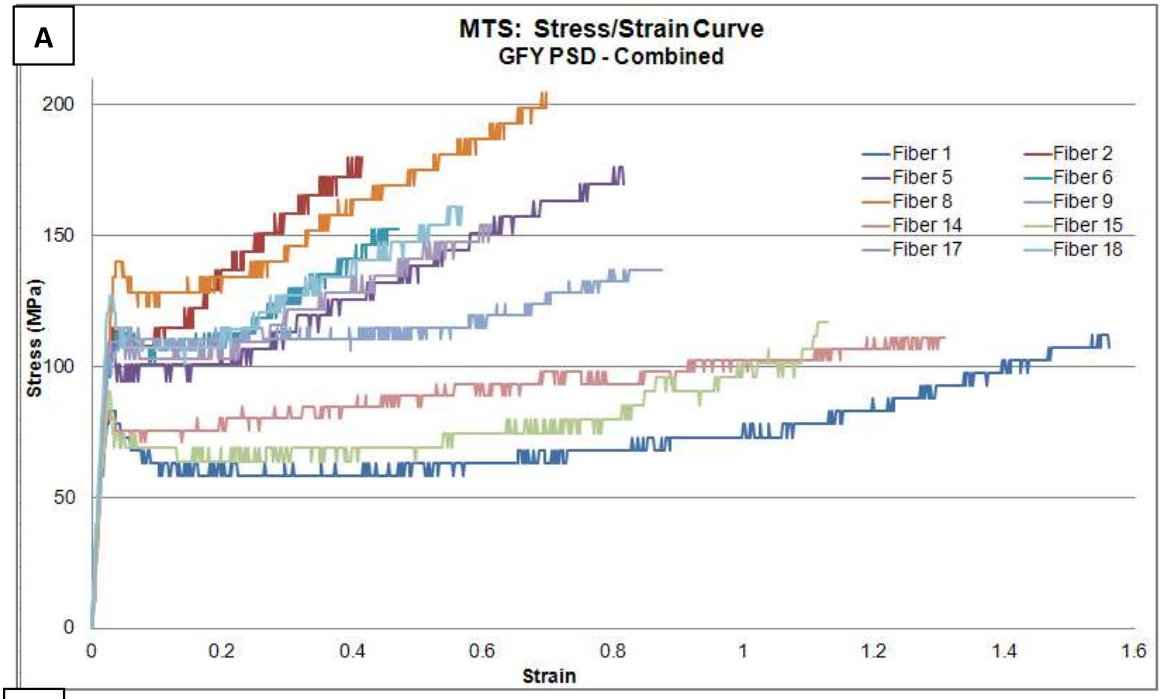
Fiber	Diameter microns	Maximum % Extensibility	Maximum Stress Mpa	Young's Modulus Mpa	Toughness $MJ.m^{-3}$
5	27.48	1.10	25.29	2415.1	0.13
7	26.38	0.29	9.15	***	0.01
8	27.86	0.59	44.48	2933.7	0.04
Mean	27.24	0.66	26.31	2674.38	0.06
Std. Dev.	0.77	0.41	17.69	1565.68	0.06

*** $r^2 < 0.9$ excluded

Figure A.5. As-spun GFY fiber mechanical properties.

A) Stress/strain analyses of GFY as-spun fibers. B) Compiled averaged mechanical parameters for GFY AS fibers.

A



B

Fiber	Diameter microns	Maximum % Extensibilitiy	Maximum Stress Mpa	Young's Modulus Mpa	Toughness $MJ.m^{-3}$
1	16.16	156.02	112.18	3539.3	115.09
2	13.30	41.33	179.33	4212.1	55.29
5	14.23	81.72	176.11	4608.9	104.52
6	15.02	47.12	152.46	4155.8	55.93
8	14.76	69.76	204.57	4115.6	107.02
9	16.96	87.39	137.17	3958.6	100.83
14	16.91	130.82	111.35	3796.0	119.99
15	15.46	112.88	117.2	3721.6	87.68
17	14.07	61.21	154.29	4651.5	73.49
18	13.77	56.73	161.11	4904.5	70.70
Mean	15.06	84.50	150.58	4166.39	89.05
Std. Dev.	1.29	37.82	31.32	441.25	23.93

Figure A.6. Post-spin drawn GFY fiber mechanical properties.

A) Stress/strain analyses of GFY processed (PSD) fibers. B) Compiled averaged mechanical parameters for GFY PSD fibers. This fiber is most similar to native Flag.

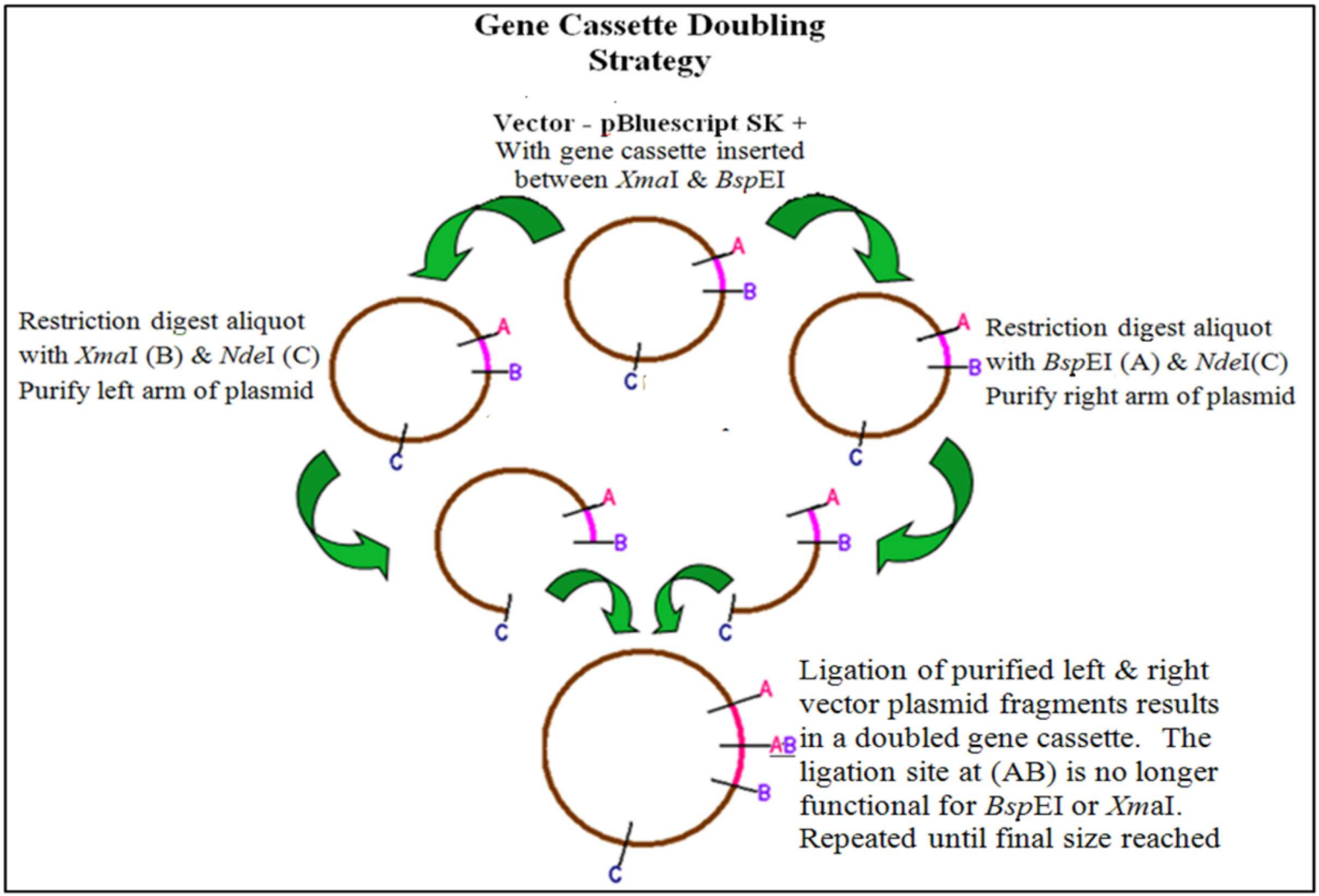


Figure A.7. The doubling strategy for dsDNA sequences in pBluescript

This page is intentionally left blank.