Biomimetic manufacturing of fibers

F. Teule¹, S. Jung¹, J. Wood¹, W.R. Marcotte¹, M.S. Ellison² & A.G. Abbott¹
¹Department of Genetics and Biochemistry, Clemson University, USA
²School of Materials Science and Engineering, Clemson University, USA

Abstract

Several fiber protein encoding genes based on several published sequences of natural protein fibers such as silk and collagen were engineered. These engineered spider silk like homopolymer (spidroin 1 or spidroin 2) and collagen-silk copolymer genes were cloned and expressed in yeast (Pichia pastoris).

To test the importance of alanine residues in spidroin 1 proteins and their impact on the mechanical properties of the resulting fibers, several spidroin 1 constructs were made that contained sequences encoding for different amounts of alanine repeats in the protein (normal, low, and no alanine residues). One of the copolymer genes was also engineered in a plant vector and introduced in tobacco for expression. All these synthetic genes will be expressed in higher plants particularly seed storage protein producing plants such as legumes, as possible good candidates for large-scale production of newly designed fiber proteins. For production in plants, the use of natural high-level seed specific promoters is being investigated.

Expression of these synthetic genes in transgenic systems such as yeast and plants should provide sufficient quantities of recombinant protein for fiber and film production. With sufficient quantities of protein available, we are currently designing spinning technologies based on biological systems and we are investigating the role that various protein primary structural components play in fiber production.

1 Introduction

Like all more evolved spiders, the golden orb weaver spider (Nephila clavipes) possesses seven types of silk glands that she uses to make different types of silk depending on their use (e.g. web, cocoon, dragline). The spider then spins its
silks using one of the three sets of spinnerets located on the ventral part of her abdomen and pulls out the silk thread with its back legs. Those silks like silkworm silks are natural protein based fiber polymers. However, in comparison to silkworm silk, spider silks exhibit very different mechanical and physical properties. The most extensively studied silk is the dragline silk of *Nephila clavipes* that is produced by the spider’s major ampullate glands. This silk combines high extensibility and strength resulting in very tough fibers.

Dragline silk is made of two highly repetitive proteins called Masp1 and Masp2 (major ampullate spidroin 1 and 2). Two partial cDNAs sequences corresponding to genes encoding these two spidroins were isolated and published by Xu [1], and Hinman [2]. The primary structure of Masp1 and Masp2 is composed of alanine motifs sandwiched between glycine-rich (Masp1) or proline containing glycine-rich regions (Masp2). Hypotheses regarding the roles of individual protein structural motifs to the final fiber’s mechanical and physical properties were proposed by Hayashi [3] and need thorough testing.

Since spiders are cannibalistic, spider farming for silk production is not feasible. Alternative means to manufacture spider silk using the new tools provided by biotechnology are the subject of investigation by many laboratories worldwide. As a result, synthetic spider silk-like genes were cloned and expressed in *E. coli* or in yeast (*P. pastoris*) as reviewed by Heslot [4] and more recently in mammalian cells by Lazaris [5]. Attempts to artificially spin the genetically engineered proteins were reported by Lazaris [5]. In addition, Scheller [6] produced synthetic spidroin 1-like proteins in tobacco and potato demonstrating that plants can also serve as bioreactors to produce engineered fiber proteins.

For our research, *N. clavipes*’ major ampullate silk cDNA sequences (Masp1 and Masp2) were used as a model to set up a technology allowing the production and expression of designed fiber protein encoding genes in yeast (*Pichia pastoris*) and plants. We are currently investigating the relationship between primary structure components of a protein and their role in the resulting protein fiber physical and mechanical properties. In particular, customized spidroin 1 gene constructs were made to determine the role of alanine motifs in fiber proteins. Spidroin 1-like genes encoding variable amounts of alanine residues were made and the final structures and properties of the resulting recombinant proteins produced in yeast are the subject of current studies. In addition, we are also studying copolymeric proteins inspired from natural fibrous proteins such as those from which the mussel byssal threads are composed. In this case, the sequences encoding the helical part (GXY)$_n$ of a root-knot nematode cuticle collagen (*Meloidogyne incognita*) cloned by Wang [7] were juxtaposed to either spidroin 1 or spidroin 2 customized gene sequences.

Ultimately, for large-scale production of any designed fiber proteins, we are investigating plant transgenic production in seed storage protein-producing plants such as the legume species. For production in plants, we are exploring the use of a natural high-level seed specific promoter (gene regulatory sequence or switch) such as the peanut omega-9 desaturase gene promoter. Such a
promoter, once juxtaposed to our synthetic fiber protein encoding genes should target the fiber protein production and accumulation in plant seeds by switching on the gene expression in these tissues. The engineered fiber proteins could then be purified from the harvested plant seeds. As of now, we are using available promoters for plant transgenic production of our synthetic fiber proteins. We have introduced a collagen/spidroin 2 gene controlled by a constitutive CaM35S promoter in tobacco plants and the expression of this gene is currently under study.

2 Materials and methods

2.1 Synthetic fiber gene construction

2.1.1 Spidroin-like homopolymers
To construct both spidroin 1-like and 2-like genes, gene sequences corresponding to the consensus repeat sequences of *Nephila* clavipes' spidroin 1 or 2 proteins (Masp1 and Masp2) were used. For each spidroin gene, two or more sets of 5' phosphorylated complementary oligonucleotides spanning the corresponding gene consensus repeat were synthesized (IDT technologies, Inc).

Several types of spidroin 1 genes encoding for proteins containing variable amounts of alanine repeats were made using different sets of oligonucleotides: regular amount, lesser amount and no alanine repeat at all.

Each set of complementary oligonucleotides were annealed in 2X SSC at 65°C for 5 hours and cooled down slowly to room temperature. Two point five micrograms of each appropriate annealed reaction were ethanol precipitated. DNA pellets were briefly dried and then resuspended in 1X T4 DNA ligase buffer (3mM Tris-HCl pH 7.8, 1mM DTT and 0.1 mM ATP) and 3 units of T4 DNA ligase (Promega) were added to the 20µl ligation reactions. The head to tail ligation reactions were performed overnight at 4°C and produced multimeric constructs that were electrophoretically separated on 0.8% agarose gels. A size selection of gene products ranging between 500bp and 1000bp was performed by electroelution following the protocol described by Sambrook [8]. Once recovered, these products were “filled” and “A-tailed” at the same time by incubation 1 hour at 72°C with Taq DNA polymerase (Fisher Scientific) and dNTPs. The 50µl filling reactions contained 2µg of size-selected constructs mixed with 1X Taq DNA polymerase Buffer A (50 mM KCl, 10 mM Tris-HCl pH 9, 1.5 mM MgCl2), 0.1 mM dNTPs and 5 units of Taq DNA polymerase. These genes products were cloned in a TA cloning plasmid vector pGEMT-Easy (Promega) and sequenced using an ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer). Sequences were resolved on an ABI 373 Stretch sequencer (Applied Biosystems Inc.).

2.1.2 Collagen/spidroin-like copolymers
All the following PCR reactions were performed using a high fidelity *Pfu* DNA polymerase (Stratagene).
To obtain the collagen-like gene sequence, a PCR reaction was performed using the *mi-col-2* collagen cDNA clone isolated by Wang [7] as a template. Specific 5' and 3' PCR primers designed to add an EcoRI and a BamHI restriction sites in 5' and 3' respectively were used to amplify the 423 bp (GlyXY) (Gly = Glycine, X and Y= any amino acid) coding region of the *mi-col-2* gene.

Using specific primers designed to add a Bam HI restriction site and a Sac II restriction sites in 5' and 3' respectively, several PCR reactions were performed to amplify the three types of cloned spidroin 1-like genes (normal, less and no alanine encoding genes) ranging between 530bp and 600bp, and a 108 bp cloned spidroin 2-like gene.

All of these collagen, spidroin 1-like and spidroin 2-like amplified products were separately cloned in pGEMT-Easy and used to transform competent *E. coli* strain JM109 (Promega). The recombinant plasmids were sequenced using M13 universal primers. The plasmids were then double digested with Bam HI and Sac II restriction enzymes for spidroin 1-like and spidroin 2-like clones and with Eco RI and Bam HI enzymes for the collagen clone. All these Eco RI/Sac II inserts were purified by electroelution, ethanol precipitated and stored at -20°C until use for collagen/spidroin 1 or 2-like copolymer gene construction.

2.1.2 Cloning and expression of the synthetic genes in yeast

2.1.2.1 Homopolymer gene cloning The same PCR-based strategy allowing the addition of restriction enzyme sites on each end of the cloned genes was used. Specific primers designed to add an Eco RI and a Sac II restriction sites in 5' and 3'respectively were used to PCR amplify each of the three types of cloned spidroin 1 homopolymers. The amplification products were cloned in a pGEMT-Easy vector and used to transform competent *E. coli* strain JM109 (Promega). The recombinant plasmids were sequenced. The plasmids containing each of the spidroin-like genes were each submitted to an Eco RI/Sac II double restriction digestion. The isolated Eco RI/Sac II genes were ligated separately to a yeast expression vector pPICZαA (Easy select™Pichia Expression Kit, Invitrogen) previously digested with the same cloning enzymes.

2.1.2.2 Collagen/spidroin 1 or 2 copolymer gene cloning Copolymer genes of collagen/spidroin 1 or 2 were “assembled” directly in the pPICZαA shuttle vector for collagen/spidroin 1 genes and in the pPICZαC vector for collagen/spidroin 2 gene. Both vectors were submitted to an Eco RI/Sac II restriction digestion and were electroeluted separately following the protocol described by Sambrook [8]. The vector, collagen insert and spidroin 1 or 2-like inserts were mixed and ligated with T4 DNA ligase following the instructions specified by the manufacturer (Promega). The different ligation reactions were used separately to transform competent *E. coli* strain TOP10F’ (Invitrogen). The recombinant plasmids containing the collagen/spidroin 1-like copolymer genes (three types) and the collagen/spidroin 2-like copolymer genes were sequenced using the 5’AOX1, 3’AOX1 and α Factor primers specific to the shuttle vector.
All the yeast transformations of Chlorella pastoris strain GS115 and the following expression studies were performed as described by the manufacturer (Invitrogen). In the case of spidroin 1-like and collagen/spidroin 1-like recombinant yeast, “super expressers” were obtained by plating the transformation reaction on a media containing 5 times the normal zeocin concentration.

2.2 Protein extraction and purification

Spidroin-like and collagen/spidroin-like proteins were produced and extracted from recombinant yeast cells following the manufacturer’s recommendations (Invitrogen). Collagen/spidroin 2 proteins secreted by recombinant yeast clones were harvested from the culture media after removing all the yeast cells by centrifugation. The supernatant was submitted to a 30% ammonium sulfate cut allowing the recovery of the recombinant proteins after centrifugation, resuspension, and dialysis against distilled water. Further purification of the histidine tagged proteins was done using Talon Metal affinity chromatography (Clontech). Subcellular fractionation of recombinant yeast cells was performed for the spidroin 1-LA and the collagen/spidroin 2 clones according to the protocol described by Ouzzine [9]. Protein concentrations were evaluated by the Bradford method using a BCA Assay (Pierce). To purify the spidroin 1 LA proteins from the other yeast cellular proteins, the crude cellular protein extracts were heated at 95°C for 5, 10 and 15 minutes and the samples were subjected to centrifugation at 16000g for 10 minutes at room temperature. The supernatant of the heat-treated protein samples was recovered and analyzed.

2.3 SDS-PAGE and Western Blot analysis

Recombinant proteins (intracellular or secreted fractions) were subjected to SDS-PAGE analysis on 4%-12% polyacrylamide gels followed by Coomassie-blue staining or Western blot analysis using an anti myc epitope antibody detection system (ECL™ Western Blotting Analysis Systems, Amersham Pharmacia Biotech). A prestained protein marker broad range (New England Biolabs) was used for this analysis.

2.4 Tobacco transgenics

2.4.1 Copolymer gene cloning in Agrobacterium tumefascien

The collagen/spidroin 2 gene cloned in pPICZαC was amplified by PCR using primers adding an Nco I site and an Xba I site in 5’ and 3’ respectively. The amplified gene product was subjected to a Nco I/Xba I double restriction digestion and ethanol precipitated. The pBM437 cloning vector containing a CaMV35S promoter was also subjected to an Nco I/Xba I double restriction digestion, electroeluted and ethanol precipitated. The Nco I/Xba I copolymer gene was ligated to the modified pBM437 vector at the Nco I and Xba I restriction sites and used to transform competent E. coli strain JM109.
The recombinant plasmids were sequenced. The 5' Hind III/3' Eco RI region of the recombinant plasmid, containing the CaMV35S promoter juxtaposed to the copolymer gene region, was isolated by restriction digestion with these two enzymes, purified by electroelution, and ethanol precipitated. The plant vector pBI101.1 (derived from pBI121) was subjected to restriction digestion with Hind III and Eco RI, electroeluted and ethanol precipitated. Purified insert and plant vector were ligated to each other using T4 DNA Ligase (Promega) and the recombinant pBM460 plasmid was used to transform Agrobacterium tumefaciens strain JM101.

2.4.2 Tobacco transformation using *Agrobacterium tumefaciens*

Tobacco leaves were transformed with the recombinant *A. tumefaciens* following a protocol described by Lin and others [10]. Kanamycin resistant transformants were obtained and screened by PCR using vector specific primers flanking the copolymer gene region to confirm the presence of the synthetic gene. Seeds were harvested from the identified primary transformants (T0) and grown on MS complete media containing kanamycin to obtain T1 seedlings.

2.4.3 RT-PCR analysis:

Total RNA was extracted from T1 seedlings’ leaves following the protocol described by Chomczynski and Sacchi [11]. For RT-PCR analysis, first strand cDNAs synthesis was performed using 1 μg of DNase treated total RNA and 5 units of AMV-RT enzyme (Promega). RT-PCR reactions were performed according to the protocol described by Nelsen and Marcotte [12]. Following amplifications using first strand cDNAs as templates were carried on using primers specific to the copolymer gene or to a tubulin gene included as a positive control. PCR reactions using gene specific primers were performed to amplify the copolymer collagen/spidroin 2 gene insert from the pBM460 plasmid clone and the tubulin gene from the NP16 plasmid clone. Aliquots of the RT-PCR and PCR reactions were separated by electrophoresis on a 0.7% agarose gel. The gel was stained with ethidium bromide for visualization under UV light.

3 Results

3.1 Yeast expression experiments

Strategies enabling the production and cloning of repetitive gene sequences were designed for the engineering of novel gene sequences specifying structural proteins such as silk spidroins. We have developed and utilized a method to make synthetic spidroin 1 or 2-like multimeric genes (based on Masp1 and Masp2 sequences) and collagen-spidroin-like copolymer genes based on ligation and PCR reaction using specific primers. Small homopolymer genes (spidroin 1 or 2-like) and copolymer genes (collagen spidroin 1 or 2-like) were engineered and cloned in yeast for protein production. All the gene constructs were engineered in a vector that allowed the secretion of the recombinant proteins...
thus facilitating their purification. Although the multimerization process can easily generate genes of a size superior to 1kbp (up to 6 kbp), constructs of 500 bp were size selected for our study for two reasons: 1) to confirm by sequencing that the synthetic genes do not contain any errors or missing bases due to defective oligonucleotides, and 2) to build gene cassettes encoding for different amino acid motifs that could later be put together to form bigger genes.

3.1.1 Spidroin 1-like constructs
Polyalanine motifs present in spider silk spidroins or silkworm silk fibroins are thought to be responsible for the fiber’s strength by forming β-sheet crystalline structures in the final fiber. To test this hypothesis, spidroin 1 gene constructs containing different amounts of alanine motifs: regular amount (ALA), lesser amount (LA) and no alanine motifs (NA) were engineered (Figure 1).

A spidroin 1 gene of 534 bp encoding for less alanine residues (SPI1LA17) was made and introduced in yeast for protein production. Since spidroin-like proteins are very weakly stained by most standard protein stains, Western blot analysis was necessary to determine the presence of the recombinant proteins. Western blot analysis showed the presence of the spidroin 1 LA proteins in the yeast cells (Figure 2).
Our results show that the protein was weakly secreted in the culture media and that most of it was sequestered in the microsomal fraction of the cells making its purification a little more difficult due to the abundance of native yeast proteins.

Figure 2: Spidroin 1-LA production by recombinant yeast.
A: SDS-PAGE analysis of the spidroin 1 LA protein produced by the recombinant yeast followed by Coomassie blue staining. B is the Western blot analysis of A. For A and B, lane 1: prestained protein marker broad range (New England Biolabs); lane 2: culture media; lane 3: intracellular soluble protein fraction; lane 4: intracellular microsomal protein fraction. The arrows denote the spidroin 1 LA protein (a is a 20 KDa monomer; b is a 40 KDa dimer, and c is a 80 KDa tetramer approximately).

However, heat treatment of the cellular protein extract at 95°C for 5 to 15 minutes resulted in the elimination the majority of the native cellular proteins allowing the recovery of the spidroin 1 LA protein (Figure 3). We are currently testing methods to further purify this protein in order to obtain a sample for structural characterization and spinning experiments.

Figure 3: Purification of spidroin 1 LA proteins by heat treatment
A: SDS-PAGE analysis of the spidroin 1 LA protein extracted from the recombinant yeast followed by Coomassie blue staining. B is the Western blot analysis of A. For A and B, lane 1: prestained protein marker broad range (New England Biolabs); lane 2: cellular protein extract; lanes 3, 4, and 5: recovered cellular protein extracts after heat treatment at 95°C for 5, 10 and 15 minutes respectively. The arrows denote the proteins (a is a 20 KDa monomer form, b is a 30 KDa monomer form with peptide signal, and c is a 40KDa dimer form approximately).
Two other spidroin 1-like clones of each category (ALA and NA) were obtained. The spidroin 1 ALA clones obtained are currently being characterized. A spidroin 1-NA homopolymer gene is currently being engineered in the yeast vector and will soon be introduced in yeast for protein production.

Many spidroin 2 genes of different sizes were obtained and one of them was introduced in yeast for protein production. Results presented by Teulé [13] indicated that the spidroin 2 protein was produced but was not secreted by the recombinant yeast.

3.1.2 Collagen/spidroin-like constructs
A copolymer collagen/spidroin 1 LA was engineered. The expression pattern study of this clone is currently under way and will be compared with the spidroin 1 LA’s. A spidroin 1-NA copolymer gene is currently being engineered in the yeast vector and will soon be introduced in yeast for protein production.

A copolymer collagen/spidroin 2 gene was made using a 108 bp spidroin 2 sequence juxtaposed to a customized collagen sequence and cloned in yeast for expression (Figure 4). Western blot analysis of culture media and cellular protein fractions (soluble and microsomal) showed that the protein was produced and successfully secreted. However, although multimeric forms of the protein were detected in the cellular fractions, only the dimeric form of the protein was secreted. No monomeric form of the collagen spidroin 2-like protein was detected. Aggregation or protein multimerization could be caused by naturally occurring cross-linking reactions involving lysine residues present in the collagen part of the recombinant protein. Affinity chromatography was successfully used to purify this copolymer protein.

Figure 4: Collagen/spidroin 2 protein production in yeast.
A: SDS-PAGE analysis of the copolymeric protein produced by the recombinant yeast followed by Coomassie blue staining. B is the Western blot analysis of A. For A and B, lane 1: prestained protein marker broad range (New England Biolabs); lane 2 culture media; lane 3: intracellular soluble protein fraction; lane 4: intracellular microsomal protein fraction. The arrows show the detected recombinant collagen/spidroin 2 proteins that are mostly dimeric (a is 38.8 kDa and b is 42.3 kDa), and trimeric forms (c is 54.9 kDa).
3.2. Plant transgenics experiments

The gene encoding the collagen/spidroin 2 copolymer was also engineered in a plant vector and introduced in tobacco using *Agrobacterium tumefascien* mediated transformation. RT-PCR experiments performed using total RNA extracted from leaf tissues of T1 generation seedlings showed that the copolymer gene is being transcribed in some transgenic lines (Figure 5). RT-PCR amplification of a tubulin gene from the same tobacco lines tested was used as a positive control. PCR amplifications of each copolymer and tubulin genes from plasmid clones were also performed to confirm that the T1 transformants contained the correct size genes. Efforts are under way to characterize the recombinant copolymer protein in these transgenic lines.

![Figure 5: RT-PCR analysis of tobacco T1 transformants. Stained agarose gel showing the RT-PCR and PCR amplification products.](image)

Lane 1: λ*Pst* I DNA molecular weight marker, lane 2: PCR amplification of the cloned copolymer gene (pBM460 plasmid), lanes 3 and 4: RT-PCR amplification of the copolymer gene transcript from two separate transgenic tobacco lines, lane 5: PCR amplification of a cloned tubulin gene (plasmid clone NP16), lanes 6 and 7: RT-PCR amplification of a tubulin gene transcript from the same two transgenic tobacco lines. The amplified copolymer collagen/spidroin 2 and tubulin genes are indicated (a and b respectively).

4 Conclusion

Natural fibrous proteins are numerous (silk, collagen, elastin, lamprin) and all of them display outstanding physical properties such as elasticity, strength, and toughness. They are all built using the same tools (amino acid motifs) with varying composition and repeat numbers from one molecule to another. It is remarkable that biomaterials like spider silk, or collagenous fibers such as Achilles’ tendon or mussel byssal threads are biodegradable, manufactured under biological conditions (temperature and pressure of the living organism), and superior to the best synthetic polymer material (often synthesized in extreme conditions). This is why there is such an interest in engineering similar or novel biobased protein fiber materials. The challenge is first to develop an understanding of the rules that govern the relationship between primary protein structure and final fiber properties, second to create a way to abundantly
produce the designed proteins, and finally to be able to spin the resulting recombinant fiber proteins into fibers.

In this paper, we report the development of a strategy to produce synthetic fiber proteins based on natural fibrous proteins such as spider silk and collagen. Protein structural components such as polyalanine motifs are being investigated to determine their impact in the strength of the resulting fiber. Other components need to be studied such as the GPGXX motif and GGX motif present in spider silk spidroins to establish their roles. Results from our study should be informative and are the first step in order to design customized novel fiber proteins with desired properties. Plants such as legumes represent ideal candidates for large-scale production of these novel fiber proteins and should provide enough material for the implementation of fiber spinning technologies.

Acknowledgments

We would like to thank Dr R. S. Hussey (University of Athens, GA) for supplying the mi-col-2 clone and the National Textile Center for their financial support.

References

Design and Nature


