12-2018

Production and Purification of Synthetic Minor Ampullate Silk Proteins

Danielle A. Gaztambide
Utah State University

PRODUCTION AND PURIFICATION OF SYNTHETIC MINOR AMPULLATE SILK PROTEINS

by

Danielle A. Gaztambide

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Biological Engineering

Approved:

_____	_____
Randolph V. Lewis, Ph.D Major Professor	Jixun Zhan, Ph.D Committee Member
_____	_____
Ronald Sims, Ph.D Committee Member	Richard S. Inouye, Ph.D School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2018

ABSTRACT

Production and Purification of Recombinant Minor Ampullate Silk Proteins

by

Danielle A. Gaztambide, Master of Science

Utah State University, 2018

Major Professor: Randolph V. Lewis, Ph.D.
Department: Biological Engineering

Spider silks long have been recognized as some of nature's most impressive materials. By divergent evolution, orb-weaving spiders produce and utilize six distinct silk fibers and one glue for a variety of biological roles. Each silk is the product of proteins, highly conserved across species, called spidroins. Because it is not possible to farm spiders, these proteins must be produced synthetically in order to study the protein sequences and use these proteins for material applications.

This research uses *Escherichia coli* (*E. coli*) as a production system to investigate the first known expression and purification of synthetic minor ampullate silk, a lesser studied silk the spider uses to make up the auxiliary spiral of the orb web. The auxiliary spiral creates a template for the capture spiral and provides additional stability during web construction. Natural minor ampullate silk has a high tensile strength, does not supercontract in water, and is inelastic with fiber properties similar to Kevlar. This investigation completes the construction and expression of six synthetic analogs of minor ampullate silk protein ranging from 40 to 134 kilodaltons (kDa). Using fed-batch aerobic

fermentation, this thesis investigates the efficacy and optimization of *E. coli* as a production platform for recombinant minor ampullate silk proteins. This study also develops purification protocols for these proteins following production and performs the first preliminary characterization of synthetic recombinant minor ampullate silk proteins based on sequence and molecular weight.

(93 pages)

PUBLIC ABSTRACT

Production and Purification of Recombinant Minor Ampullate Silk Proteins

Danielle A. Gaztambide

Spider silks are incredible natural materials that have a wide variety of properties that can rival or outperform even common synthetic materials like Nylon and Kevlar. As nature's architects, orb-weaving spiders spin seven different silks that are used for very specific roles throughout the spider's lifecycle. These silks are comprised of proteins called spidroins. Each of these spidroins has evolved to have properties such as strength and/or stretch that make these silks successful and highly adapted in their designated roles in web construction, prey capture and reproduction.

This study involves the production of minor ampullate silk by genetically modifying the bacteria *Escherichia coli*. Minor ampullate is a lesser studied silk that is used for the first spiral of the orb web. This spiral is a template that the spider uses to finish the web and provides stability during the web construction. Minor ampullate silk is strong, however it does not stretch so it may be well-suited for certain applications such as ballistic materials.

By producing and purifying different arrangements of minor ampullate silk protein, it is possible to learn how this protein can be expressed without using the spider itself. This investigation sheds light on how deviations in the protein sequence and motif

arrangement can produce different properties, which can potentially be used to make new materials.

ACKNOWLEDGMENTS

I would like to thank my major professor, Dr. Randy Lewis, for his continued support and expertise throughout this project, as well as my lab manager and mentor, Dr. Justin Jones, for his guidance and assistance with protein purification and chromatography techniques. I would also like to thank my committee members, Dr. Jixun Zhan and Dr. Ronald Sims, for their guidance and teaching throughout my career at Utah State.

I give special thanks to all of my friend and colleagues at the spider silk lab for their encouragement, moral support, and teaching as I learned the techniques needed to complete this project.

Finally, I would like to thank my family and friends for their unwavering support and encouragement.

Danielle A. Gaztambide

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

NATIVE SPIDER SILKS

Since ancient times, spider silk has inspired intrigue with its strength, elasticity, toughness, ductility, thermostability, and biocompatibility when compared with both natural and synthetic materials. All 41,000 species of spiders produce and utilize silk. Depending on the species, they use silk for a variety of biological acts including prey capture, prey wrapping, shelter, egg casings, mating, and even parachuting[1]. Most spiders spin more than one type of silk. Notably, the orb-weavers are the best representation of millions of years of gene duplication and divergence that led to the variety of silks that these spiders utilize today[2].

Female orb-weaving spiders produce six distinct silk fibers and one glue. Each originates from a separate silk gland in the spider's abdomen as Figure 1-1 demonstrates. These silks include: major ampullate (dragline and web radii); minor ampullate (temporary auxiliary spiral); flagelliform (capture spiral); aggregate (aqueous glue); piriform (fiber attachment); aciniform (prey wrapping); and tubuliform (egg sacs and cocoons)[3]. Major ampullate silk proteins traditionally have been the most studied spider silk proteins due to their impressive combination of strength and elasticity that result in a toughness two to three times greater than synthetic fibers such as Kevlar and nylon. However, each of the silks possess different mechanical properties that may be ideally suited for material applications as diverse as ballistic material and very fine sutures[4].

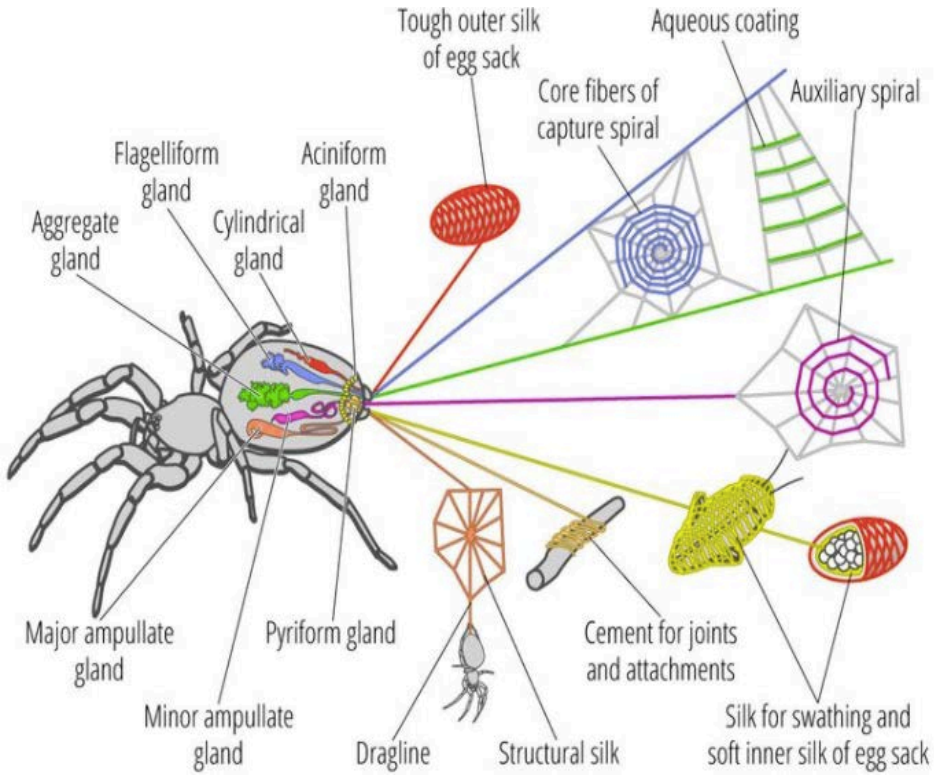

Figure 1-1. Spider silk glands and silk biological roles

Table 1-1 illustrates how these silks compare with each other and with other natural and synthetic fibers[5]. Natural minor ampullate silk has properties similar to Kevlar and tendon, due to its high tensile strength (1×10^9 N/m^2) and decreased elongation.

The self-assembling fibroins that make up these silks result in a diverse set of fibers with mechanical properties that are functionally related to their amino acid motifs and thus protein sequence and folding structures[6]. In nature, spider silk proteins are large (~350 kDa) and include nonrepetitive N- and C-termini, with the rest of the sequence consisting entirely of repetitive amino acid motifs.

Table 1-1: Fiber mechanical property comparisons.

material	strength ($N\ m^{-2}$)	elongation (%)	energy to break ($J\ kg^{-1}$)
dragline silk	4×10^9	35	4×10^5
minor ampullate silk	1×10^9	5	3×10^4
flagelliform silk	1×10^9	>200	4×10^5
tubuliform	1×10^9	20	1×10^5
aciniform	0.7×10^9	80	6×10^9
Kevlar	4×10^9	5	3×10^4
rubber	1×10^6	600	8×10^4
tendon	1×10^6	5	5×10^3

[a] Some data from Gosline, J. M.; Dennv, M. W.; DeMont, M. E. *Nature* **1984**, *309*, 551.

The N-terminus has a role in transport and encodes a signal peptide[7]; however, the role of the highly conserved C-terminus has remained somewhat elusive, with some investigations showing that it may be important for stability in the gland and fiber formation[8]. There are four types of shared amino acid motifs among the spider silk proteins: GPGGX/GPGQQ, GGX, poly-A/poly-Gly-Ala, and a nonrepetitive spacer region[3]. While much of the structure-to-function relationship in spider silks has remained elusive, fiber x-ray diffraction and nuclear magnetic resonance spectroscopy has shown that the poly-alanine and $(GA)_n$ repeats correlate to crystalline β-sheet structures, while the glycine-rich (GGX) repeats form GlyII-helices[9]. The GPGGX regions primarily found in major ampullate silk correlate to β-spiral structures, providing the dragline silk its elasticity[10].

Substantial sequence information for the minor ampullate spidroins (MiSps) was first published in 1997 and has since been corroborated by additional investigations;

however, there are currently no published investigations of synthetic analogs of these spidroins[12]. Minor ampullate silk, like many of the spider silks, is comprised of two proteins: MiSp1 (9.5 kb) and MiSp2 (7.5 kb)[13,14]. These two spidroins have similar repetitive regions; however, MiSp2 has a shorter and less conserved repetitive region when compared with MiSp1. Nevertheless, the repetitive regions in the MiSps are more highly conserved than either of the major ampullate silk proteins[13]. These repetitive regions of the MiSps sequence resemble amorphous regions in the heavy chain of the *Bombyx mori* silkworm[15]. MiSp1 and MiSp2 also contain identical 137 amino acid nonrepetitive spacer regions that occur every 10 repeats and are not homologous to any other spider silk sequence.[13] No secondary structure has been definitively assigned to this spacer region, however it is hypothesized to be the cause of the inelasticity of the resulting MiSp fiber and contains long stretches of α-helical structures[16]. Figure 1-2[17]

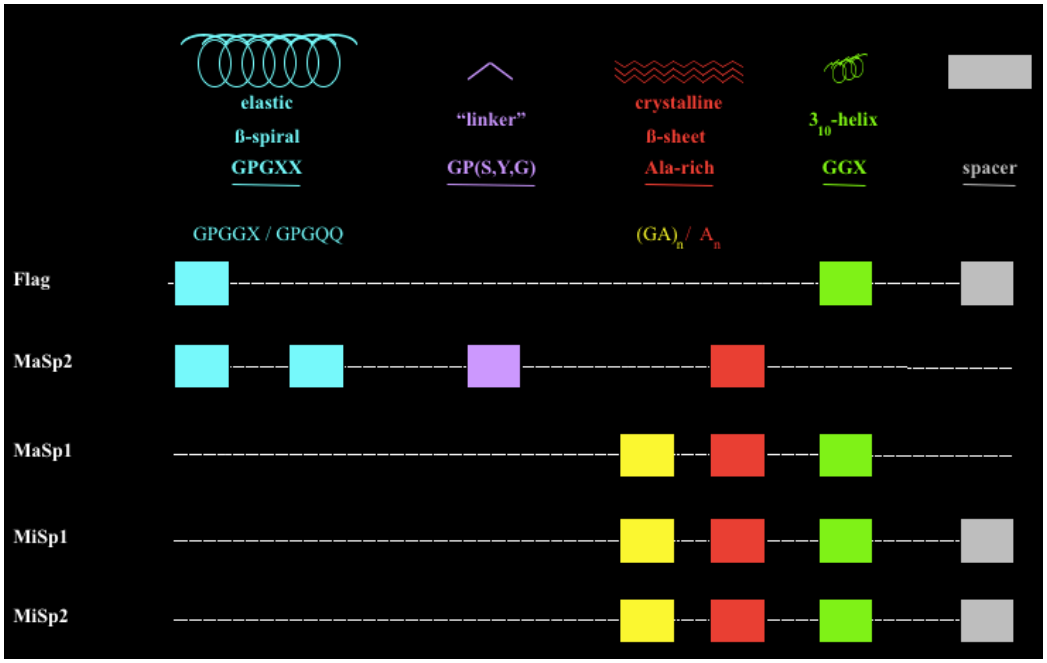

Figure 1-2. Secondary structures assigned to the prevalent motifs in spider silks

illustrates the secondary structures that have been assigned to the amino-acid motifs in spider silks. The repetitive motifs in MiSps include GX, GGX, GGGX and shorter poly-alanine repeats, when compared with major ampullate silk. While the repetitive sequence in MiSp is similar to major ampullate spidroin 1 (MaSp1), the longer poly-alanine repeats seen in major ampullate spidroins (MaSps) are replaced by $(GA)_n$ repeats in minor ampullate silk. The sequence hierarchy of minor ampullate silk proteins from the *Nephila clavipes* spider is illustrated in Figure 1-3[13]. This sequence organization is hypothesized to be the cause for the decreased tensile strength seen in MiSps, as the $(GA)_n$ β-sheet structures possess fewer hydrophobic interactions when compared with the

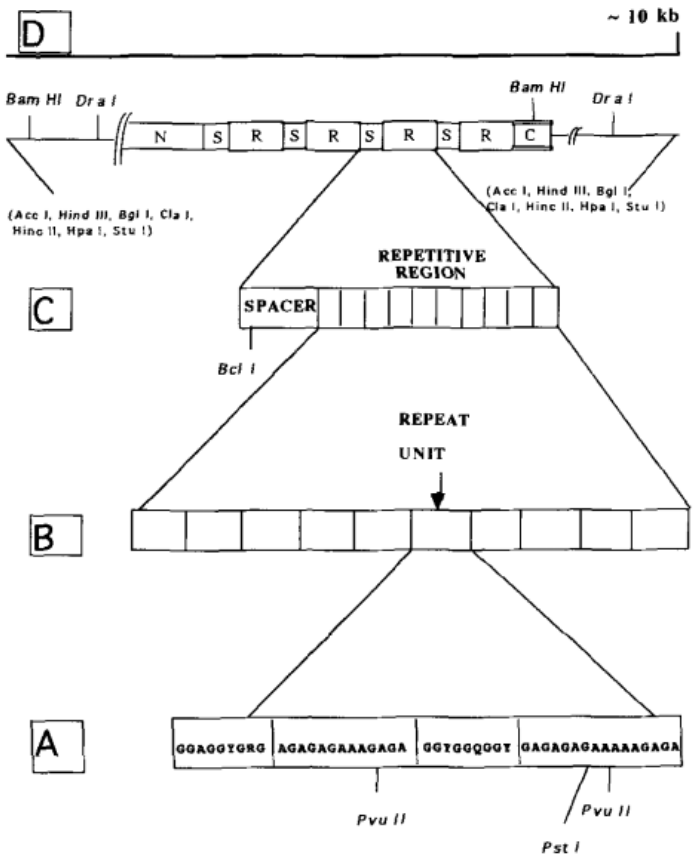

Figure 1-3. Sequence hierarchy of minor ampullate silk protein

longer poly-alanine β-sheet structures[16,18]. The MiSps also have different ratios of GGX and GA repeats and the repetitive region lacks the proline seen in MaSp2, which is hypothesized to give major ampullate silk its elasticity[9,19,20].

SYNTHETIC RECOMBINANT SPIDER SILK PRODUCTION

Due to spiders' cannibalistic and territorial behavior, and the fact that it is difficult to obtain specific spider silk proteins, farming spiders to study and use their silk is not economically or practically feasible. For reference, it took eight years to make a single spider silk cape with fibers harvested from 1.2 million orb webs[21]. Therefore, it is imperative to use transgenic hosts to produce spider silk synthetically for large-scale production of recombinant spider silk proteins (rSSps) so that these proteins may be utilized for functional materials[12,22]. These transgenic hosts have included both unicellular and multicellular organisms as heterologous hosts. While each transgenic host has its pros and cons, spider silk proteins are inherently difficult to produce synthetically regardless of their host due to their large size (~350 kDa) and highly repetitive sequence[21,23,24]. Multicellular organisms typically produce proteins at the size scale of spider silk proteins; however, genetic instability, protein aggregation post-translation, low mechanical properties, low yields, and high costs remain significant problems for large-scale production. Multicellular organisms that have been investigated for the expression of spider silk proteins include goats, mice, silkworms, and plants[12]. A timeline of spider silk production in a variety of hosts is shown in Figure 1-4[21].

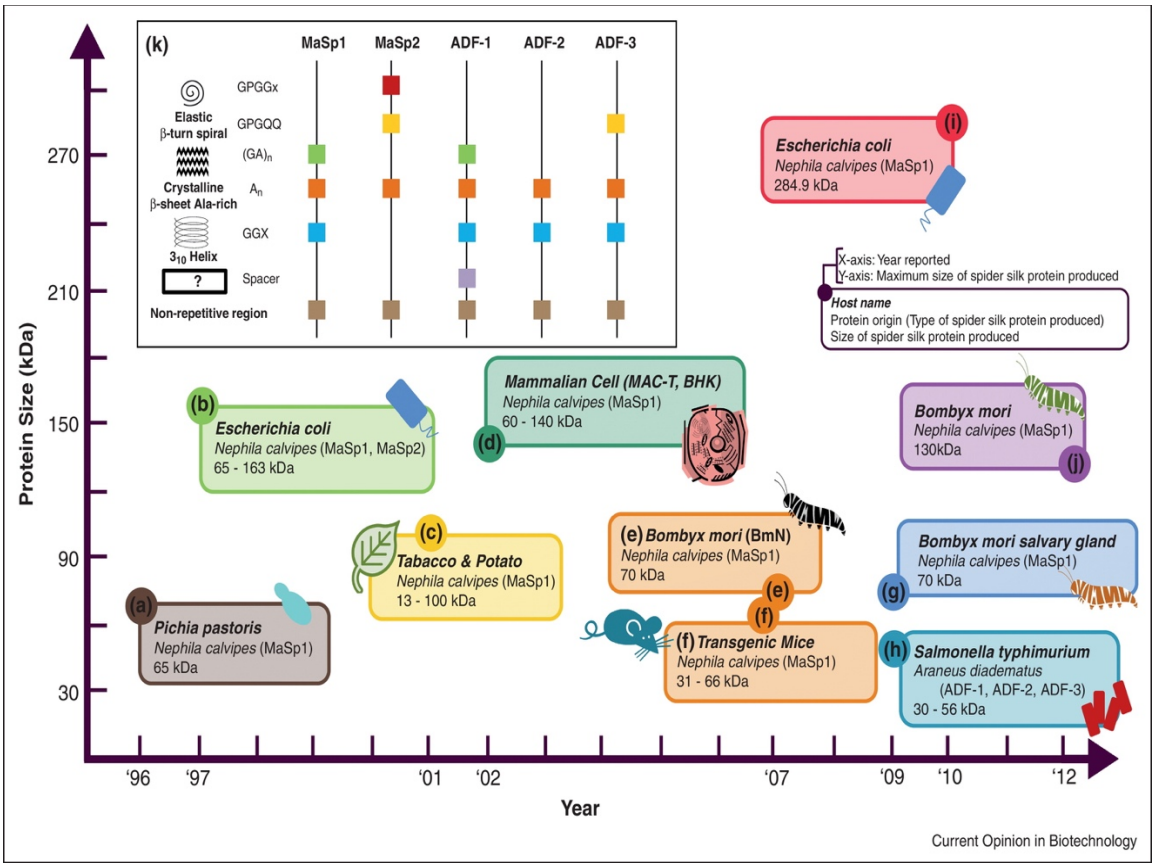

Figure 1-4. Chronological representation of synthetic recombinant spider silk production in transgenic hosts

Because of the low cost of production, established protocols for scalability, and ease of genetic manipulation, unicellular organisms such as bacteria, and more specifically *E. coli*, have hosted most of the research into the production of synthetic spider silk proteins[25–27]. *E. coli* is a well-suited choice for large-scale protein expression and production because a large library of expression plasmids, molecular protocols, and cultivation parameters that are already established[28]. Adding to its cost-effectiveness, *E. coli* is also a much faster expression system to develop than many of the multicellular systems described above.

While synthetic production of spider silk proteins has been the subject of a large amount of investigation in recent years, most of these studies have been done with the major ampullate silk proteins[17,26,29–31]. Minor ampullate silk is one of the lesser studied silks due to the fact that the minor ampullate gland is much smaller and thus harder to isolate than the major ampullate gland. Minor ampullate fibers also typically have smaller diameters than the major ampullate silk fibers and their biological role is not as readily understood[19]. As well, based upon evidence presented in this thesis, it is very possible that when expressed in *E. coli* the proteins become so insoluble as to be unrecoverable by common methods. In essence, their levels of insolubility make it appear as if no protein was expressed. This may be one of the leading reasons for the lack of study of recombinant forms of other spider silk proteins like the minor ampullate silk proteins.

Aside from the type of spider silk being produced, there are also many other hurdles to protein expression for synthetic spider silk proteins that are the result of the large size and repetitive nature of the protein sequence. Minimizing truncation of these high molecular weight products will need to be considered, as translation pauses and proteolytic degradation are all factors for high-density, cell culture expression of these repetitive proteins[12,17,21–23,28,32]. These proteins also have a tendency to fold incorrectly post-translation and, due to their self-assembling nature, create insoluble aggregates in a manner similar to inclusion bodies[21]. These aggregated proteins create even more challenges for purification, as the solubility and protein-protein interactions can defy conventional purification techniques such as affinity chromatography.

Producing these large proteins in a cost-effective manner will also be a deciding factor if spider silk is to replace synthetic fibers such as nylon or Kevlar. Figure 1-5[33]

shows a cost-analysis profile of rSSps produced in *E. coli.* As you can see, the cost of production of these rSSps depends on the capacity of the production system, an area that has seen continuous improvement in recent years[23,24].

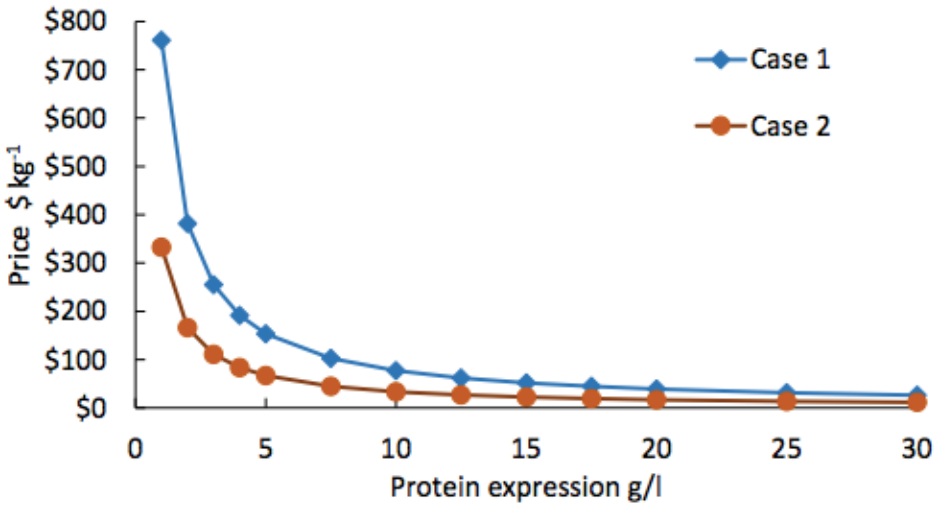

Figure 1-5. Cost analysis for production of rSSps in *E. coli*. Case 1: baseline scenario. Case 2: optimized scenario.

MATERIAL APPLICATIONS OF SPIDER SILK PROTEINS

Harnessing the properties of natural spider silk proteins synthetically will create possibilities for new, lightweight polymer materials that outperform current synthetic and natural materials. Figure 1-6[34] illustrates the native and synthetic processes for creating these silk materials and their subsequent potential product applications. Understanding the processing conditions necessary in synthetic systems to replicate silk properties will be crucial for future silk production and material characterizations.

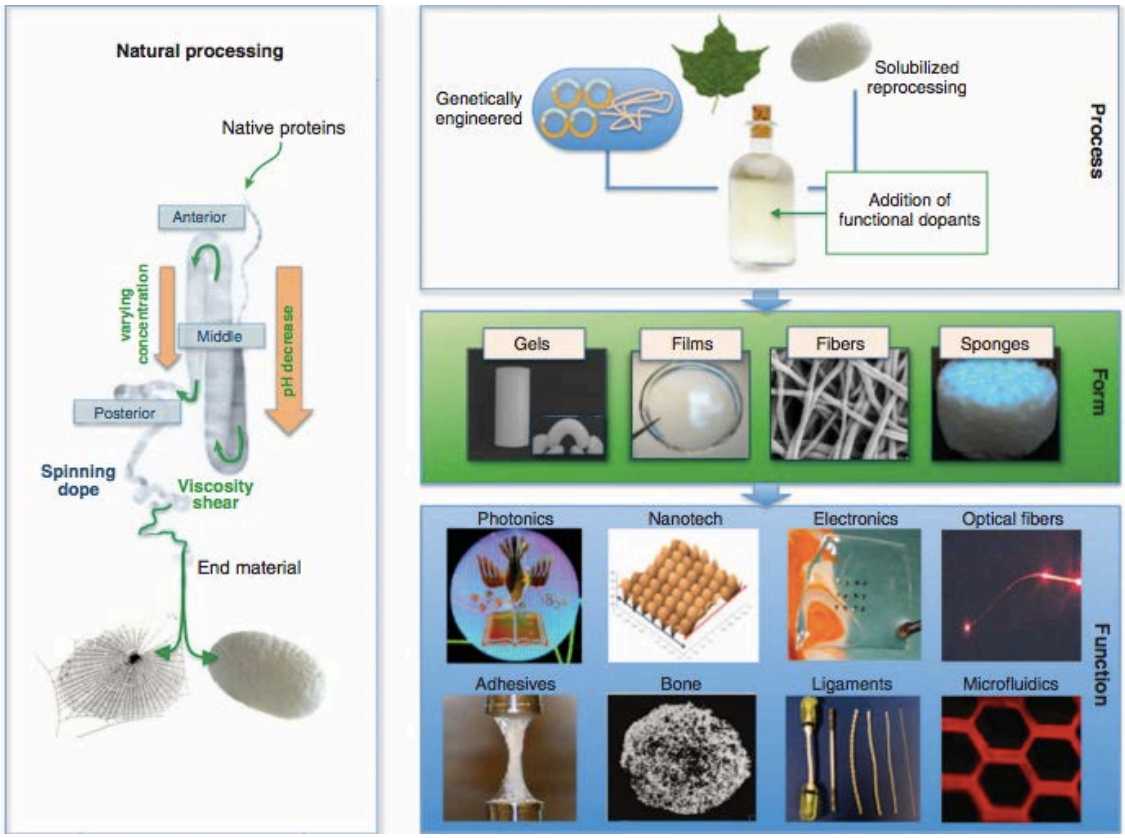

Figure 1-6. Native and synthetic processing of spider silk materials and potential product applications.

While the spider utilizes these proteins for mainly fiber applications, synthetic spider silk proteins have also been solubilized in water to produce environmentally friendly adhesives, films, sponges, lyogels, hydrogels, coatings, and fibrous mats as seen in Figure 1-7[35–37]. When solvated in water, these materials produce no harmful by-products and can be modified to contain antibiotics, antimicrobials, growth factors, etc[37–39]. For comparison, Kevlar production utilizes concentrated sulfuric acid and yields hydrochloric acid as a by-product, which translate to additional costs to store and remove these compounds. Films have traditionally been the most studied material application outside of fibers because they typically require less processing than fibers and have many

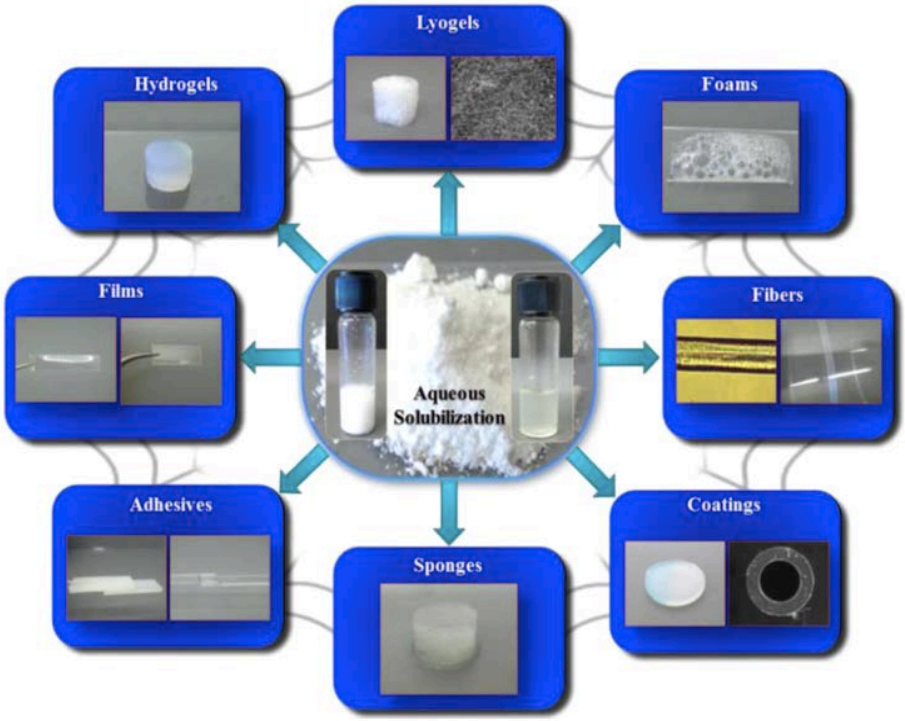

Figure 1-7. Aqueous recombinant spider silk materials

diverse functional applications. Silk film applications include medical device coatings, cell scaffolds, drug delivery systems, and *in vivo* implants[40–44].

To the author's knowledge, there are no published characterizations of films or other materials using synthetic minor ampullate spider silk proteins. The solvation of rSSps in water also has raised the possibility that spider silk may be used in medical and other industries that restrict use of toxic substances. Previously, large and naturally insoluble rSSps had been solvated in 1,1,1,3,3,3-hexafluoroisopropanol (HFIP), a toxic organic solvent that disqualifies the protein products for many industrial uses. These products typically require much less protein than fibers, which increases the cost-effectiveness for material applications. Synthetic MiSps may offer new and interesting

properties to these aqueous materials, allowing them to fill new niches in a variety of applications.

Using the known sequence information from *Nephila clavipes*, this investigation focused on: (i) the first known production of synthetic recombinant minor ampullate silk protein analogs in genetically engineered *E. coli*, (ii) the purification of these silk proteins from the soluble phase, and (iii) mechanical property characterization of thin films derived from the purified proteins.

CHAPTER 2

EXPRESSION AND PURIFICATION OF SYNTHETIC MINOR AMPULLATE SILK PROTEINS

MATERIALS AND METHODS

Design and Construction of Synthetic Minor Ampullate Silk Protein Analogs

The repetitive DNA core sequence and spacer region of MiSp1, based on *Nephila clavipes*, has been sequenced (GenBank:AAC14589.1) and was used for the synthetic constructs of this investigation. Prior to this study, this repetitive sequence was codon optimized and synthesized (Life Technologies). From this baseline construct, the next five constructs were built using molecular biology techniques. Figure 2-1 shows the full protein sequence used for each sequence module of the MiSps in these gene constructs. It matches exactly the repeat consensus sequence and spacer region in MiSp1 of *Nephila clavipes* shown in Figure 1-3. Notice the short poly-alanine stretches, GGX and $(GA)_n$ residues, and non-repetitve spacer and C-terminus. The repetitive consensus sequence is mostly composed of glycine and alanine, with highly conserved tyrosine and glutamine residues[13].

For purposes of this project, six DNA constructs were synthesized and built for expression of minor ampullate silk proteins in *E. coli* (MiSp8, MiSp8s, MiSp16, MiSp16s, MiSp24, and MiSp24s). Each construct was named in reference to how many core MiSp repeats they contained and whether they include the spacer region(s) indicative of these proteins in nature. Referring again to the sequence hierarchy in Figure 1-3, natural MiSps contain 10 repetitive units per repeat. The first synthetic construct will

contain eight repetitive units (1,189 bp) as shown in the MiSp8 region of Figure 2-1. This gene insert and all subsequent spider silk gene inserts were flanked by restriction sites NdeI and BamHI at the N and C-terminus, respectively. The C-terminus region shown in Figure 2-1 had been previously cloned into the chosen expression vector and thus do not need to be included in the construction of the repetitive clones. The cloning vector that was used for the creation and propagation of these constructs was pMK (2829 bp) and TOP10 chemically competent cells (Thermo Fisher) were used for all subsequent transformations during cloning.

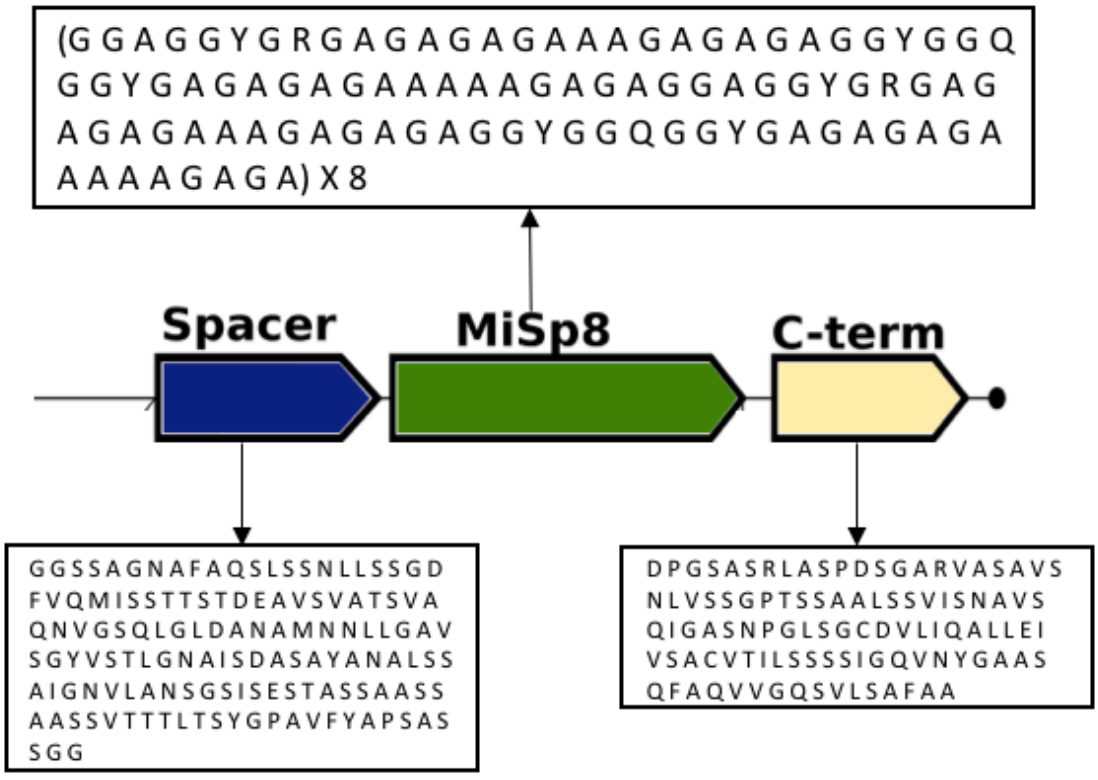

Figure 2-1. Protein sequence information for modules of synthetic minor ampullate silk protein constructs

From this baseline MiSp8 construct shown in Figure 2-1, the next five silk constructs were created using molecular biology techniques and gene synthesis. The 411 bp nonrepetitive spacer region created to match the amino acid sequence of the spacer region of MiSps of *Nephila clavipes* was also optimized for *E. coli* and synthesized at the N-terminus of the MiSp8 construct downstream from the NdeI and AgeI restriction sites and in frame with the repetitive MiSp8 core sequence. This 1,595 bp construct was named MiSp8s.

Taking advantage of the modular and repetitive nature of spider silk, the remaining gene constructs were built from the MiSp8 and MiSp8s base constructs using compatible and nonregenerable restriction sites to extract and combine the repetitive modules resulting in increasingly larger DNA constructs[15]. Double digests were first performed on two aliquots of the pMK-MiSp8 plasmid in order to build pMK-MiSp16, using AgeI and NcoI on one aliquot and NcoI and BspeI on the other. The NcoI restriction site interrupts the kanamycin resistance gene in the pMK cloning vector. Following these digests, the two larger gene pieces from each of these digests (2,407 bp and 2,255 bp) were extracted using a QIAquick gel extraction kit. Once extracted and purified, these oligos were ligated using quick ligase (NEB) and transformed into One Shot TOP10 chemically competent cells per the manufacturer's directions (Thermo Fisher). Colonies were incubated overnight at 37°C on an agar plate containing kanamycin (100 mg/ml) for additional selective pressure. Single colonies were transferred to 5 ml of LB media and incubated at 37°C for 12 hours. Plasmids were purified using a QIAGEN Plasmid Miniprep kit. Digestion checks with BamHI and NdeI were performed on the purified

plasmid to verify a successful ligation following agarose gel electrophoresis (0.2% agarose) run for 35 minutes at 100 V and stained with ethidium bromide. Plasmids that digested as expected were sequenced using the T7 forward and reverse universal primers (ACTG Inc.) to confirm. Successful ligation of this construct resulted in a complete 4,662 bp plasmid that contained two sets of eight repeat units (MiSp16) in the pMK cloning vector. The mechanisms of construction are shown in Figure 2-2. Restriction enzymes AgeI and BspeI are compatible and nonregenerable. Therefore, following ligation these enzymes only cut at the N- and C-termini of the full MiSp16 repetitive sequence, because the site in the middle of these repeats has been eliminated. This method using compatible, nonregenerable restriction sites allows for the construction of larger MiSps. This basic protocol was repeated another three times to yield the final three MiSp constructs (MiSp16s, MiSp24, MiSp24s). Double digests of AgeI with NcoI and BspeI with NcoI were repeated on two aliquots of pMK-MiSp8s (3,900 bp). The larger DNA pieces from each digest (2,824 bp and 2,672 bp) were extracted and ligated to yield the complete pMK-MiSp16s plasmid (5,496 bp). This ligation resulted in two MiSp8 repeats separated by two spacer regions (3,203 bp) because the spacer region was synthesized downstream of the AgeI restriction site in the pMK-MiSp8s construct to make it compatible with this protocol. To build pMK-MiSp24, these same double digests were performed again on an aliquot of pMK-MiSp8 and an aliquot of pMK-MiSp16 and the two larger pieces of each digest (3,587 bp and 2,255 bp) were extracted, purified, and ligated. Finally, to build pMK-MiSp24s, these double digests were performed on two aliquots of MiSp8s and

MiSp16s. Again, the two larger pieces from each double digest were extracted, purified,

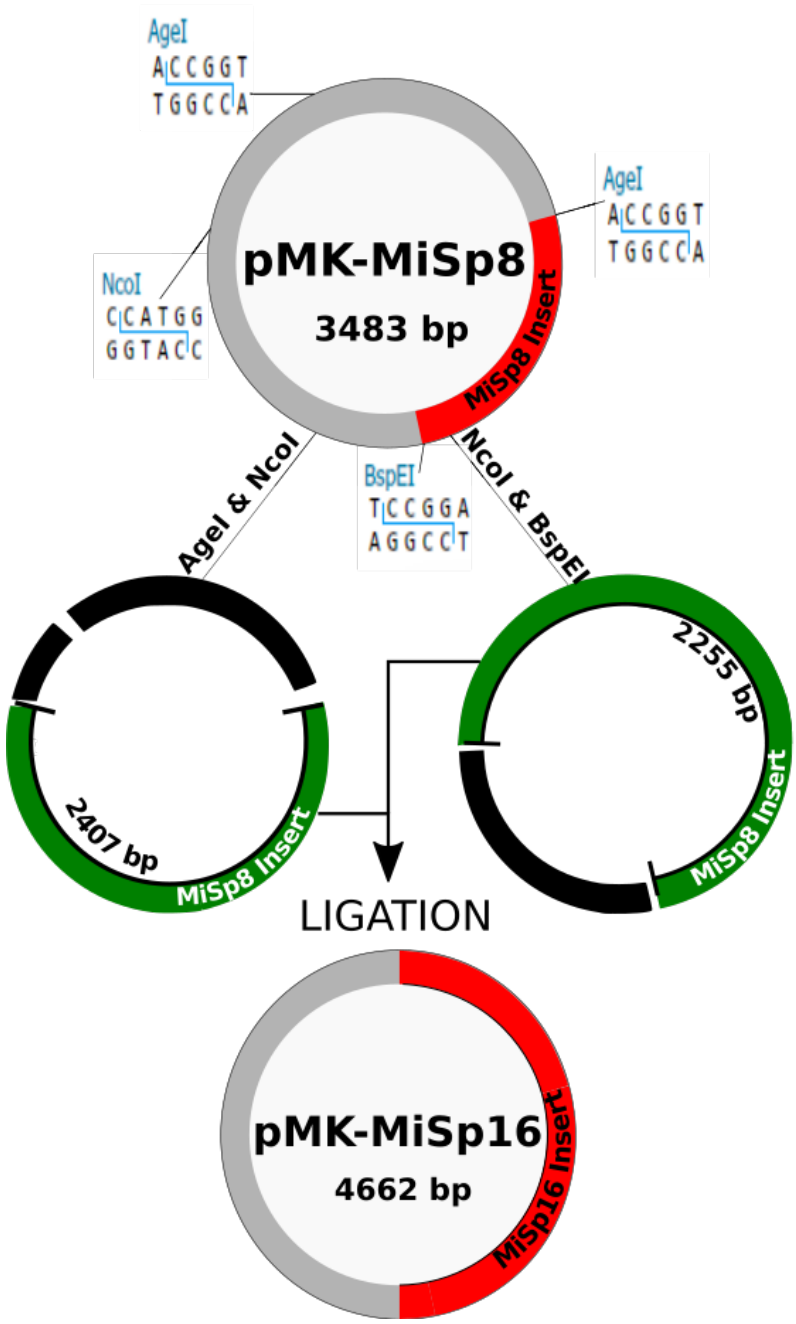

Figure 2-2. Modular build of repetitive MiSp constructs illustrating the build from pMK-MiSp8 to pMK-MiSp16

and ligated (4,421 bp and 2,671 bp).

Once verified by digestion checks and sequencing, these synthetic spider silk constructs were digested out of the pMK cloning vector using enzymes BamHI and NdeI and ligated into a modified pET19 series expression vector that contains the nonrepetitive spider silk C-terminus region and a 6X-histidine tag. The C-terminus region was cloned into this vector before this project and was provided for this research (publication pending). This in-house vector is named pET19kt (~6,000 bp).

A helper plasmid was also transformed into a stock of chemically competent BL21(DE3) cells, which codes for three glycine tRNAs and chloramphenicol resistance as part of a modified pACYC vector (4,245 bp). This cell stock was used for all subsequent transformations for both flask studies and benchtop bioreactor runs. These glycine tRNAs were expected to ease the burden of translation for the glycine-rich repetitive constructs, because previous studies have identified that in rSSp production, *E. coli* cells exhibited up-regulation for a glycine biosynthetic enzyme and the β-subunit of glycyl-tRNA synthetase, indicating cellular stress due to the expression of these highly repetitive silk sequences[45].

Expression of Minor Ampullate Silk Proteins

Preliminary Flask Studies

The pET19kt-MiSp plasmids constructed for protein expression were pET19kt-MiSp8, pET19kt-MiSp8s, pET19kt-MiSp16, pET19kt-MiSp16s, pET19kt-MiSp24, and

pET19kt-MiSp24s. A BL21(DE3) chemically competent cell stock previously transformed with a plasmid coding for three glycine tRNAs and chloramphenicol resistance was made and used for transformation with the MiSp constructs. Preliminary investigations included flask studies for all six of the synthetic MiSps to verify protein expression and examine the expression capabilities for each protein. One liter flask studies were done for all six proteins using Terrific Broth media (yeast 24g/L, glycerol 4 ml/L, tryptone 20 g/L, potassium phosophate monobasic 0.17 M, potassium phosphate dibasic 0.72 M) and the antibiotics chloramphenicol (60 mg/ml) and kanamycin (100 mg/ml). Flasks were inoculated from a single-colony 10 ml overnight culture and grown at 37°C in a baffled shaker flask at 220 RPM to an OD_{600} of between 0.6 and 0.8. Protein expression was induced with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Before induction, the temperature of the culture was lowered to 30°C, 25°C, or 22°C, to mitigate truncation and result in more full-size, spider silk protein in the soluble fraction[38]. The culture was grown for four hours after induction at each of these temperatures and hourly samples were taken. Protein expression was analyzed for each protein by SDS-PAGE and Western blot using a 6X Anti-His primary antibody (Rockland Scientific) and an Anti-Mouse IgG (H+L) secondary antibody (Rockland Scientific) or a custom rSSp C-terminus antibody with a Donkey Anti-Rabbit IgG (H+L) secondary antibody (Rockland Scientific). Blotting, antibody exposure time, and antibody concentration were according to manufacturer's instructions. To prepare the samples for SDS-PAGE, hourly samples were standardized to a specific OD (50 for high-density culture samples) and this cell suspension was then sonicated with a Qsonica Q500

sonicator using a microtip (1/8” diameter) at 1A for 15 seconds and then centrifuged in a benchtop centrifuge (Beckman Coulter Microfuge 18) at 18,000 rcf to yield a soluble fraction and an insoluble fraction. The soluble fraction was added directly to 2x Laemmli sample buffer (Bio-Rad) at a 1:1 ratio and heated at 100°C for 15 minutes prior to loading. The insoluble fraction was resuspended in 8M urea and sonicated again for 15 seconds at 1A and then centrifuged to remove any remaining insoluble material. The clarified solution was added to 2x Laemmli sample buffer at a 1:1 ratio and heated for 15 minutes at 100°C before sample loading.

Benchtop Bioreactor Experiments

As with the flask study experiment, a BL21(DE3) cell stock previously transformed with the glycine helper plasmid was used for transformation with the pET19kt-MiSp constructs. A New Brunswick BioFlo/CelliGen 115 benchtop bioreactor (BF115) was used for all high-density culture experiments for the expression of the MiSps. This 2 L vessel has a 1 L working volume and was assembled for fed-batch aerobic fermentation per the manufacturer’s directions. Figure 2-3 shows the assembly of the reactor. The reactor was autoclaved before use. Vessel media components included glucose (25 g/L), trace metals (1x), magnesium sulfate (.002 M), yeast (5 g/L), phosphates (10 g/L), thiamine (2.5 g/L), and 0.3% v/v antifoam. The antibiotics kanamycin (100 mg/ml) and chloramphenicol (60 mg/ml) were added to both the vessel media and the feed media. 510 ml of feed media was made from an autoclaved 70% glucose solution at a volume of 220 ml. Additional feed media components included Hy-Express System II (0.6% v/v),

Figure 2-3. Benchtop bioreactor setup for MiSp expression

magnesium sulfate (.012 M), thiamine (.0124 g), trace metals (0.6x), and antibiotics kanamycin (100 mg/ml) and chloramphenicol (60 mg/ml)). A 100 ml seed culture inoculated with the pET19kt-MiSp construct was grown to an OD_{600} of ~3 at 37°C and 220 RPM and used to inoculate the BF115 bioreactor for a starting OD_{600} ~0.3. New Brunswick Biocommand software was used to monitor the aerobic fermentation and programmed to start the exponential feed pump when the cells ran out of glucose (pH spike combined with a dissolved oxygen drop).

A cascade program was set to control oxygen input, air flow, agitation, and pH dependent on probe readout values. A separate feed pump of ammonium hydroxide was used to maintain pH and was controlled by this cascade program. The fixed parameters

for growth were a pH of 7.0 and a dissolved oxygen saturation of 85%. The cascade program controlled the oxygen input and agitation to ensure 85% dissolved oxygen in the vessel was maintained. The culture was allowed to grow at 37°C for ~8 hours or until an OD_{600} of ~80 was reached. At this point, temperature in the vessel was decreased to 22°C and the culture was induced with a final concentration of 1 mM IPTG. Following induction, glucose concentration was kept between 5 and 10 g/l and hourly samples were taken. The aerobic fermentation was continued for 4 hours after induction.

Hourly samples were standardized for a final OD_{600} of 25 by dilution with lysis buffer (2 M urea, 20 mM Tris-HCL, 0.5 M NaCl, 2 mM EDTA) for SDS-PAGE and Western blot analysis. Samples were prepped using the same protocol as above for the flask study experiments.

Purification of Synthetic Minor Ampullate Silk Proteins

Purification of the four synthetic MiSps expressed in *E. coli* was achieved using three main purification methods following an initial extraction and purification step. Due to the different molecular weights and sequence information of these recombinant proteins, purification methods had to be adapted and optimized for each construct. Moreover, purity, solubility, and yield were factors that rendered certain methods unviable. Purification methods included immobilized metal affinity high performance liquid chromatography, fractional ammonium sulfate precipitations, and sequential precipitations using ethanol, isopropanol, and acetone.

MiSp Extraction and Initial Purification Steps

Following the harvesting of the cells from the BioFlo/Celligen 115 bioreactor, the cell slurry was centrifuged at 11,652 rcf for 45 minutes. The cell pellet was collected and the media was discarded. This cell pellet was weighed and 1.5x wt/vol of a urea lysis buffer at a pH of 8 was used to suspend the cell pellet (2 M urea, 20 mM Tris-HCL, 0.5 M NaCl, 2 mM EDTA). This suspension was sonicated with a QSonica sonicator with 3/4" diameter probe at 40 A for 45 to 60 minutes or until mixture was homogenous and cell lysis was completed. This mixture was then centrifuged at 11,652 rcf for 45 minutes and the supernatant was collected in a 1 L glass bottle. The lysate was then heated in an 80°C water bath for 15 minutes with intermittent mixing to denature and precipitate *E. coli* proteins. Polyethylenimine (PEI) was added at a final concentration of 0.2% v/v and allowed to mix for 5 minutes to precipitate nucleic acids and then centrifuged for 1 hour at 11,652 rcf. The supernatant was collected and stored at 4°C.

Nickel Affinity High Performance Liquid Chromatography

The polyhistidine (6X) tag at the N-terminus of the MiSp constructs was utilized for immobilization on a Ni-NTA resin using an AKTA Avant chromatography system. A GE XK-16 column was packed for a bed height of 14 cm. To maximize protein solubility, accessibility of the His-tag, and to mitigate protein-protein interactions MiSp sample solutions were prepared using a salt buffer with 2M urea, 8M urea, and/or 50% glycerol. Running buffer compositions were adjusted accordingly to minimize rapid concentration changes on the column. A custom program was created for the purification

of MiSp8 using Unicorn 6.0 workstation software. This program set sample injection rate to 1 ml/minute for slow application of sample onto the column and the flow through was collected. The column was washed with running buffer (0.5 M NaCl, 20 mM Tris-HCL, 0.5 mM EDTA) and wash fractions were collected. 5 ml fractions were collected for 3 column volumes at 5%, 10% and 100% imidazole elutions. Samples of these fractions were mixed 1:1 with 2x Laemmli sample buffer, heated for 15 minutes at 100 °C and analyzed by SDS-PAGE and western blot to confirm MiSp in solution. The His-tag antibody was used in western blotting to confirm MiSp in solution.

Following SDS-PAGE analysis to confirm MiSp in the fractions that elicited a peak in the chromatographic spectra, both dialysis and ammonium sulfate precipitations were performed to precipitate the MiSps. Ammonium sulfate was slowly added to the chromatography fractions for a final saturation of 15% at 4°C. This was mixed for 35 minutes and centrifuged for 1 hour at 11,652 rcf and the pellet was collected. The pellet was washed and agitated in diH_2O until the conductivity of the supernatant was <10 microsiemens and lyophilized for storage.

As an alternative to ammonium sulfate precipitation, dialysis was performed to remove the salts from the chromatography fractions and precipitate the MiSps as well. 30 kDa dialysis tubing (Sigma-Aldrich) containing the MiSp fractions was suspended in a large volume of stirred ddH20 solution that was replaced every 6-8 hours for ~3 days or until the conductivity of the dialysate remained constant. The sample was then centrifuged for 1 hour at 11,652 rcf and the pellet was lyophilized for storage.

Fractional Ammonium Sulfate Precipitations

Following the PEI and heat treatments, fractional ammonium sulfate precipitations were done to target a saturation point where the MiSp precipitated, but left most other soluble proteins and non-protein impurities in solution. Following the heat and PEI treatments, ammonium sulfate was mixed with the clarified lysate to a final saturation of 5% at 4°C, centrifuged at 11,652 rcf at 4°C for 45 minutes and the pellet was discarded. The remaining supernatant was mixed with ammonium sulfate at 4°C for a final saturation of 15%. The mixture was then centrifuged for 1 hour at 11,652 rcf and 4°C. The pellet was collected and washed with diH_20 until conductivity fell below 10 microsiemens. The pellet was then lyophilized for storage and dry weight calculations.

Ethanol/Isopropanol/Acetone Precipitations

As an alternative precipitation method, ethanol, isopropanol (IPA), and acetone precipitations were investigated to attempt to increase the solubility of the precipitated protein without sacrificing purity.

Clarified lysates pre-treated with heat and PEI, containing MiSp8, MiSp8s, MiSp16, and MiSp16s were first mixed with varying concentrations of cold ethanol, centrifuged at 4°C and 8,500 rpm for 45 minutes, and tested by SDS-PAGE for presence of MiSp in both the soluble and insoluble fractions. In the case of the larger proteins, fractional precipitations using cold IPA and acetone were needed to precipitate the protein as a final purification step. After SDS-PAGE analysis, protein pellets were

washed with 70% IPA until supernatant conductivity was below 10 microsiemens. Pellets then were lyophilized for storage and dry weight calculations.

Preliminary Characterization of Synthetic Minor Ampullate Silk Proteins

Aqueous Recombinant Minor Ampullate Silk Films

Lyophilized MiSp protein pellet was weighed and added to distilled water for a final concentration of 6% w/v in a sealed glass vial. The MiSps were solubilized in water using high heat and pressure as described in previously published methods using a microwave[35]. Propionic acid (99%) at a concentration of 1% v/v was added to the dope if the protein exhibited poor solubility. Once the silk was in solution, the dope was carefully spread onto clean 7x30 mm PDMS strips and allowed to dry (~400 ul of dope per strip). The films were peeled off the PDMS strips for mechanical testing. No additional post-treatments were used on films before testing due to the small sample sizes at this project stage; however, it is likely that additional treatments such as film stretching may help enhance the film properties[36,46]. The films were measured for gauge length, width, and thickness, then glued to plastic C-card testing strips across an 8 mm gap as described in previously published methods[36]. Tensile testing was done using an MTS Synergie 100 machine with a 50 N load cell. The C-card was positioned on the MTS machine clamped on both ends. The C-card was cut before testing so that only the film was being tested. Tensile data was collected using MTS TestWorks 4 (2001) and analyzed using Microsoft Excel for stress and strain values.

Fourier Transform Infrared Spectroscopy

FTIR spectra were recorded using dry protein powder samples on an Agilent 660 instrument with a diamond ATR accessory. The spectra were collected with 32 scans and a resolution of 4 cm^{-1}. OriginPro 2018 software was used for baseline correction and deconvolution of the spectra. The Amide I band was fitted using three Gaussian peaks centered at 1620 cm^{-1} (β-sheet), 1650 cm^{-1} (α-helix/random coil), and 1698 cm^{-1} (β-sheet), in accordance with the literature[15,19].

Statistical Analysis

Mechanical properties were analyzed for significance using a two-tailed t-test assuming unequal variance with a null hypothesis that sample means were equal. A p-value of <0.05 was deemed significant.

RESULTS AND DISCUSSION

Synthetic Minor Ampullate Silk Protein Constructs

Using the build protocol outlined above in Figure 2-2, the double digests performed to build the complete MiSp contructs are shown in Figure 2-4. These gel extractions show the DNA pieces that were extracted and ligated together to create the complete MiSp plasmids.

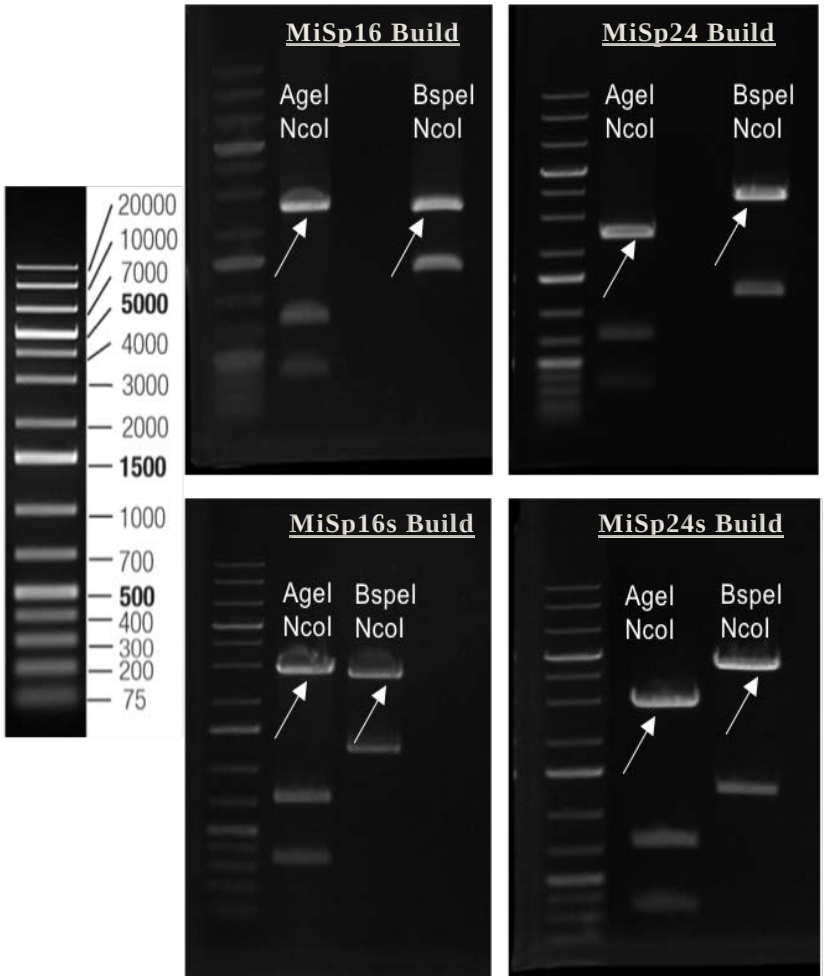

Figure 2-4. Gel extractions used to build the pMK-MiSp constructs

Using the build protocol outlined above in Figure 2-2, the double digests performed to build the complete MiSp contructs are shown in Figure 2-4. These gel extractions show the DNA pieces that were extracted and ligated together to create the complete MiSp plasmids.

Following ligation and verification by sequencing, the complete MiSp gene constructs were removed from the pMK cloning vector and ligated into the pET19kt expression vector. The digested expression plasmids using BamHI and NdeI are shown in

Figure 2-5 with the pET19kt vector at ~6000 bp and the MiSp insert running below it for each of the six inserts. Figure 2-6 shows the coding regions of these final minor ampullate silk protein constructs, with final protein molecular weights ranging from 40 to 134 kDa.

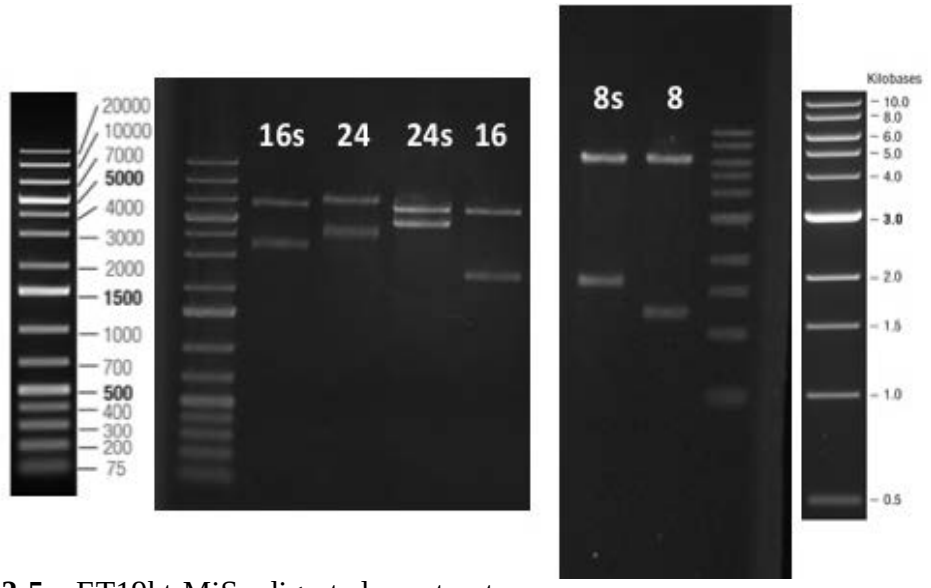

Figure 2-5. pET19kt-MiSp digested constructs

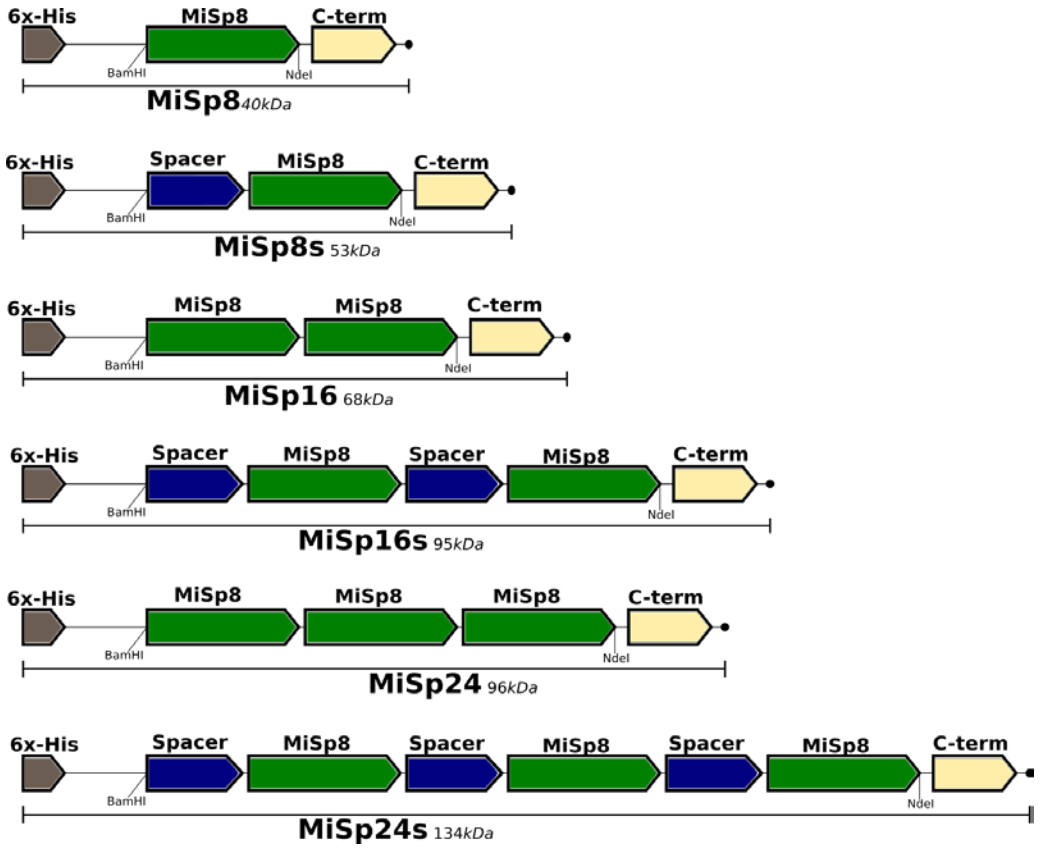

Figure 2-6. Coding regions for synthetic minor ampullate silk proteins

Expression of Synthetic Minor Ampullate Silk Proteins

The goals of the high-density cell culture outlined in this thesis were to maximize the amount of MiSp in solution (soluble fraction) and minimize truncation so that the majority of this protein was full-size MiSp. Flask culture studies were effective in understanding how well these protein constructs expressed at low cell densities and how the temperature after induction affected the expression levels of these large and repetitive

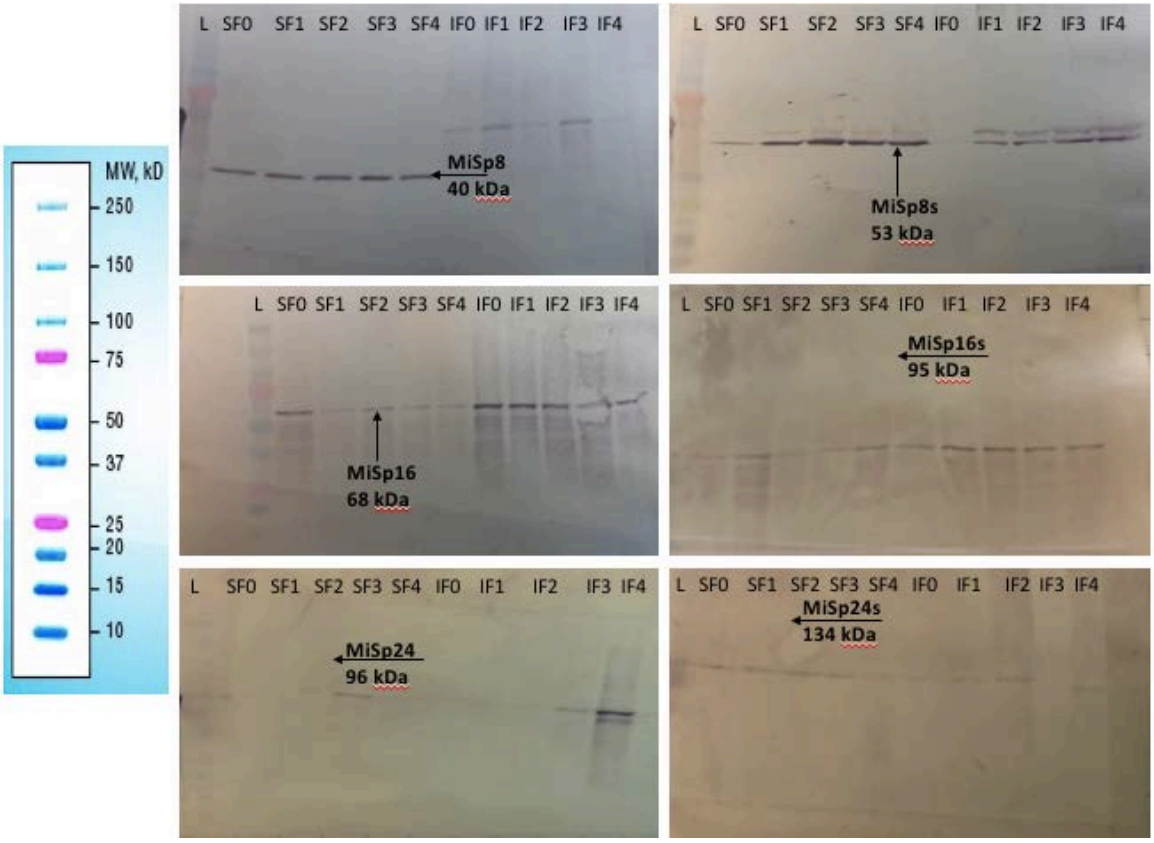

Figure 2-7. Western blots of flask study hourly samples to determine expression levels and solubility of MiSps; SF and IF denotes the soluble and insoluble fraction at each hour after induction; arrow indicates full-size MiSp

proteins. The flask study experiments also yielded information on the solubility of the protein constructs and how much of the expressed protein is stable in the soluble fraction after translation. Figure 2-7 shows the western blot analysis for the hourly samples taken during these flask study experiments when the temperature at induction was lowered to 30°C. Lanes 1-5 (SF0-SF4) on each membrane show the soluble fractions, while lanes 6-10 (IF0-IF4) show the insoluble fraction of each cell sample for each MiSp construct. At the flask scale with the Terrific Broth media, OD_{600} varies between 1 and 3, so slight

changes in protein expression and solubility are more readily noticeable than at high cell densities. As expected, expression for the two smallest protein constructs (MiSp8 and MiSp8s) is much higher than for the other four MiSp constructs of larger molecular weight as seen in the top row of Figure 2-7. Furthermore, as the constructs get larger, full-size protein is either not present, present in low quantities, or most of the full-size protein is in the insoluble fraction. The largest construct that saw expression at this 30°C induction temperature was MiSp16 at 68 kDa (second row left in Figure 2-7). A small amount of full-size protein is present in the soluble fraction, with the majority in the insoluble fraction. MiSp16s, MiSp24, and MiSp24s did not express at full size at an induction temperature of 30°C and the bands present in the expression checks for these constructs are highly truncated versions of the MiSp. At the flask scale, full-size expression was not seen for MiSp24 and MiSp24s at any of the induction temperatures tested. However, when the induction temperature was dropped to 22°C, full-size expression was seen for MiSp16s (95 kDa) as shown in Figure 2-8, but MiSp16s still showed significantly less expression than MiSp16. At this point, it was unclear whether enough MiSp16s would be produced for purification even after high density culture. Interestingly, the C-terminus antibody was also much more effective in detecting MiSp in solution at low cell densities than the His-tag antibody. This was the first indication that the His-tag at the N-terminus of the MiSp is not as readily available as the C-terminus when the proteins are folded in solution.

In the interest of time and the goals of this thesis, benchtop bioreactor runs were only performed for MiSp8, MiSp8s, MiSp16, and MiSp16s because full-size protein was

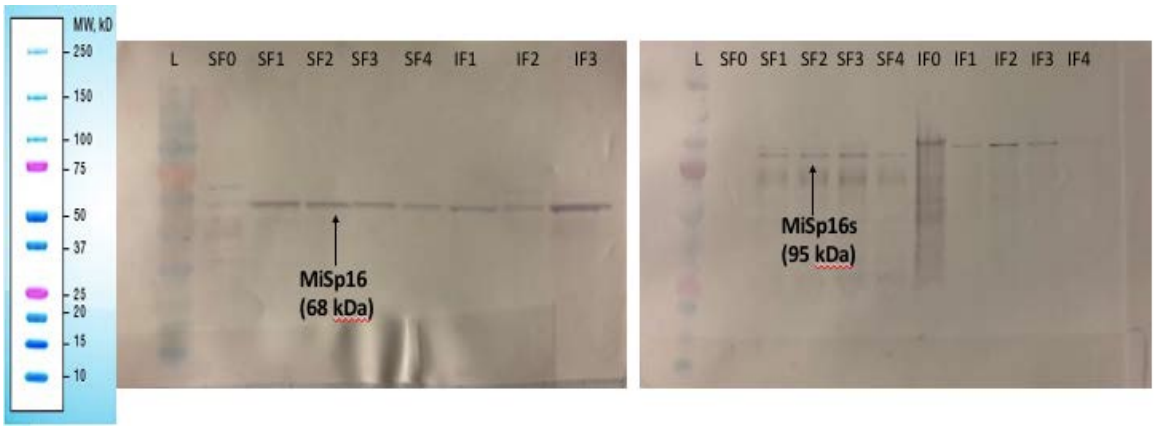

Figure 2-8: Western blot of hourly samples from flask study expression checks for 22°C induction; SF and IF denote soluble and insoluble fractions at each hour after

not expressed at any induction temperature for the MiSp24 and MiSp24s constructs at the flask scale. It is possible that the optimized expression of these constructs will involve the investigation of alternative expression conditions and molecular machinery that may nclude a different vector system[28,47]. The flask studies show that lower induction temperatures are necessary especially for the larger MiSp constructs, although the expression levels for MiSp16 and MiSp16s are lower than expected overall. The benchtop bioreactor experiments were used to analyze how high-density culture with OD_{600} values ranging from 80 to 100 may affect these expression conditions. Conditions for growth in these experiments were kept constant, with the exception of temperature and growth time after induction. Figures 2-9 and Figure 2-10 show expression at the benchtop bioreactor scale for MiSp8, MiSp8s, MiSp16, and MiSp16s.

At this scale, , MiSp8, MiSp8s, MiSp16, and MiSp16s are expressing at full size in the soluble fractions. As molecular weight increases, truncation becomes a more significant issue and there is less full-size protein in the MiSp16 and MiSp16s soluble

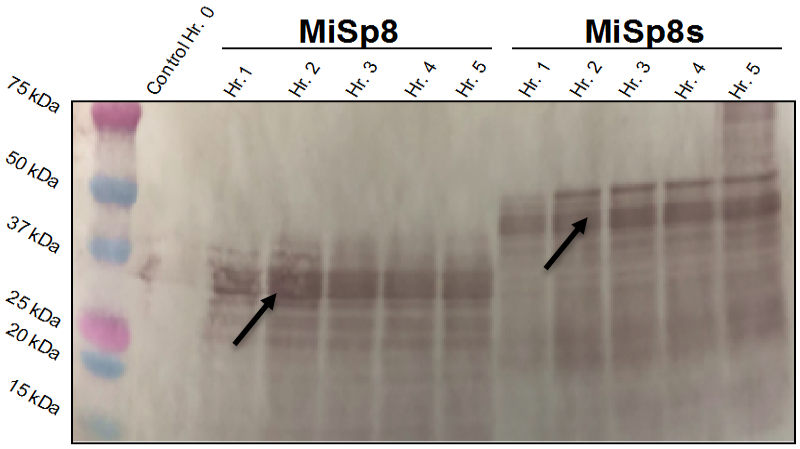

Figure 2-9. MiSp8 and MiSp8s hourly protein expression at the bioreactor scale in the soluble fraction standardized to OD_{600} 25

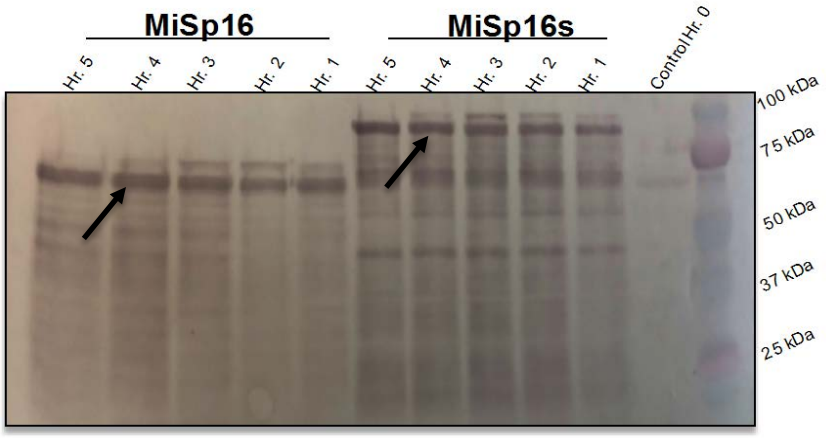

Figure 2-10. MiSp16 and MiSp16s hourly expression in the soluble fraction at the bioreactor scale; samples standardized to OD_{600} 25

fractions. As molecular weight increases, truncation becomes a more significant issue and there is less full-size protein in the MiSp16 and MiSp16s soluble fractions than for MiSp8 and MiSp8s. For the MiSp16s expression, there is a higher concentration of the

truncated products overall. The His-tag antibody was used for all western blots at the bioreactor scale because the His-tag is useful for detecting truncated forms of the MiSp. However, especially in the case of the smallest constructs (MiSp8 and MiSp8s), the His-tag antibody more readily attached to MiSps that are running ~10 kDa lower than full size as seen in Figure 2-10. When the cell lysate samples are run with the C-terminus antibody, two distinct conformations of MiSp are seen: one running at full size on the SDS-PAGE, and one running ~10 kDa low as seen in the MiSp8s expression checks in Figure 2-11. Because the C-terminus is evident in both conformations, it is reasonable to assume both are full-size MiSp. Two other possible explanations could be that the protein was degraded after translation or that the custom rSSp C-terminus antibody is not specific enough. Degradation of the protein at the consistency seen in this high-density culture is unlikely because only two conformations are seen when the C-terminus antibody is used and they are seen in relatively high concentrations. Furthermore, the His-tag attaches to both conformations of the protein in solution; however, it attaches to the lower band with a much higher affinity as seen in Figures 2-9 and 2-10. Figure 2-11 also illustrates how the C-terminus antibody does not show the truncated versions of the protein seen when the His-tag antibody is used as in Figures 2-9 and 2-10.

To rule out the possibility that the lower band is an *E. coli* protein impurity that binds to the C-terminus antibody, a blank BF115 benchtop bioreactor run was performed under the same constraints that were used for the MiSp aerobic fermentations. Expression checks were then performed on the hourly samples and a western blot was done using

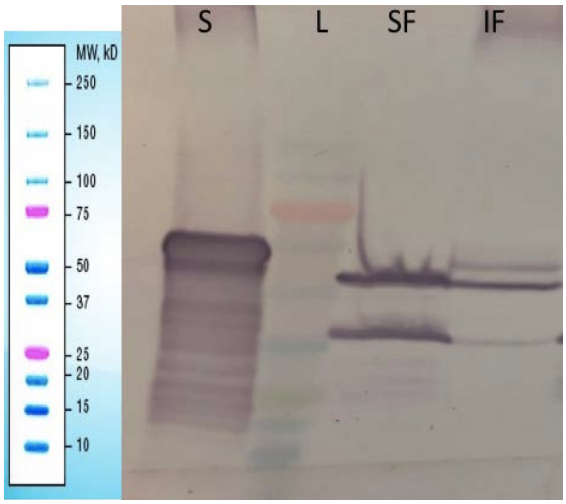

Figure 2-11. Western blot using C-terminus antibody; Lane S: 65 kDa MaSp1 (M4) standard; Lane L: ladder; Lane SF soluble fraction MiSp8s lysate; Lane IF- insoluble fraction from MiSp8s lysate

both the His-tag antibody and the C-terminus antibody. Figure 2-12 shows the soluble and insoluble fractions of this "blank" expression test.

The C-terminus antibody did attach to an *E. coli* impurity in the insoluble fraction of the cell lysate; however, no signal was seen in the soluble fraction as is evident in (A) of Figure 2-12. Therefore, it can be determined that the two bands seen in the expression checks of the MiSp high-density culture are both the target spider silk proteins. Additionally, the purification of MiSp involves the isolation of only the soluble fraction, so this *E. coli* impurity with an affinity to the C-terminus will be removed or mostly removed in the first centrifugation. Further investigation showed that the C-terminus antibody is less specific when used for insoluble pellet samples. Therefore, all additional pellet samples were analyzed using the His-tag antibody because this antibody did not result in any signal in either of the fractions as seen in (B) of Figure 2-12.

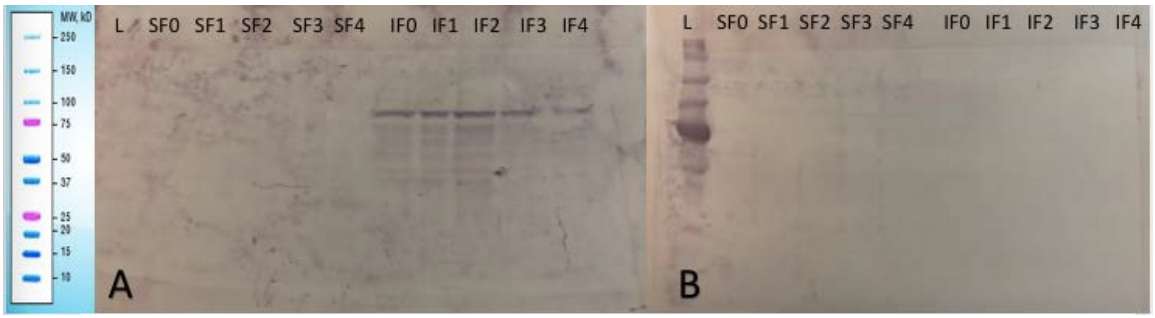

Figure 2-12. Hourly expression checks of soluble and insoluble fractions for blank BF115 run. SF: soluble fraction, IF: insoluble fraction; (A) Western blot using C-terminus antibody. (B) Western blot using His-tag antibody.

Purification of Synthetic Minor Ampullate Silk Proteins

MiSp Extraction and Initial Purification Steps

MiSp8, MiSp8s, MiSp16, and MiSp16s were all extracted from the cell pellet using the protocol described above. Each of these constructs was tested for their stability at 80°C and with the addition of PEI to precipitate *E. coli* proteins and nucleic acids. These treatments were necessary to substantially clean up the lysate in order for future purification steps to be successful. Figure 2-13 illustrates this initial purification showing the MiSp8 (40 kDa) lysate supernatant before and after the PEI and heat precipitations. As seen in Lane 1 of Figure 2-13, the PEI and heat treatment are highly effective in removing most of the protein impurities present in the lysate, leaving the target protein as the most concentrated protein in solution.

These precipitations did result in minor protein losses for the MiSp8 (40 kDa), MiSp8s, (53 kDa) and MiSp16 (68 kDa) constructs and more severe losses for the

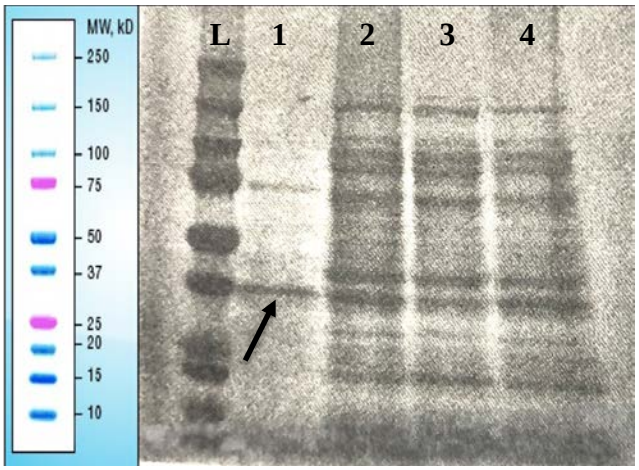

Figure 2-13. MiSp8 SDS-PAGE showing lysate initial purification with heat and PEI. Lane 1- MiSp8 supernatant post-PEI treatment; Lane 2- MiSp8 pellet post-PEI treatment; Lane 3- MiSp8 supernatant post-lysis with no treatment; MiSp8 pellet post-lysis with no treatment

MiSp16s (95 kDa). While a fraction of MiSp16s protein is still present after PEI and heat treatments, the concentration is substantially reduced when compared with the other MiSp constructs, indicating that MiSp16s is not as stable in the soluble fractions as the other three MiSps. Figure 2-14 shows the difference in concentration between MiSp16 and MiSp16s after PEI and heat treatments. This stability difference is notable and is most likely due to the high molecular weight of the protein and not the presence of the two spacer regions in this construct. You can see in Figure 2-14 that after PEI and heat the majority of protein in solution is no longer the full-size MiSp16s depicted by the arrow. By comparison, MiSp16 (68 kDa) showed good stability in solution after lysis, heat, and PEI treatments and the majority of MiSp is full size, with acceptable levels of truncated product still present.

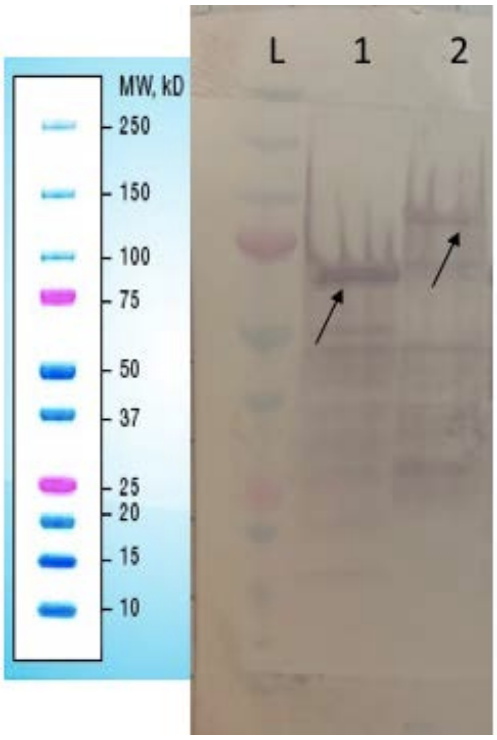

Figure 2-14. Lysate processing using heat and PEI as an initial purification. Lane 1: MiSp16 (68 kDa) soluble fraction after heat and PEI treatments. Lane 2: MiSp16s (95 kDa) soluble fraction after heat and PEI treatments

The amount of full-size MiSp16 and MiSp16s produced by the cells in high density culture was substantially lower than the smaller MiSp8 and MiSp8s constructs as described in this chapter. The losses experienced by MiSp16s in this initial purification step made downstream processes less viable for this construct from both a purity and a yield standpoint. This initial purification step was constant for each of the MiSp constructs going forward. The downstream processes and overall purifications for each of the MiSps are shown in Figure 2-15.

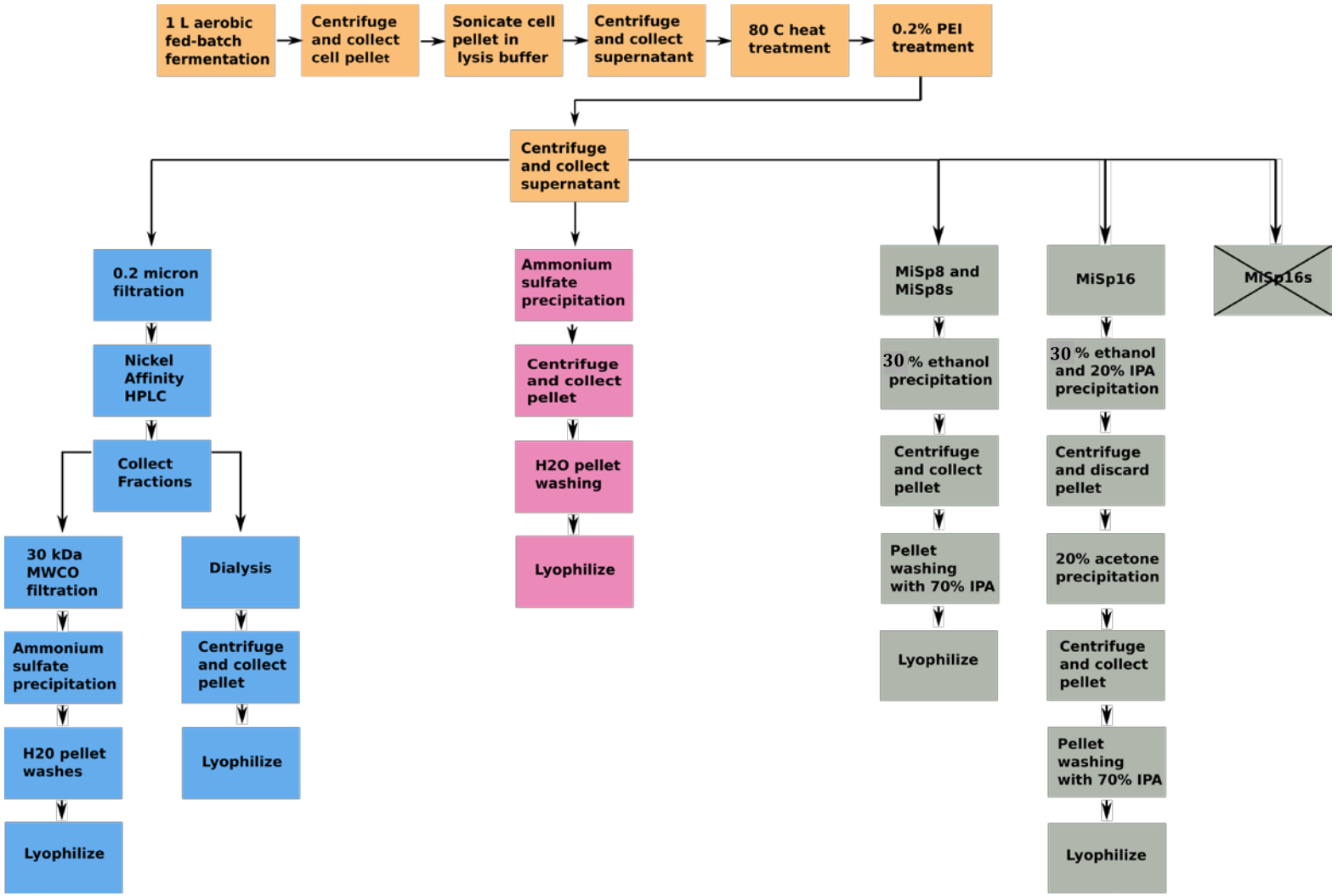

Figure 2-15. Purification strategies used in synthetic minor ampullate silk protein purification

Nickel Affinity High Performance Liquid Chromatography

Taking advantage of the 6X histidine tag at the N-terminus of the MiSp constructs, MiSp8 (40 kDa) and MiSp8s (53 kDa) were purified using this chromatography method as they were soluble. These constructs were chosen as the trial proteins for this purification method because their smaller size and increased expression levels allowed for the best-case scenario in terms of yield and solubility of the four MiSp constructs being expressed in the BF115 reactors. Previous studies have identified metal affinity chromatography as a suitable small-scale purification method for synthetic spider silk protein; however, most of these studies involved major ampullate spidroins[25,29,45,48,49].

The fractions collected after stepwise imidazole elutions from the column showed full-size MiSp8 and MiSp8s in solution. Figure 2-16 shows the SDS-PAGE analysis of the flow-through, wash, and elution fractions for the purification of MiSp8s.

There are substantial amounts of low molecular weight impurities in the eluted fractions from the nickel column, as well as one high molecular weight impurity at ~90 kDa. These impurities are only seen on the Coomassie blue stained SDS-PAGE as they do not show on a western blot analysis with the His-tag antibody, indicating that these impurities are not full-size spider silk protein and should not bind independently to the column. These issues persisted after repacking the column with fresh resin and troubleshooting with a series of both MiSp8 and MiSp8s purifications.

Furthermore, the impurities could not be removed from the fractions by centrifugal filtration, pH precipitation, or dialysis. This indicated that it was probable

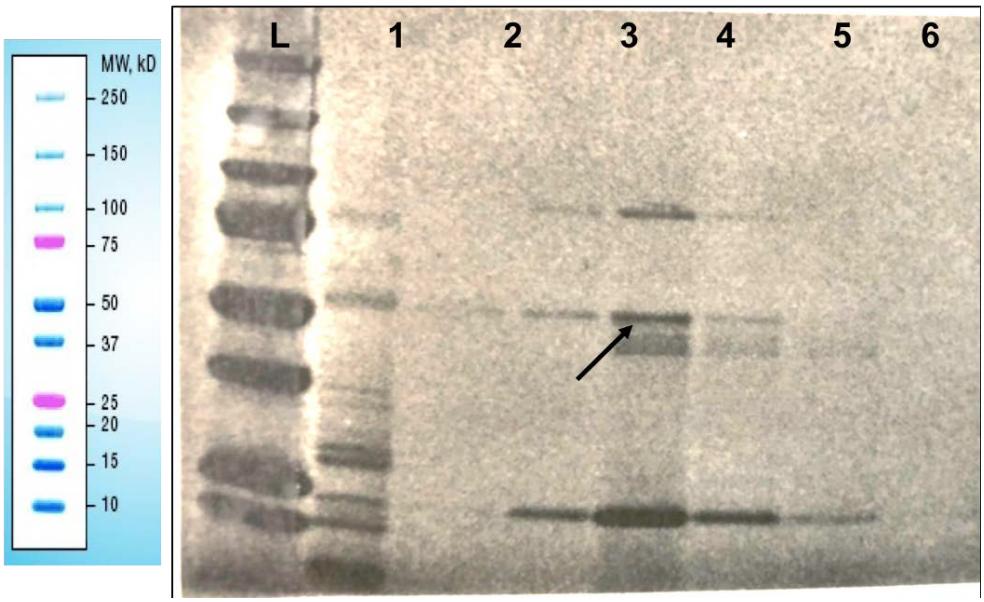

Figure 2-16. SDS-PAGE of MiSp8s nickel affinity chromatography purification. Arrow indicates full-size MiSp8s. Lane 1: column flow through. Lane 2: column wash. Lane 3: 5% imidazole elution. Lane 4: 10% imidazole elution. Lane 5-6: 100% imidazole elution.

these protein impurities are interacting with the MiSp in solution, allowing them to bind and be eluted from the column with the protein of interest.

Nickel affinity chromatography is one of the more laborious and expensive methods for spider silk purification, especially at larger scales. Combined with the purity issues seen with the MiSp purification on the Ni-NTA column, the yields for the MiSp8 and MiSp8s purifications averaged only ~15 mg rSSp/L aerobic fermentation. The low yields may be the result of inconsistent column binding and elution of the MiSp in solution—an issue that is likely due to the folding of the MiSps upon expression in *E. coli*, which limits the accessibility of the His-tag. Because of these low yields, alternative methods were explored to achieve yields that could sustain the goals of this thesis.

Fractional Ammonium Sulfate Precipitations

As with nickel affinity chromatography, fractional ammonium sulfate precipitations are frequently used in the purification of recombinant spider silk proteins to obtain a protein pellet that can be then resolubilized in a concentrated “dope” for material applications. For each synthetic MiSp construct, precipitations were done at varying saturations with ammonium sulfate to identify a saturation point where all the MiSp was removed from the soluble fraction. At 15% saturation and 4°C, each MiSp construct (MiSp8, MiSp8s, MiSp16, and MiSp16s) could no longer be detected by western blot in the soluble fraction, indicating that both conformations of MiSp seen in the lysate were now in the pellet. The yields for this method were much better than with the chromatography method, at about ~100-200 mg/L for each construct. Table 2-1 shows the dry yields for each protein construct.

Table 2-1. Average dry weight totals for MiSp purification with fractional ammonium sulfate precipitations

MiSp Construct	Average Dry Yield (mg/L)
MiSp8	200
MiSp8s	187
MiSp16	182
MiSp16s	86

Upon precipitation, however, a marked decrease in solubility was seen to the point that it was extremely difficult to resolubilize the pellet for analysis using SDS-

PAGE. It was determined that the pellets were also completely insoluble under high heat and pressure, in harsh organic solvents such as hexafluoroisopropanol (HFIP), and in strong chaotropic salts such as guanidine thiocyanate and sodium dodecyl sulfate (SDS). Because of this incredible level of insolubility, it was difficult to evaluate the purity of the pellets and to verify that the protein was indeed present in the pellet sample due to the fact that SDS-PAGE requires protein denaturation and solubility. Qualitative analysis during the doping process under high heat and pressure indicated that the purity level was high, as there was no observed color change or phase separation in the vial with continued doping, even with the addition of propionic acid and ammonium hydroxide (0.5M) at concentrations up to 5%.

To assess purity and verify the protein was not degraded or altered during the precipitation, a wide variety of methods was attempted to solubilize the pellet and the MiSps. Figure 2-17 shows sequential resuspensions and sonications of MiSp ammonium

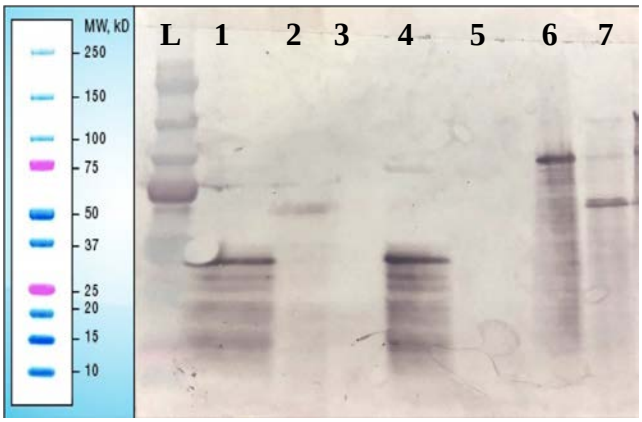

Figure 2-17. Urea and sarkosyl washes to verify MiSp in ammonium sulfate pellets. Lane 1-MiSp8s; Lane 2- MiSp16; Lane 3- MiSp8 insoluble; Lane 4- MiSp8s; Lane 5- MiSp8 insoluble; Lane 6- MiSp16s; Lane 7- MiSp16

sulfate pellets with 8M urea and 2% sarkosyl, which eventually resulted in a very small amount of solubilization and denaturation of the MiSps so that they could be analyzed using SDS-PAGE and western blotting. In most of these cases, the protein concentration was only high enough for western blotting and bands were not visible on the Coomassie-stained polyacrylamide gel. To evaluate the purity of the pellet, enough of the MiSp8 pellet was eventually solubilized using sequential sonications with 8M urea, 2% sarkosyl, and doping under high heat and pressure. Figure 2-18 shows this gel. This was the first

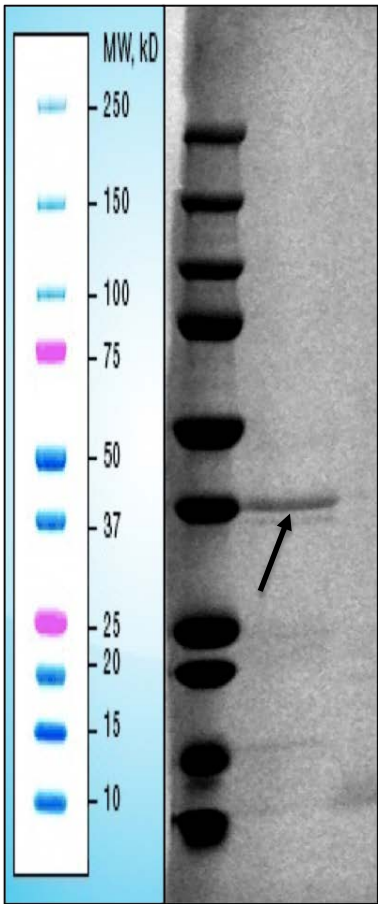

Figure 2-18. SDS-PAGE of partially solubilized MiSp8 (43 kDa) ammonium sulfate pellet

and only instance where enough of the ammonium sulfate protein pellet was solubilized for visualization on a Coomassie-stained polyacrylamide gel. In every other case, the concentration of soluble and denatured protein was only high enough for western blotting visualization (<100 ng).

Although the concentrations of the other constructs in solution did not meet the threshold to see a signal with the Coomassie stain, the MiSp8 pellet seen in Figure 2-18 reinforces what was qualitatively evident. The ammonium sulfate precipitation results in an MiSp protein pellet with a high degree of purity. The insolubility of these pellets, however, was a significant problem because they could not be solubilized in any solvent for analysis, even at concentrations as low 1% w/v. Known issues in the expression and purification of synthetic spider silks are the aggregation and natural insolubility of these proteins both at the expression and purification stages; however, purified synthetic MaSps are still soluble in HFIP and SDS in the worst-case scenarios. Information from the literature has identified ammonium sulfate as one of the most viable methods for protein stability and solubility[50]. Data from the Lewis lab and others report that the level of aggregation and insolubility seen with MiSps at the ammonium sulfate step does not occur in the purification of MaSps and chimeric silks from a bacterial platform[31,35,25,38].

Further investigation with the precipitation of MiSps using dialysis showed a less drastic, but similar level of insolubility seen with the ammonium sulfate precipitations. Analysis of the pellets using FTIR did not show an increase in beta sheet structures; aggregation of the MiSps in the insoluble fraction therefore is the likely cause of the insolubility and it is not a result of structural changes occurring upon precipitation.

Increasing the precipitation time and slowly (0.5 ml/minute) adding solubilized ammonium sulfate to the clarified lysate to avoid the “crash” of the MiSp out of solution failed to improve the solubility of the pellet. Because of this, no material applications could be explored following this purification method.

Ethanol/Isopropyl Alcohol/Acetone Precipitations

Although the fractional ammonium sulfate precipitations succeeded from a purification standpoint, alternate precipitation methods were explored so that the synthetic MiSps could be investigated for material applications. Ethanol, isopropyl alcohol (IPA), and acetone were chosen as potential alternatives because preliminary investigations showed that upon precipitation with these solvents, the MiSps were much more soluble so they could be denatured in SDS and analyzed. The reasons for this solubility difference are relatively unclear as the modes of action of precipitation between these methods aren’t inherently that different. In the alcohol/IPA/acetone precipitations however, the precipitation of MiSps happens instantaneously, which may not allow for the levels of protein binding and aggregation seen with methods like dialysis and ammonium sulfate where the concentration of free water diminishes more slowly. Alcohol precipitations also result in more impurities in the precipitate and are far less specific when it comes to the complete precipitation of both conformations of MiSp in a single step, as was achieved with the ammonium sulfate precipitation. The presence of these impurities may also interrupt the aggregated complex of MiSps that exhibited such incredible levels of insolubility. While obtaining more soluble MiSps was a positive

aspect of these methods, significant investigation was necessary to evaluate which combinations and concentrations of these solvents were needed to effectively precipitate both conformations of MiSp and still arrive at an acceptable purity level for material applications to be successful.

The MiSp8, MiSp8s, and MiSp16 constructs underwent initial precipitations with each of these solvents and ethanol was identified as the best of the three solvents for precipitation of impurities because MiSps were very stable in high concentrations of ethanol. Interestingly, the stability in solution and solubility of each of these constructs decreased with molecular weight. Figure 2-19 shows the complete purification of MiSp8 from the heat/PEI step through to the ethanol precipitation pellet. The MiSp8 (40 kDa) precipitated after the concentration of ethanol reached 30% (v/v), while the MiSp8s (53

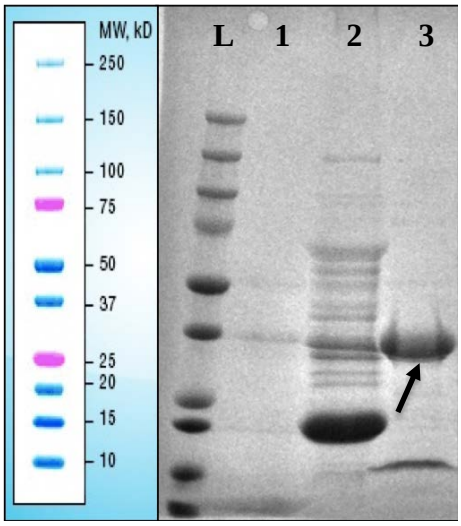

Figure 2-19. MiSp8 purification. Lane 1: MiSp8 clarified lysate post heat and PEI treatments; Lane 2: discarded pellet from PEI and heat treatments; Lane 3: MiSp8 pellet from a 30% ethanol cold precipitation, arrow indicates MiSp8 in final pellet

kDa) required an overnight precipitation or an additional 10% (v/v) precipitation to fall out of solution and much of the MiSp8s remained in solution under these circumstances.

MiSp8 was best-suited for the alcohol precipitations because it exhibited the highest purity and best recovery when compared with the other constructs. The MiSp16 was the most difficult protein to precipitate out of solution using the alcohol/IPA/acetone precipitation method. Figure 2-20 shows MiSp16's stability in sequential precipitations using both ethanol and IPA. MiSp16 was stable in solution at concentrations up to 40% ethanol and 20% IPA, with minimal losses. This was convenient from a purification standpoint because the ethanol and IPA precipitations remove many non-protein impurities that have been known to cause problems downstream. However, these precipitations did not remove the low molecular weight protein impurities that also plagued the chromatography method and resulted in minimal losses of the smaller conformation of MiSp, which do not show with the His-tag antibody but do show with the C-terminus antibody. The lower molecular weight conformation of MiSp that was described in this chapter was shown to be less soluble than the higher molecular weight band; however, it is still believed to be full-size MiSp. To precipitate the rest of both conformations of MiSp16, an additional 20% acetone purification was used following the ethanol and IPA treatments. Figure 2-21 shows the pellets with varying concentrations of acetone.

The purification of MiSp16s using these methods was not viable because it was not stable enough to stay soluble following the initial PEI and heat purification step, which was an imperative step in this process due to the purity problems that presented

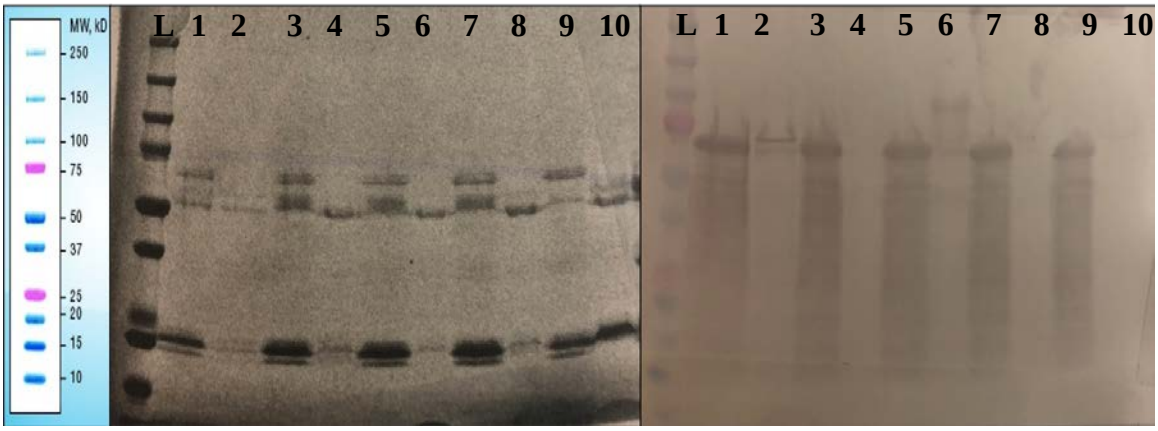

Figure 2-20. MiSp16 ethanol/IPA precipitations. Left: SDS-PAGE; Right: Western blot with His-tag antibody; Lane 1: 10% IPA/25% EtOH super; Lane 2: 10% IPA/25% EtOH pellet; Lane 3: 15% IPA/25% EtOH super, Lane 4: 15% IPA/25% EtOH pellet; Lane 5: 20% IPA/25% EtOH super; Lane 6: 20% IPA/25% EtOH pellet; Lane 7: 30% IPA/25% EtOH super; Lane 8: 30% IPA/25% EtOH pellet; Lane 9: 20% IPA/40% EtOH super; Lane 10: 20% IPA/40% EtOH pellet

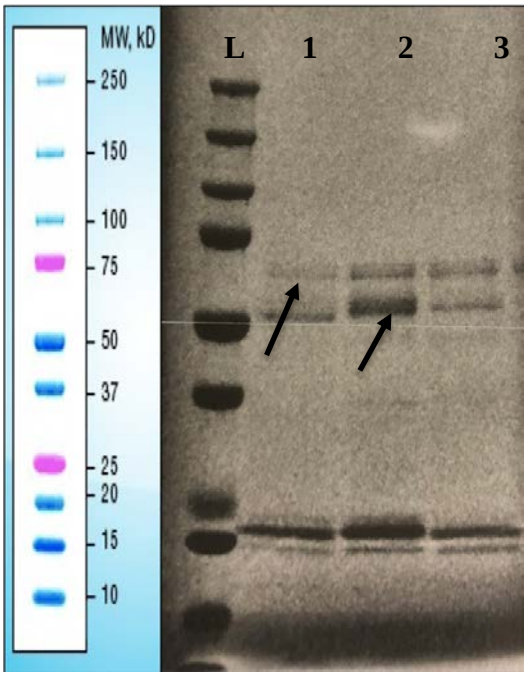

Figure 2-21. MiSp16 ethanol/IPA/acetone pellets; Lane 1: 10% acetone pellet; Lane 2: 15% acetone pellet; Lane 3: 20% acetone pellet

themselves even after most of the *E. coli* proteins and nucleic acids were removed from that initial clean-up step. Table 2-2 shows the yields from these methods. Note that the yields are slightly higher than seen in the ammonium sulfate method; however, purity levels have to be considered in that analysis.

Table 2-2. Average dry weight yields for ethanol/IPA/acetone precipitations

MiSp Construct	Average Dry Yield (mg/L)
MiSp8	250
MiSp8s	228
MiSp16	192
MiSp16s	n/a

These precipitations using alcohols and acetone were effective in that they produced synthetic MiSps that were markedly more soluble in the pellet form. This pellet could then be utilized in concentrated dopes for material applications; however, the purity was not at the level seen with the ammonium sulfate precipitations. These purity issues are unavoidable with this method because it requires high concentrations of alcohol or acetone to precipitate the MiSps, which also precipitates protein and non-protein impurities that remain in the clarified lysate. These impurities have been known to

present problems with protein workability and mechanical properties; however, the sacrifice of purity was necessary in these cases to avoid the solubility issues and move forward with the project goals.

PRELIMINARY BIOPHYSICAL CHARACTERIZATION OF SYNTHETIC MINOR AMPULLATE SILK PROTEINS

The preliminary characterization of aqueous minor ampullate silk proteins and their material applications allows for the first comparison between synthetic analogs of minor ampullate silk proteins and how the properties of MiSp films depend on purification processing, protein sequence, and protein molecular weight.

FTIR spectra including the Amide I (1600-1700 cm^{-1}) and Amide II (1500-1600 cm^{-1}) regions for MiSp8, MiSp8s, MiSp16, and MaSp1 (transgenic goat-derived) are shown in Figure 5-2. Goat-derived MaSp1 (65 kDa) was used as a control/comparison in these studies because it has similar sequence and structural motifs to the MiSps. Data is limited on *Nephila clavipes* MaSp1 produced from a bacterial platform; therefore, the goat-derived protein was the best alternative for a comparison. Figure 2-22 shows that MaSp1, MiSp8, and MiSp16 deviate little in their spectra and exhibit predominantly β-sheet structures. However, the MiSp8s and MiSp16s spectra have a clear left shift. This indicates a transition to primarily α-helical/random coil structures.

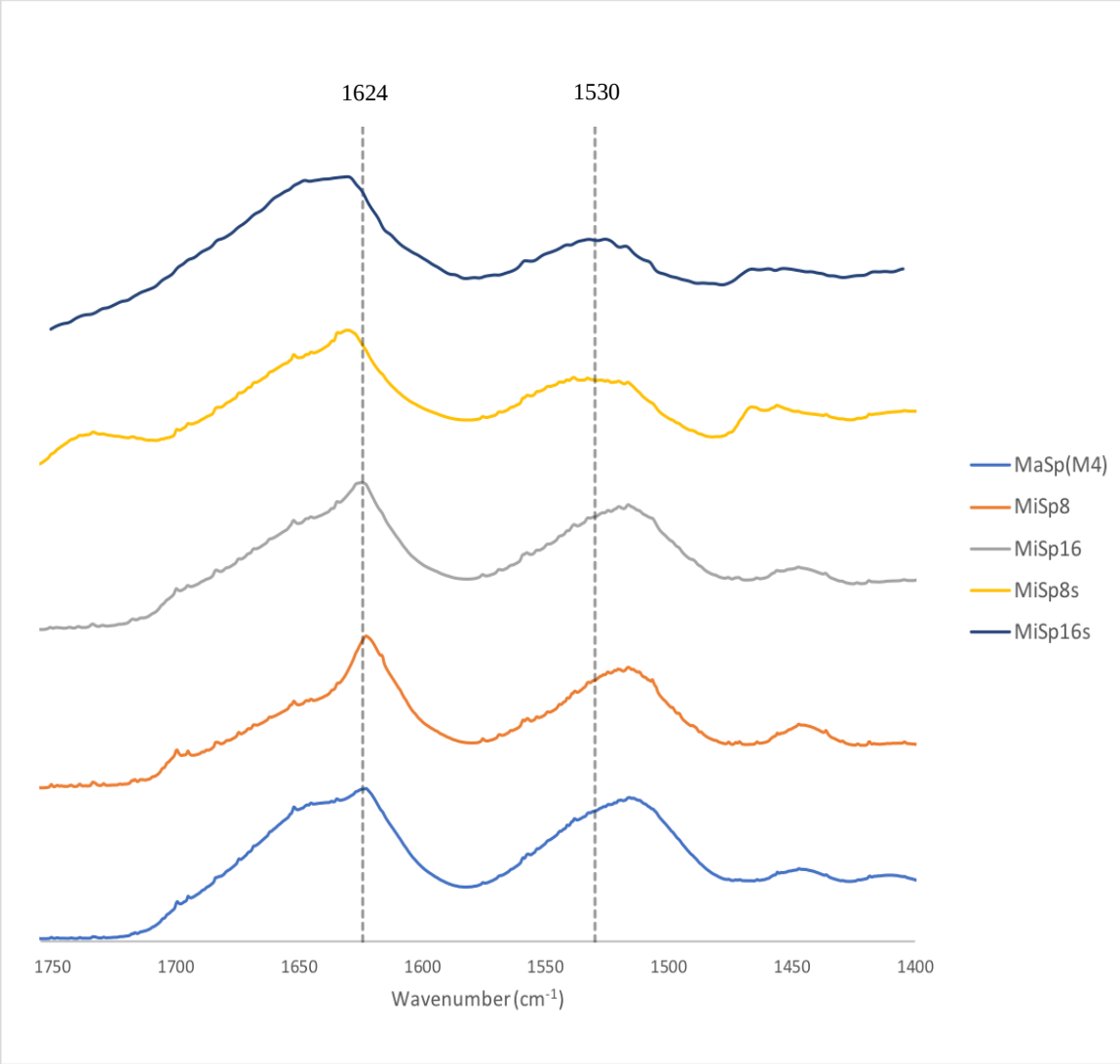

Figure 2-22. FTIR spectra of rSSps; 1624 cm^{-1} and 1530 cm^{-1} peaks mark prominent β-sheet regions in Amide I and Amide II

To further analyze the spectra and this shift, the Amide I region was deconvoluted and Figure 2-23 shows these deconvolution results. The Amide I region was chosen for deconvolution because it is the primary amide vibration band and its vibration depends

A

Construct	β- sheet (%)	α-helix/random coil (%)
MaSp1	60	29
MiSp8	65	25
MiSp8s	23	76
MiSp16	60	36
MiSp16s	19	77

B

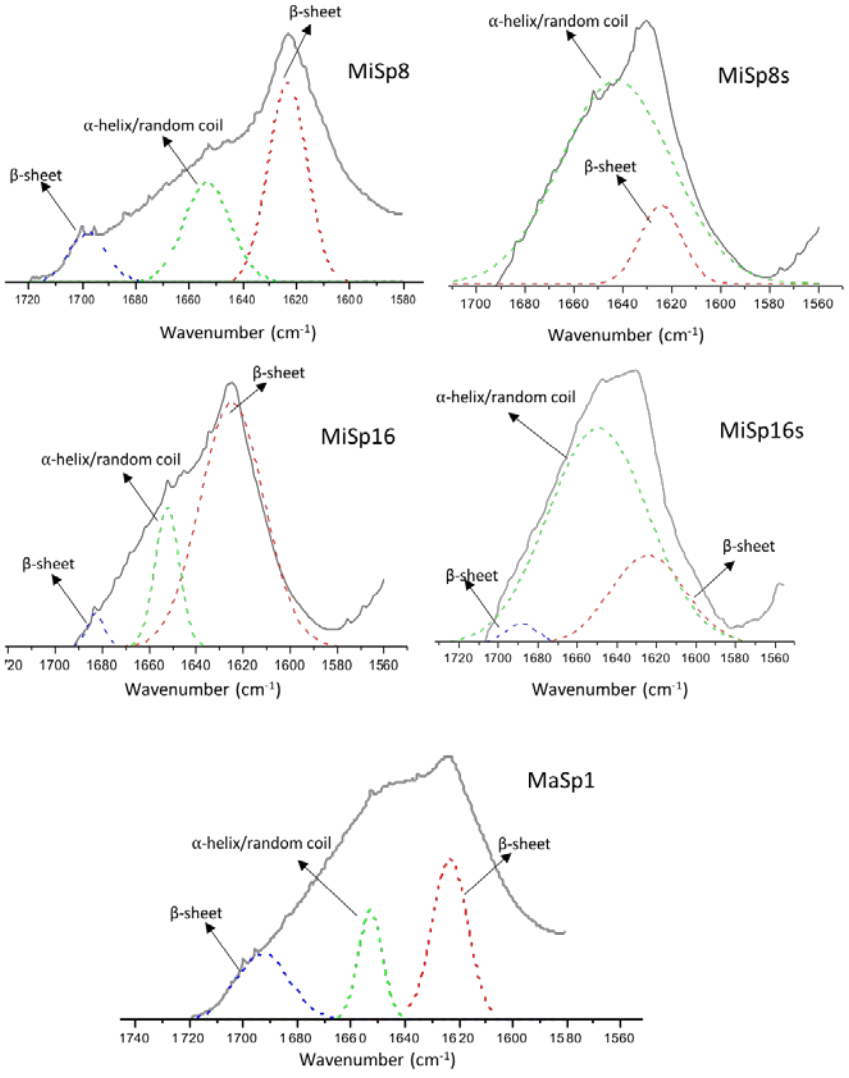

Figure 2-23. FTIR deconvolution results; A: Summary of deconvolution results; B: Deconvolution of Amide I of the IR spectra with peak centers at 1620, 1650, and 1698 cm^{-1}

on the secondary structure of the protein backbone, without hinderance from the nature of the residues' side chains[52]. The three characteristic peaks used for the deconvolution were 1624 cm^{-1}, 1650 cm^{-1}, 1698 cm^{-1}, which have been identified as reliable indicators of protein secondary structure in silk structural analyses using both FTIR and Raman spectroscopy[51,52]. Figure 2-23 (B) indicates the secondary structure assignments to these characteristic peaks[53]. All of the constructs had similar relative percentages of β-sheet and α-helical/random coil structures, with the exception of MiSp8s and MiSp16s. Figure 2-23 (A) shows these integration results. These were very interesting results because no definitive secondary structure has been directly associated with the spacer region; however, it has been hypothesized that the spacer region is composed mainly of α-helical structures, does not participate in β-sheet structures, and may also serve as a matrix to embed the crystalline regions[13]. The FTIR spectra of the MiSp8s and MiSp16s protein in this investigation adds evidence to these hypotheses because there is a definitive and significant shift towards α-helical/random coil structures compared with the MiSp8 and MiSp16 constructs.

From this preliminary characterization, it can be deduced that the addition of the spacer region adds significant α-helical structures to the protein's identity; however, it is unclear if the repetitive regions of MiSp8s and MiSp16s were unable to fold into their correct secondary structure due to the low β-sheet percentages seen from the deconvolution compared with MiSp8 and MiSp16.

To further characterize these spidroins and test their properties in material applications, films from an aqueous dope were created and tested. The protein pellets from the ethanol/IPA/acetone precipitations had to be used to complete this project goal because the insolubility of the ammonium sulfate pellets eliminated them for material applications. These pellets were significantly less pure than the ammonium sulfate pellets, and that may have had an effect on the mechanical properties of the aqueous films. Table 2-3 displays tensile mechanical properties. As expected based on the findings of prior publications, higher molecular weight was correlated with a higher average stress[45]. However, none of the stress values were statistically different from each other. Interestingly, the MiSp8s had just as high of an average stress value as the other rSSp constructs, which were determined to have a much higher β-sheet content from FTIR analysis; therefore, the helices in MiSp8s contribute substantially to tensile strength.

Table 2-3. Tensile data for rSSps aqueous films

rSSp	Average Max Stress (MPa)	Average Max Strain (%)
MiSp8 (40 kDa)	20.54 ± 8.37	2.9 ± 1.1*
MiSp8s (53 kDa)	25.9 ± 3.0	2.7 ± .1*
MiSp16 (67 kDa)	30.67 ± 4.24	3.3 ± 0.2*
MaSp1 (65 kDa)	39.1 ± 8.52	1.8 ± .2

* denotes significance $p < .05$

It has also been hypothesized that the presence of the spacer region may be the cause of the inelasticity of minor ampullate silk because it irreversibly deforms upon stretching[13]. The MiSp aqueous films used in this study therefore could not be stretched before testing. This stretching has been shown to increase tensile strength and strain in recombinant MaSps, possibly as a result of the alignment of the protein chains[36]. Therefore, the low strain value of the MaSp1 in Table 2-3 is somewhat deceptive, as unstretched MaSp1 was used for comparison with the MiSp aqueous films. Using established methods for film post-treatment and stretching, aqueous MaSp1 films can experience an increase in tensile stress by 2.5x and strain by 18x[36]. Due to the small sample sizes available in this study, significant investigation into post-treatments and stretching ratios could not be explored for the MiSp films. However, compared with unstretched MaSp1, the MiSp films all had higher strain values, which were significant. Based on what is known about the sequence of MiSp and its mechanical properties in nature, it is not known whether custom post-treatment methods can be used to increase these strain values. However, it is likely that future work with MiSp film treatments and processing could significantly increase the tensile strength, as is seen with the recombinant MaSps. Possible film treatments include dope additives, alcohol baths, and ratios of film stretching.

CHAPTER 3

CONCLUSIONS

This is the first known investigation involving synthetic recombinant minor ampullate silk proteins. The study of synthetic spider silk proteins has largely been confined to the major ampullate silks, however understanding the expression and purification considerations for other silk proteins may expand the material possibilities for synthetic spider silk industrial use. Using synthetic protein analogs may also shed light on how protein sequence motifs translate to secondary structure and mechanical properties. In this investigation, six synthetic MiSp constructs with varying iterations of characteristic protein motifs were created for expression in *E. coli*. These proteins were then successfully expressed at full-length in high-density cell culture at molecular weights ranging from 40 to 95 kDa.

Protein expression and extraction were more efficient with the lower molecular weight constructs and two conformations of MiSp could be detected in the soluble fraction depending on the antibody used, which was likely due to folding inconsistencies in high-density culture.

In order to study these protein analogs, four general purification methods were used: an initial lysate purification using PEI and heat treatments, nickel affinity chromatography, fractional ammonium sulfate precipitations, and ethanol/IPA/acetone precipitations. MiSp's folding patterns introduced challenges in purification, as the expressed proteins exhibited a limited accessibility to the 6x-His tag. As a result, nickel

affinity chromatography produced yields that were too low to support the goals of this project. Furthermore, low molecular weight impurities interacting with the MiSps in solution created purity problems throughout the purification processes.

Fractional ammonium sulfate precipitations produced an MiSp pellet that exhibited the highest purity level of all the purification methods, however protein aggregation resulted in a totally insoluble pellet that prevented use of the protein for all material applications. In order to precipitate the protein and solubilize the pellet in high concentration dopes for material applications, ethanol/IPA/acetone precipitations had to be used to preserve the solubility of the MiSps. The solubility difference of these solvent-extracted pellets could be due to the lower purity of the final pellet or it could be due to the rapid phase change the proteins undergo when exposed to the solvents. This method, although effective from a solubilization standpoint, was not ideal due to the fact that the solvent precipitations were not selective, which resulted in some MiSp left in solution and/or more impurities in the final pellet that may have interfered with mechanical properties of the materials.

A preliminary structural investigation and comparison was done using FTIR-ATR spectroscopy, which showed very little difference in secondary structure between MiSp8, MiSp16, and goat-derived MaSp1. These proteins all exhibited predominantly β-sheet structures, with ~20% α-helices contributing to the spectra. However, the MiSp8s and MiSp16s proteins including the spacer region(s) showed a definitive shift in the spectra and its secondary structure exhibited significant α-helices. This is the first defined characterization of the spacer region.

For information on how these secondary structures transfer to mechanical properties, aqueous MiSp films were created for tensile testing and compared to each other and unstretched, goat-derived MaSp1. The tensile stress of these films were not significantly different from each other or MaSp1. However, the strain values were all significantly higher than the strain of unstretched MaSp1. This result is interesting, but it must be also considered that the more characterized MaSp1 films experience increases in tensile properties when they are processed in a series of post-treatments that include a variation of stretch baths[36]. The sample sizes in this study were inadequate for an in-depth investigation of film processing as has been done previously for the MaSps, however it is likely that with processing MiSp films would also see an increase in tensile strength. Since these proteins are naturally inelastic, it is unlikely that MiSp films would see the elongation increases seen for MaSps—allowing them to fill different niches in product applications, such as ballistics.

Future Work

While this investigation yielded the first information on synthetic minor ampullate silk proteins, more work needs to be done to improve extraction efficiency, yields, protein purity, and material mechanical properties. While four MiSps were produced in the soluble fraction of high-density culture, there are folding inconsistencies of the full-size protein that are not yet fully understood. This causes variation in the behavior of the MiSps in solution, which complicates purifications that take advantage of protein charge, solubility, and folding and ultimately affects the yield and purity of the processes.

MiSp16s showed a high instability in solution after extraction, which resulted in an inability to purify it for use in aqueous films. Future work will require an investigation of the stability requirements for these proteins after lysis, as well as what can be done to prevent the aggregation and insolubility issues presented with the fractional ammonium sulfate precipitations. This may also require alternative purification methods such as ion-exchange chromatography, size exclusion chromatography, or reverse-phase chromatography. Investigating these other chromatography methods allows for an opportunity to significantly clean up the clarified lysate before precipitation, as the precipitation step was identified in this investigation as a critical step that always introduced either purity and/or insolubility issues.

Using these synthetically produced spidroins, more work can be done to study the structural implications of the arrangement and prevalence of MiSp's amino acid motifs. Investigating the processing and post-treatment of aqueous films may contribute to this, as it is still unclear if mechanical properties of these films may be enhanced, as is seen with MaSp aqueous films.

Engineering Significance and Innovation

Significance

The functional diversity of spider silks creates unique opportunities for robust, biocompatible, and ecofriendly alternatives for a variety of polymer applications[35,37,55–59].

Because of the unique properties of spider silks, significant research efforts have focused on the native protein sequences and folding structures. Among this research, most investigations focus on the major ampullate silks. This is due to factors that include the impressive combination of strength and elasticity seen in the dragline silk, the large size of the major ampullate gland, and the fact that the native dragline silk fibers can be more easily isolated from the spider for fiber and protein characterization[40]. While the sequences of the other six silks have been determined, significantly less research has been done on their folding structures and mechanical properties. For most of the silks, little to no research has been published on their synthetic production making them difficult to study and further characterize for potential material applications.

Minor ampullate silk is one of these lesser studied silks that possesses properties that could be well suited for many applications. Native minor ampullate silk is difficult to harvest because the natural fibers are significantly smaller in diameter than other spider silk fibers and no minor ampullate silk is easily identifiable in the orb web because it is used as a template for the capture spiral and most species later replace it with flagelliform silk[3]. Therefore, studying minor ampullate silk proteins synthetically makes it possible to study the protein sequences and their subsequent folding structures and how the production of these proteins compares with the production of other spider silk proteins, namely the major ampullate silk proteins[11].

Because of the limited published investigations into synthetic analogs of most of the spider silk proteins, engineering the production and purification of these proteins is novel and requires extensive planning, testing, and revision as was evident in this thesis.

The folding mechanisms of synthetic spider silk proteins are still not entirely understood, therefore we do see folding inconsistencies and insolubility issues in the synthetic production and purification of spider silk proteins. Interestingly, these issues present differently and are more drastic for certain silks compared with others. This may be one of the reasons that there are so few published investigations of synthetic spider silk proteins that are not the MaSps. Investigations like these are needed to understand how sequence motifs, molecular weight, and secondary structure of each of the proteins may be considered to engineer expression and purification methods that are specifically designed for each spider silk protein. Many of the problems associated with expression and purification are similar and solutions may be deduced from the further study of many different spider silk proteins that have slight modifications to their sequence and folding structures.

These investigations also have the potential to expand the opportunities for these proteins and their potential use in industry. The spider silk gene family is diverse and the mechanical properties that each silk possesses could result in material applications for a wide variety of industries that include textiles, medical, ballistics, outdoor etc.

Innovation

This thesis, combined with other investigations, has significantly broadened the knowledge of how protein sequence translates to secondary structure and mechanical properties. This may allow for the construction of new chimeric silks that could be tuned for the selection of certain properties. Each of the seven spider silks possesses properties

that could diversify the portfolio for potential material applications and products. For example, major ampullate silk is bulletproof by nature and can absorb and dissipate high amounts of energy, but its elasticity disqualifies it for real-life ballistic applications. Minor ampullate silk, on the other hand, achieves a high tensile strength while maintaining an elongation that is identical to Kevlar[5]. Ballistic recombinant spider silk materials would also be environmentally friendly to produce and more lightweight than their competitors[5].

Another interesting property of minor ampullate silk is that it does not supercontract in water, which is a phenomenon seen in the major ampullate silk that is not entirely understood. Previous studies have shown that major ampullate silk's tensile properties are affected significantly more than minor ampullate silk properties when tested in water due to the plasticizing effect of supercontraction. This might be due to the fact that MiSps are more crystalline than MaSps, causing a lower loss of orientation when exposed to water[57]. It has been hypothesized that the proline content of the silk protein may also be a contributing factor. This must be accounted for when developing materials that have high exposure to moisture and aqueous environments, such as medical materials or implants. Tuning the crystallinity of a chimeric silk or combining a variation of silk proteins could result in a polymer material that possesses a variation of the sought-after properties of silks.

Based on the findings of this thesis, the highly conserved spacer region contributes helical structures to the final protein structure. These helices have been known to contribute to tensile strength and, based on expression results in this thesis, may

allow for the construction of higher molecular weight silk proteins due to its nonrepetitive nature.

Potential Drawbacks and Cost Analysis

The synthetic production of these proteins in transgenic systems remains a limiting factor in the design of prototype materials and their subsequent use in industry. Referring Figure 1-5[33], rSSp production in bacterial systems needs to reach yields of 10 g/l of fermentation to compete with fiber materials such as Kevlar, which sell for about $20-$30/kg[60]. However, non-fiber applications such as films, coatings, adhesives, lyogels, hydrogels, and foams typically require much less protein than fibers do, which increases the cost-effectiveness of rSSps for these materials[35]. MiSp yields from *E. coli* have been shown in this thesis to be low (~100-300 mg/L) compared with other reported yields of rSSps like MaSp2 (800 mg-2 g/L)[26,29,49,54]. Similar low yields in bacterial systems, such as those seen for the MiSps, have been reported for high molecular weight constructs of MaSp1. This is likely due to the fact that MaSp2 has been the subject of more studies as far as synthetic production is concerned. More research needs to be done to optimize the parameters of cell growth, protein expression, protein extraction, and purification methods for these protein constructs to maximize yields and purity to increase the likelihood that synthetic rSSps may be used in industrial products.

REFERENCES

1. *Spider Physiology and Behaviour: Physiology*. (Academic Press, 2011).
2. Clarke, T. H., Garb, J. E., Hayashi, C. Y., Arensburger, P. & Ayoub, N. A. Spider Transcriptomes Identify Ancient Large-Scale Gene Duplication Event Potentially Important in Silk Gland Evolution. *Genome Biol. Evol.* **7,** 1856–1870 (2015).
3. Lewis, R. V. Spider Silk: Ancient Ideas for New Biomaterials. *Chem. Rev.* **106,** 3762–3774 (2006).
4. Römer, L. & Scheibel, T. The elaborate structure of spider silk. *Prion* **2,** 154–161 (2008).
5. Altman, G. H. *et al.* Silk-based biomaterials. *Biomaterials* **24,** 401–416 (2003).
6. Hayashi, C. Y., Shipley, N. H. & Lewis, R. V. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *Int. J. Biol. Macromol.* **24,** 271–275 (1999).
7. Hayashi, C. Y. & Lewis, R. V. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *J. Mol. Biol.* **275,** 773–784 (1998).
8. Ittah, S., Cohen, S., Garty, S., Cohn, D. & Gat, U. An Essential Role for the C-Terminal Domain of A Dragline Spider Silk Protein in Directing Fiber Formation. *Biomacromolecules* **7,** 1790–1795 (2006).
9. Holland, G. P., Jenkins, J. E., Creager, M. S., Lewis, R. V. & Yarger, J. L. Solid-state NMR investigation of major and minor ampullate spider silk in the native and hydrated states. *Biomacromolecules* **9,** 651–657 (2008).
10. Tatham, A. S. & Shewry, P. R. Elastomeric proteins: biological roles, structures and mechanisms. *Trends Biochem. Sci.* **25,** 567–571 (2000).
11. Dong, Z., Lewis, R. V. & Middaugh, C. R. Molecular Mechanism of Spider Silk Elasticity. *Arch. Biochem. Biophys.* **284,** 53–57 (1991).
12. Tokareva, O., Michalczechen-Lacerda, V. A., Rech, E. L. & Kaplan, D. L. Recombinant DNA production of spider silk proteins. *Microb. Biotechnol.* **6,** 651–663 (2013).
13. Colgin, M. A. & Lewis, R. V. Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like ‘spacer regions’. *Protein Sci.* **7,** 667–672 (1998).
14. Chen, G. *et al.* Full-Length Minor Ampullate Spidroin Gene Sequence. *PLoS ONE* **7,** e52293 (2012).
15. Malay, A. D. *et al.* Relationships between physical properties and sequence in silkworm silks. *Sci. Rep.* **6,** (2016).
16. Guinea, G. V. *et al.* Minor ampullate silks from Nephila and Argiope spiders: tensile properties and microstructural characterization. *Biomacromolecules* **13,** 2087–2098 (2012).
17. Hinman, M. B., Jones, J. A. & Lewis, R. V. Synthetic spider silk: a modular fiber. *Trends Biotechnol.* **18,** 374–379 (2000).

18. Shi, J., Lua, S., Du, N., Liu, X. & Song, J. Identification, recombinant production and structural characterization of four silk proteins from the Asiatic honeybee Apis cerana. *Biomaterials* **29,** 2820–2828 (2008).
19. Paquet-Mercier, F., Lefèvre, T., Auger, M. & Pézolet, M. Evidence by infrared spectroscopy of the presence of two types of β-sheets in major ampullate spider silk and silkworm silk. *Soft Matter* **9,** 208–215 (2012).
20. Papadopoulos, P., Sölter, J. & Kremer, F. Structure-property relationships in major ampullate spider silk as deduced from polarized FTIR spectroscopy. *Eur. Phys. J. E* **24,** 193–199 (2007).
21. Chung, H., Kim, T. Y. & Lee, S. Y. Recent advances in production of recombinant spider silk proteins. *Curr. Opin. Biotechnol.* **23,** 957–964 (2012).
22. Brigham, D. L. Methods, compositions, and systems for production of recombinant spider silk polypeptides. (2009).
23. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. *Front. Microbiol.* **5,** (2014).
24. Rising, A., Widhe, M., Johansson, J. & Hedhammar, M. Spider silk proteins: recent advances in recombinant production, structure–function relationships and biomedical applications. *Cell. Mol. Life Sci.* **68,** 169–184 (2011).
25. Arcidiacono, S., Mello, C., Kaplan, D., Cheley, S. & Bayley, H. Purification and characterization of recombinant spider silk expressed in Escherichia coli. *Appl. Microbiol. Biotechnol.* **49,** 31–38 (1998).
26. Fahnestock, S. R. & Irwin, S. L. Synthetic spider dragline silk proteins and their production in Escherichia coli. *Appl. Microbiol. Biotechnol.* **47,** 23–32 (1997).
27. Jana, S. & Deb, J. K. Strategies for efficient production of heterologous proteins in Escherichia coli. *Appl. Microbiol. Biotechnol.* **67,** 289–298 (2005).
28. Baneyx, F. Recombinant protein expression in Escherichia coli. *Curr. Opin. Biotechnol.* **10,** 411–421 (1999).
29. Yang, Y.-X., Qian, Z.-G., Zhong, J.-J. & Xia, X.-X. Hyper-production of large proteins of spider dragline silk MaSp2 by Escherichia coli via synthetic biology approach. *Process Biochem.* **51,** 484–490 (2016).
30. Brooks, A. E. *et al.* Distinct contributions of model MaSp1 and MaSp2 like peptides to the mechanical properties of synthetic major ampullate silk fibers as revealed in silico. *Nanotechnol. Sci. Appl.* **1,** 9–16 (2008).
31. Prince, J. T., McGrath, K. P., DiGirolamo, C. M. & Kaplan, D. L. Construction, Cloning, and Expression of Synthetic Genes Encoding Spider Dragline Silk. *Biochemistry (Mosc.)* **34,** 10879–10885 (1995).
32. Heidebrecht, A. & Scheibel, T. Recombinant Production of Spider Silk Proteins. in *Advances in Applied Microbiology, Vol 82* (eds. Sariaslani, S. & Gadd, G. M.) **82,** 115–153 (Elsevier Academic Press Inc, 2013).
33. Edlund, A. Synthetic Spider Silk Sustainability Verification by Techno-Economic and Life Cycle Analysis. *Grad. Theses Diss.* (2016).
34. Omenetto, F. G. & Kaplan, D. L. New Opportunities for an Ancient Material. *Science* **329,** 528–531 (2010).
35. Jones, J. A. *et al.* More Than Just Fibers: An Aqueous Method for the Production

of Innovative Recombinant Spider Silk Protein Materials. *Biomacromolecules* (2015). doi:10.1021/acs.biomac.5b00226

36. Tucker, C. L. *et al.* Mechanical and Physical Properties of Recombinant Spider Silk Films Using Organic and Aqueous Solvents. *Biomacromolecules* **15,** 3158–3170 (2014).
37. Harris, T. I. *et al.* Sticky Situation: An Investigation of Robust Aqueous-Based Recombinant Spider Silk Protein Coatings and Adhesives. *Biomacromolecules* **17,** 3761–3772 (2016).
38. Binder, S. Scaffolds for retinal pigment epithelium (RPE) replacement therapy. *Br. J. Ophthalmol.* **95,** 441–442 (2011).
39. Bini, E. *et al.* RGD-Functionalized Bioengineered Spider Dragline Silk Biomaterial. *Biomacromolecules* **7,** 3139–3145 (2006).
40. Fredriksson, C. *et al.* Tissue Response to Subcutaneously Implanted Recombinant Spider Silk: An in Vivo Study. *Materials* **2,** 1908–1922 (2009).
41. Elia, R. *et al.* Silk electrogel coatings for titanium dental implants. *J. Biomater. Appl.* **29,** 1247–1255 (2015).
42. Gellynck, K. *et al.* Silkworm and spider silk scaffolds for chondrocyte support. *J. Mater. Sci. Mater. Med.* **19,** 3399–3409 (2008).
43. Meechaisue, C. *et al.* Preparation of electrospun silk fibroin fiber mats as bone scaffolds: A preliminary study. *Biomed. Mater. Bristol Engl.* **2,** 181–8 (2007).
44. Van Vlierberghe, S., Dubruel, P. & Schacht, E. Biopolymer-Based Hydrogels As Scaffolds for Tissue Engineering Applications: A Review. *Biomacromolecules* **12,** 1387–1408 (2011).
45. Xia, X.-X. *et al.* Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber. *Proc. Natl. Acad. Sci.* **107,** 14059–14063 (2010).
46. Huemmerich, D., Slotta, U. & Scheibel, T. Processing and modification of films made from recombinant spider silk proteins. *Appl. Phys. A* **82,** 219–222 (2006).
47. Hayashi, K. & Kojima, C. pCold-GST vector: a novel cold-shock vector containing GST tag for soluble protein production. *Protein Expr. Purif.* **62,** 120–127 (2008).
48. Lewis, R. V., Hinman, M., Kothakota, S. & Fournier, M. J. Expression and Purification of a Spider Silk Protein: A New Strategy for Producing Repetitive Proteins. *Protein Expr. Purif.* **7,** 400–406 (1996).
49. Brooks, A. E. *et al.* Properties of Synthetic Spider Silk Fibers Based on Argiope aurantia MaSp2. *Biomacromolecules* **9,** 1506–1510 (2008).
50. Kramer, R. M., Shende, V. R., Motl, N., Pace, C. N. & Scholtz, J. M. Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility. *Biophys. J.* **102,** 1907–1915 (2012).
51. Guo, C. *et al.* Structural Comparison of Various Silkworm Silks: An Insight into the Structure–Property Relationship. *Biomacromolecules* **19,** 906–917 (2018).
52. Jenkins, J. E., Creager, M. S., Lewis, R. V., Holland, G. P. & Yarger, J. L. Quantitative Correlation between the Protein Primary Sequences and Secondary Structures in Spider Dragline Silks. *Biomacromolecules* **11,** 192–200 (2010).

53. Lefèvre, T., Rousseau, M.-E. & Pézolet, M. Protein Secondary Structure and Orientation in Silk as Revealed by Raman Spectromicroscopy. *Biophys. J.* **92,** 2885–2895 (2007).
54. Xia, X.-X. *et al.* Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 14059–14063 (2010).
55. Bai, S. *et al.* Silk scaffolds with tunable mechanical capability for cell differentiation. *Acta Biomater.* **20,** 22–31 (2015).
56. Calejo Maria Teresa, Ilmarinen Tanja, Jongprasitkul Hatai, Skottman Heli & Kellomäki Minna. Honeycomb porous films as permeable scaffold materials for human embryonic stem cell-derived retinal pigment epithelium. *J. Biomed. Mater. Res. A* **104,** 1646–1656 (2016).
57. Bauer, F., Wohlrab, S. & Scheibel, T. Controllable cell adhesion, growth and orientation on layered silk protein films. *Biomater. Sci.* **1,** 1244 (2013).
58. Ding, D. *et al.* From Soft Self-Healing Gels to Stiff Films in Suckerin-Based Materials Through Modulation of Crosslink Density and β-Sheet Content. *Adv. Mater.* **27,** 3953–3961 (2015).
59. Meinel, L. *et al.* The inflammatory responses to silk films in vitro and in vivo. *Biomaterials* **26,** 147–155 (2005).
60. Price Of Kevlar Fibers Per Kg-Price Of Kevlar Fibers Per Kg Manufacturers, Suppliers and Exporters on Alibaba.comAramid Fabric. Available at: https://www.alibaba.com/trade/search?fsb=y&IndexArea=product_en&CatId=&SearchText=price+of+kevlar+fibers+per+kg&isGalleryList=G. (Accessed: 4th July 2018)
61. Qing, G. *et al.* Cold-shock induced high-yield protein production in *Escherichia coli*. *Nat. Biotechnol.* **22,** 877–882 (2004).
62. Goldenberg, D. *et al.* Role of Escherichia coli cspA promoter sequences and adaptation of translational apparatus in the cold shock response. *Mol. Gen. Genet. MGG* **256,** 282–290 (1997).
63. Li, J. *et al.* Expression of soluble native protein in Escherichia coli using a cold-shock SUMO tag-fused expression vector. *Biotechnol. Rep.* **19,** e00261 (2018).

APPENDICES

APPENDIX A

INVESTIGATION OF A COLD SHOCK VECTOR SYSTEM

Background

In 2004, Takara Bio developed a series of cold-shock expression vectors (pColdI, pColdII, pColdIII, and pCold IV)[61]. These systems regulate gene expression using the cspA promoter, which exhibits up-regulation at growth temperatures of 15°C[62]. This, coupled with an upstream lac operator sequence prevents leaky expression at 37°C and allows for gene expression and protein translation to occur at low temperatures after the addition of IPTG. The vector also includes a translation-enhancing element (TEE) and a 6x-His tag. This process allows expression of the target protein at high yields, while suppressing cell growth and the expression of other cellular proteins. A vector map of pColdII is shown in Figure A-1.

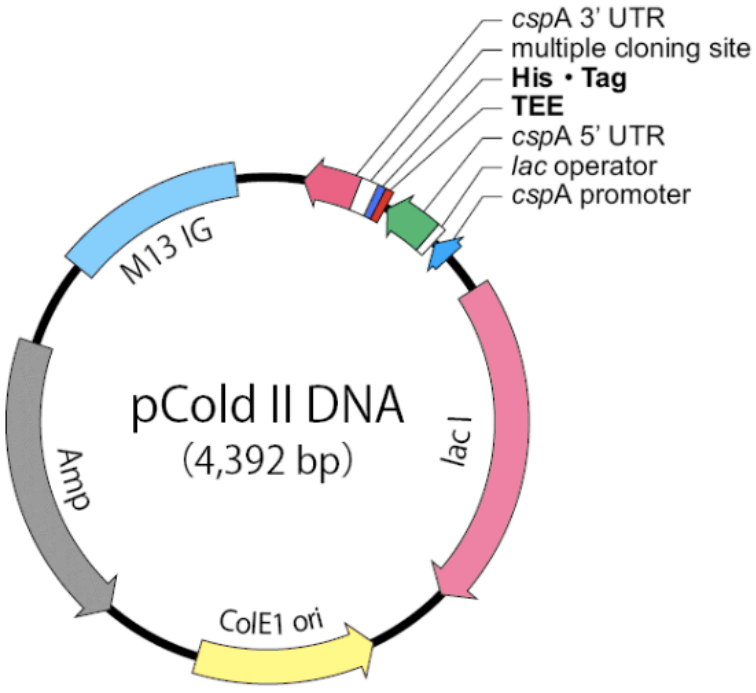

Figure A-1. Vector map of pColdII

These vectors are currently being explored as an alternate choice for recombinant spider silk protein production as well as for the production of other eukaryotic proteins[63]. These large and repetitive proteins exhibit both expression and solubility limitations upon expression in *E. coli*, which may make them prime candidates for a cold-shock expression system.

Materials and Methods

MiSp-pColdII Constructs

The linearized synthetic MiSp constructs MiSp8 and MiSp16 were ligated into pColdII using the restriction sites BamHI and NdeI in the multiple cloning site of the expression vector. Figure A-2 shows these final digested constructs, which were verified by sequencing (ACTG, Inc).

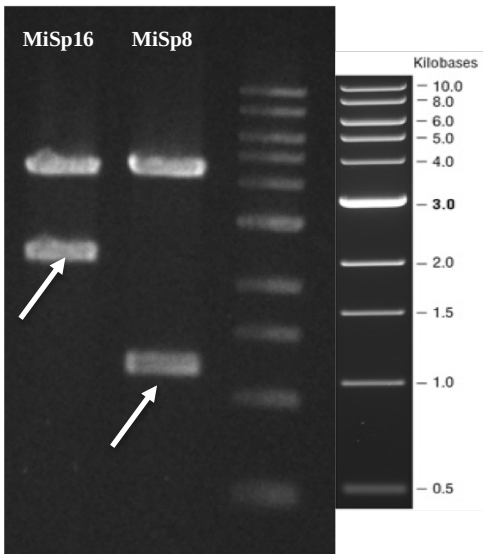

Figure A-2. MiSp-pColdII digested constructs

MiSp-pColdII Preliminary Flask Study

Once the synthetic MiSp-pColdII constructs had been verified, a preliminary flask study was performed to evaluate expression levels of MiSp8. In order to effectively compare the expression levels between the pColdII system and the pET19kt system, two side-by-side 1L flask studies were done using MiSp8-pColdII and MiSp8-pET19kt. Overnight cultures of a single colony of each construct were used to inoculate both flasks containing Terrific Broth media as was described in Chapter 2. Both flasks were incubated at 37°C and 200 RPM until OD_{600} values reached 0.6 to 0.8. At this point, the pColdII flask was removed from the incubator and chilled on ice until the temperature of culture dropped to 15°C. Both flasks were then inoculated with a final concentration of 1mM IPTG. After induction with IPTG, the pColdII flask was incubated at 15°C and the pET19kt flask was incubated at 25°C according to the protocol in Chapter 2. Both flasks were incubated for 24 hours after induction.

Unfortunately, the pColdII vector has not yet been adapted to include the C-terminus region that the pET19kt vector includes so western blot analysis of the hourly samples had to be done using the His-tag antibody to compare the expression to the pET19kt expression system. Expression tests and SDS-PAGE analysis was done according to the protocol in Chapter 2.

Preliminary Results and Discussion

Recombinant spider silk proteins have traditionally been difficult to express in bacterial systems like *E. coli* due to significant truncation of the full molecular weight

product and improper folding that results in insoluble inclusion bodies that are difficult to extract and purify.

Because the pColdII vector lacks the C-terminus coding region, analysis had to be done using the His-tag antibody instead of the C-terminus antibody. This was not ideal, as the prior flask studies carried out in this thesis showed that the folding of MiSps limit the accessibility of the His-tag and it was difficult to get a strong signal at the low protein expression levels available at the flask scale, whereas the C-terminus antibody gives a much stronger signal at these low expression levels. The expression checks showing the soluble fractions of both of these flask studies are shown in Figure A-3. It's seen in this western blot analysis that expression of MiSp8 was evident in both the pColdII and pET19kt flask, however this preliminary flask study shows that the pET19kt flask may have resulted in higher expression levels than the pColdII flask. The MiSp8 expressed in

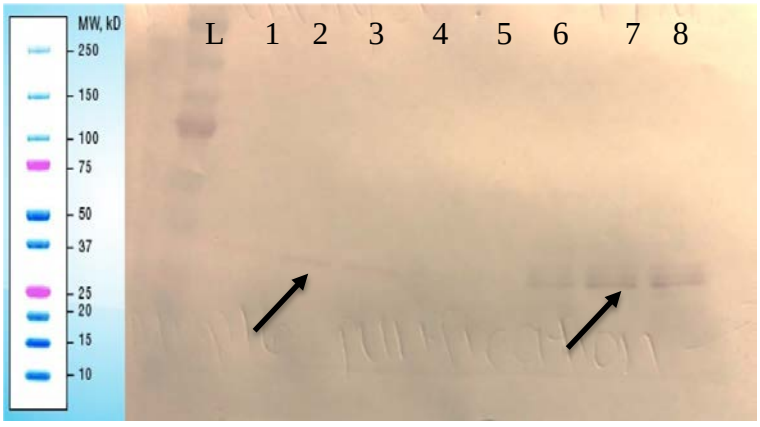

Figure A-3. His-tag western blot of pColdII and pET19kt expression of MiSp8. Lanes 1-4: soluble fractions of pColdII flask at hours 4, 16, 21, 24 after induction; Lanes 5-8: soluble fractions of pET19kt flask at hours 4, 16, 21, and 24 after induction

the pColdII system does appear to have lower levels of truncation. This lower level of truncation is to be expected based on the low induction and incubation temperature.

Conclusions

Two synthetic MiSp constructs (MiSp8 and MiSp16) were ligated into a pColdII expression vector for expression analysis compared to a pET19kt expression system. A preliminary flask study exhibiting a side-by-step 24 hour induction of MiSp8 showed that the pColdII vector did not show significantly higher expression in the soluble fraction of recombinant protein when compared to pET19kt, however it does appear to result in lower levels of truncation. This may be significant at larger scales where truncation becomes an issue at higher OD_{600} values. Because pColdII carries an ampicillin resistance gene, benchtop bioreactor studies were not done to evaluate these truncation levels at higher OD_{600} values. Ampicillin is not compatible with high-density, as beta-lactamase causes degradation problems in high-density culture.

Future work needs to be done to evaluate how well this expression system may work with the production and purification of synthetic MiSps. In order to work with this vector in high-density culture, a kanamycin adapter needs to be cloned in and more analysis needs to be done on the expression and truncation levels of each recombinant spider silk protein. Data from this thesis and from literature indicate that lower induction temperatures help with the truncation and overall yield of these proteins, therefore a cold-shock vector system like pColdII may be a suitable alternative to pET series vectors.

APPENDIX B

SYNTHETIC RECOMBINANT MINOR AMPULLATE SILK PROTEIN SEQUENCES

C-termini highlighted in blue and spacer regions highlighted in yellow

MiSp8

M T G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A
G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G
A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G
A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G
A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G
A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A
G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A
G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A
G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A
G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G
G Y G G Q G G Y G A G A G A G A A A A A G A G A S G G D P G S A S R L A S P
D S G A R V A S A V S N L V S S G P T S S A A L S S V I S N A V S Q I G A S N P
G L S G C D V L I Q A L L E I V S A C V T I L S S S S I G Q V N Y G A A S Q F A Q
V V G Q S V L S A F A A **Stop**

MiSp8s

M T G G G S S A G N A F A Q S L S S N L L S S G D F V Q Met I S S T T S T D E A
V S V A T S V A Q N V G S Q L G L D A N A Met N N L L G A V S G Y V S T L G N
A I S D A S A Y A N A L S S A I G N V L A N S G S I S E S T A S S A A S S A A S S
V T T T L T S Y G P A V F Y A P S A S S G G S G G A G G Y G R G A G A G A G A
A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G
G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A
A A A A G A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G
Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A
A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G
Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A
A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q
G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A
G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y
G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A
A A G A G A S G G D P G S A S R L A S P D S G A R V A S A V S N L V S S G P T
S S A A L S S V I S N A V S Q I G A S N P G L S G C D V L I Q A L L E I V S A C V
T I L S S S S I G Q V N Y G A A S Q F A Q V V G Q S V L S A F A A **Stop**

MiSp16

M T G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A
G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G
A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G
A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G
A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G
A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G A
G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A
G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G A
G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A
G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G
G Y G G Q G G Y G A G A G A G A A A A A G A G A S G G A G G Y G R G A G A
G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G A
G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A
G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G
G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G
A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G
G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G
A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G
Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A
G A A A G A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G
A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A
G A A A A A G A G A S G G D P G S A S R L A S P D S G A R V A S A V S N L V S
S G P T S S A A L S S V I S N A V S Q I G A S N P G L S G C D V L I Q A L L E I V
S A C V T I L S S S S I G Q V N Y G A A S Q F A Q V V G Q S V L S A F A A **Stop**

MiSp16s

M T G G G S S A G N A F A Q S L S S N L L S S G D F V Q M I S S T T S T D E A V
S V A T S V A Q N V G S Q L G L D A N A M N N L L G A V S G Y V S T L G N A I
S D A S A Y A N A L S S A I G N V L A N S G S I S E S T A S S A A S S A A S S V T
T T L T S Y G P A V F Y A P S A S S G G S G G A G G Y G R G A G A G A G A A A
G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y
G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A
A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G
G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G A G A G A A A G
A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G
R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A
A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G
Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A
G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R
G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A
G A G A S G G G S S A G N A F A Q S L S S N L L S S G D F V Q M I S S T T S T D

E A V S V A T S V A Q N V G S Q L G L D A N A M N N L L G A V S G Y V S T L
G N A I S D A S A Y A N A L S S A I G N V L A N S G S I S E S T A S S A A S S A
A S S V T T T L T S Y G P A V F Y A P S A S S G G S G G A G G Y G R G A G A G
A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G
G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G
A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G
Y G G Q G G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G A G A
G A A A G A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G
A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A
G A A A A A G A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y
G G Q G G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G A G A G
A A A G A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A
G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G
A A A A A G A G A G A S G G D P G S A S R L A S P D S G A R V A S A V S N L V S S
G P T S S A A L S S V I S N A V S Q I G A S N P G L S G C D V L I Q A L L E I V S
A C V T I L S S S S I G Q V N Y G A A S Q F A Q V V G Q S V L S A F A A **Stop**

MiSp24

M T G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A
G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G
A G G Y G G Q G G Y G A G A G A G A A A A A A G A G A G A G G A G G Y G R G A G
A G A G A A A G A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A A G A G
A G G A G G Y G R G A G A G A G A A A G A G A G A G A G G Y G G Q G G Y G A G
A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G A
G G Y G G Q G G Y G A G A G A G A A A A A A G A G A G A G G A G G Y G R G A G A
G A G A A A A G A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A
G G A G G Y G R G A G A G A G A A A G A G A G A G A G G Y G G Q G G Y G A G A
G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G A G
G Y G G Q G G Y G A G A G A G A A A A A A G A G A S G G A G G Y G R G A G A
G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A A G A G A
G G A G G Y G R G A G A G A G A A A G A G A G A G A G G Y G G Q G G Y G A G A
G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G
G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G
A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G A
G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G
A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G
Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A
G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G
A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A
G A A A A A G A G A S G G A G G Y G R G A G A G A G A A A G A G A G A G G
Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A
G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G

A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A
G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y
G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G
A A A G A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A
G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G
A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G
G Q G G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G A G A G A
A A G A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A S G G D
P G S A S R L A S P D S G A R V A S A V S N L V S S G P T S S A A L S S V I S N
A V S Q I G A S N P G L S G C D V L I Q A L L E I V S A C V T I L S S S S I G Q V
N Y G A A S Q F A Q V V G Q S V L S A F A A **Stop**

MiSp24s

M T G G G S S A G N A F A Q S L S S N L L S S G D F V Q M I S S T T S T D E A V
S V A T S V A Q N V G S Q L G L D A N A M N N L L G A V S G Y V S T L G N A I
S D A S A Y A N A L S S A I G N V L A N S G S I S E S T A S S A A S S A A S S V T
T T L T S Y G P A V F Y A P S A S S G G S G G A G G Y G R G A G A G A G A A A
G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y
G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A
A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G
G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G A G A G A A A G
A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G
R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A
A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G
Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G A G A G A A A G A
G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R
G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A
G A G A S G G G S S A G N A F A Q S L S S N L L S S G D F V Q M I S S T T S T D
E A V S V A T S V A Q N V G S Q L G L D A N A M N N L L G A V S G Y V S T L
G N A I S D A S A Y A N A L S S A I G N V L A N S G S I S E S T A S S A A S S A
A S S V T T T L T S Y G P A V F Y A P S A S S G G S G G A G G Y G R G A G A G
A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G
G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G
A G A A A A A G A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G
Y G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A
G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G
A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A
G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y
G G Q G G Y G A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G
A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G G A
G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G
A A A A A G A G A G A S G G G S S A G N A F A Q S L S S N L L S S G D F V Q M I S

S T T S T D E A V S V A T S V A Q N V G S Q L G L D A N A M N N L L G A V S G
Y V S T L G N A I S D A S A Y A N A L S S A I G N V L A N S G S I S E S T A S S A
A S S A A S S V T T T L T S Y G P A V F Y A P S A S S G G S G G A G G Y G R G
A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G
A G A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G
A G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A
G A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A
G A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A
G A G G A G G Y G R G A G A G A G A A A G A G A G A G G A G G Y G G Q G G Y G A
G A G A G A A A A A G A G A G G A G G Y G R G A G A G A G A A A G A G A G
A G G Y G G Q G G Y G A G A G A G A A A A A G A G A G A G G A G G Y G R G A G
A G A G A A A G A G A G A G G Y G G Q G G Y G A G A G A G A A A A A G A G
A G G A G G Y G R G A G A G A G A A A G A G A G A G G Y G G Q G G Y G A G
A G A G A A A A A G A G A S G G D P G S A S R L A S P D S G A R V A S A V S N
L V S S G P T S S A A L S S V I S N A V S Q I G A S N P G L S G C D V L I Q A L L
E I V S A C V T I L S S S S I G Q V N Y G A A S Q F A Q V V G Q S V L S A F A
A **Stop**