



# Advances in Plant-Derived Scaffold Proteins

Congyue Annie Peng\*, Lukasz Kozubowski and William R. Marcotte Jr\*

Department of Genetics and Biochemistry, Clemson University, Clemson, SC, United States

Scaffold proteins form critical biomatrices that support cell adhesion and proliferation for regenerative medicine and drug screening. The increasing demand for such applications urges solutions for cost effective and sustainable supplies of hypoallergenic and biocompatible scaffold proteins. Here, we summarize recent efforts in obtaining plant-derived biosynthetic spider silk analogue and the extracellular matrix protein, collagen. Both proteins are composed of a large number of tandem block repeats, which makes production in bacterial hosts challenging. Furthermore, post-translational modification of collagen is essential for its function which requires co-transformation of multiple copies of human prolyl 4-hydroxylase. We discuss our perspectives on how the GAANTRY system could potentially assist the production of native-sized spider dragline silk proteins and prolyl hydroxylated collagen. The potential of recombinant scaffold proteins in drug delivery and drug discovery is also addressed.

**Keywords:** spider silk, collagen, elastin, scaffold, regenerative medicine, extracellular matrix

## OPEN ACCESS

### Edited by:

Poul Erik Jensen,  
University of Copenhagen, Denmark

### Reviewed by:

Randy V. Lewis,  
Utah State University, United States  
John George Hardy,  
Lancaster University, United Kingdom

### \*Correspondence:

Congyue Annie Peng  
congyup@clemson.edu  
William R. Marcotte Jr  
marcotw@clemson.edu

### Specialty section:

This article was submitted to  
Plant Biotechnology,  
a section of the journal  
Frontiers in Plant Science

**Received:** 16 November 2019

**Accepted:** 27 January 2020

**Published:** 25 February 2020

### Citation:

Peng CA, Kozubowski L and Marcotte  
WR Jr (2020) Advances in Plant-  
Derived Scaffold Proteins.  
Front. Plant Sci. 11:122.  
doi: 10.3389/fpls.2020.00122

## INTRODUCTION

Scaffold proteins, synthetic or natural, provide promising innovative solutions to regenerative medicine (Lavine et al., 2012; Bourzac, 2015; Gould, 2015). A broad range of functional proteins with superb biocompatibility and biodegradability such as helical collagen (Parry et al., 2005), elastin (Mithieux and Weiss, 2005), silkworm silks (Dicko et al., 2006), and spider silks (Gosline et al., 1999) are candidates for proteinaceous scaffold biomaterials. The source of proteins to be processed to biomaterial scaffolds is often limited, unsustainable, or sometimes carries the risks of human pathogen contamination (Wong Po Foo and Kaplan, 2002). Recombinant production systems, such as mammalian cells, insect cells, silkworms, yeast (*Pichia pastoris*), *Escherichia coli*, and plants provide opportunities to produce scaffold proteins in full length or as representative motifs. Recombinant production of scaffold proteins also allows the flexibility of engineering variants or combinations of motifs that are difficult or impossible to obtain from natural sources. Each system has gained some successes, but the intrinsic nature of the host systems and economic feasibility are often a source of limitations.

Plant host systems demand less in energy input, chemical reagents, and contamination management (Shoseyov et al., 2014). In theory, sustainable and cost-effective production of recombinant scaffold proteins are feasible using plant hosts (Scheller and Conrad, 2005). Depending on the recombinant

proteins, eukaryotic post-translational modifications can presumably be achieved inside plant cells as the proteins mature (Ruggiero et al., 2000). This review summarizes the advances and challenges in plant host-derived scaffold proteins: spider silk protein, collagen, elastin, and a bone matrix protein, with the focus on fusion tags that facilitate protein accumulation in the endoplasmic reticulum (ER) and subsequent protein purification. Techniques with the potential to improve yield to meet the requirement of economical application, methods allowing efficient transformation of several genes or large repeat domains, and the utilization of scaffold proteins for drug screening are also discussed.

## SPIDER SILK-BASED SCAFFOLD PROTEIN PRODUCED FROM PLANTS

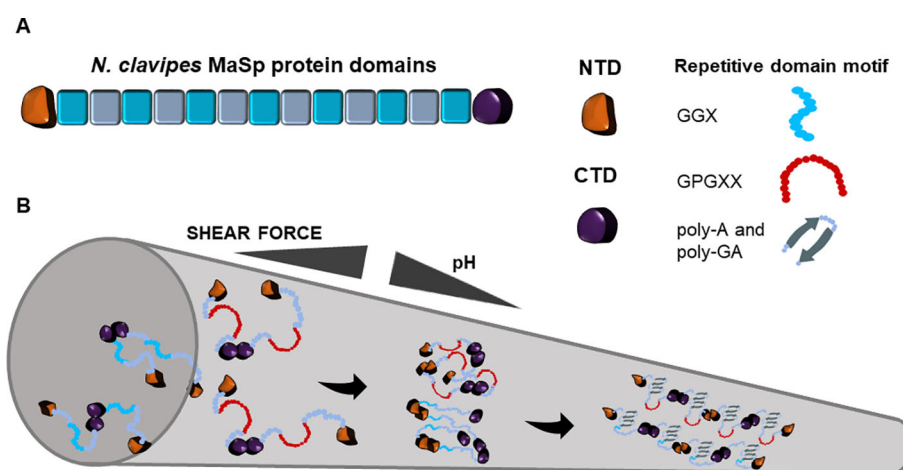
### Major Ampullate Spidroin Sequence, Structure and Folding

Both spider silk and silkworm silks are superior materials for tissue regeneration (Holland et al., 2019). The abundant availability of silkworm silk extracted from silkworm cocoons permits its extensive application in tissue engineering through materials fabrication (Huang et al., 2018). Purification of silkworm silks requires a degumming process to remove sericin, a protein component in the silkworm cocoon and the outer layer of the silkworm silks that stimulates immune response (Altman et al., 2003). Spider dragline silk, the life-line of an orb-weaving spider, has impressive mechanical strength and toughness, which makes it a potential candidate for applications such as bone regeneration (Gosline et al., 1999; Gomes et al., 2011; Andersson et al., 2016; Lee M. et al., 2016). In addition, spider silk propagates light as an optical fiber (Huby et al., 2013). Although the cannibalistic nature of

spiders precludes the possibility of farming (Nentwig, 2013), recombinant production of spider silk proteins provides alternatives to obtain spider silk-like proteins. Recombinant production also allows generation of biomaterials in a diverse range of forms, such as hydrogels, films, coatings, meshes, and nanoscale particles and tubes.

### Major Ampullate Spidroin Types

Two major spidroin proteins were identified from *Nephila clavipes* spider dragline silk, namely major ampullate spidroins 1 and 2 (MaSp1, MaSp2). Each protein contains conserved tandem block repeats and the flanking non-repetitive N- and C-terminal domains that are conserved in all orb-weaving spiders (Xu and Lewis, 1990; Hinman et al., 1992, **Figure 1**). The predicted 3.45 Gb genome of *N. clavipes* encodes eight potential MaSps, ranging in sizes from ~100 to over 3,000 amino acids with motifs mapped to the originally discovered MaSps. To differentiate MaSps discovered from the genome assembly, the new nomenclature set of MaSp-a through MaSp-h is used (Babb et al., 2017). Studies prior to the genome assembly have used the previous nomenclature: MaSp1 and MaSp2. MaSps are diverse but some motifs are conserved among arachnid species (Gatesy et al., 2001). Proteotranscriptomic study of the *N. clavipes* major ampullate gland detected transcripts and protein products of the two primary spidroins: *MaSp1* and *MaSp2*. Importantly, proteins involved in ion transport, folding and conformation regulation, post-translational modification, and fibrillar preservation and protection were simultaneously detected, indicating the complexity of spidroin production and the following transition into solid fibers (Santos-Pinto et al., 2019). Despite similar domain structures, the isoelectric point of the two spidroins differs. Isoelectric focusing analysis indicate the pI of MaSp1 is above 8.5 and the pI of MaSp2 is between 5.1 and 5.9 (Sponner et al., 2005a).



**FIGURE 1** | Schematic representation of *Nephila clavipes* major ampullate spidroin domains **(A)** and the fiber assembly model in major ampullate spinning duct **(B)**. The colored blocks represent repeat blocks with the color variation indicating the glycine-rich motif and the poly-alanine motif. Along the tapered spinning duct, with the increase of shear force and the decrease of pH, *N. clavipes* major ampullate spidroins transit from the liquid unfolded state into oligomeric state through the dimerization of N-terminus and the facilitation of C-terminus. The oligomeric state ultimately transitions into a two-phased state, which is comprised of unordered amorphous regions and ordered  $\beta$ -sheet crystalline regions.

### Major Ampullate Spidroin Repeat Domain

The tandem block repeat domains of MaSp1 and MaSp2 differ slightly from each other but can be generalized into the following consensus sequences (Hayashi et al., 1999):

MaSp1: (GGA) GQ (GGY) (GGL) (GGQ) GAGR (GGL) (GGQ) (GA)<sub>2</sub> (A)<sub>3</sub>

MaSp2: (GPGGY) (GPGQQ) (GPGGY) (GPGQQ) GPSGPS (A)<sub>9</sub>

The ~100 block repeats cover about 95% of each protein sequence. The main difference between MaSp1 and MaSp2 is that MaSp1 repeat region is essentially devoid of proline. Proline is believed to contribute to the major structural and mechanical difference between MaSp1 and MaSp2 (Marhabaie et al., 2014). The proline kink in the polypeptide backbone causes a reduction in hydrogen bond donor capacity and adversely affects  $\beta$ -sheet extension of the poly alanine residues.

### Major Ampullate Spidroin C-Terminal Domain (CTD)

The short non-repetitive C-termini of MaSp1 and MaSp2 share 75% identity (Beckwitt and Arcidiacono, 1994). Three conserved sites were identified between *N. clavipes* MaSp and minor ampullate spidroin (MiSp) CTDs (Collin et al., 2018). The CTD folds into five parallel helices and forms a homodimer stabilized through a disulfide bond between conserved cysteine residues and two salt bridges (Hagn et al., 2010). *In vitro* experiments show that recombinant CTDs attached to the repeat domain lead to faster aggregation of the repeat domain and are more susceptible to buffer-induced fibril assembly (Huemmerich et al., 2004a; Huemmerich et al., 2004b; Sponner et al., 2005b; Hedhammar et al., 2008). The CTD reinforces protein alignment, exposes its hydrophobic region to attract CO<sub>2</sub> for binding, and facilitates  $\beta$ -amyloid like nucleation (Ittah et al., 2006; Eisoldt et al., 2010; Andersson et al., 2014).

### Major Ampullate Spidroin N-Terminal Domain (NTD)

The non-repetitive, hydrophilic NTD is conserved among species, within different types of spider silk proteins (dragline, flagelliform, and cylindrical silk), and among waxmoth and silkworm fibroins (Bini et al., 2004; Motriuk-Smith et al., 2005; Rising et al., 2006; Collin et al., 2018). In spun dragline silk, the NTD can be detected in both inner and outer core regions (Andersson et al., 2013), indicating it is not being cleaved in the process of fiber assembly. MaSp1 NTD is composed of five anti-parallel  $\alpha$ -helices, forms thermally stable homodimers at pH 6, and restricts CO<sub>2</sub> binding (Askarieh et al., 2010; Hagn et al., 2011; Andersson et al., 2014). Asymmetric interaction of amino acid residues and salt bridge formation is detected and is believed to contribute to the plasticity of the *N. clavipes* NTD (Atkison et al., 2016). Folding at the area surrounding a conserved tryptophan residue in the stable NTD homodimer compels the downstream repeat domains to initiate  $\beta$ -sheet orientation (Askarieh et al., 2010; Gaines et al., 2010). NTD motifs, essential for folding and response to acidic pH, are structurally homologous between species (Heiby et al., 2017). The dipole-dipole interaction of NTD monomers at acidic pH may reduce the free energy barrier required to initiate dimer association

and become the driving force for fiber assembly (Ries et al., 2014; Barroso da Silva et al., 2016).

### Major Ampullate Spidroin Post-Translational Modification (PTM)

Major ampullate spidroins obtained from dragline silk fibers and gland extractions contain post-translational modifications, such as tyrosine or serine phosphorylation (Michal et al., 1996). L-Dopa (3,4-dihydroxyphenylalanine) and dityrosine are detected from hydrolyzed major dragline silk solutions, but it is not clear if this result may be the result of tyrosine oxidation during sample preparation or the exposure of fibers post-spinning to uv radiation (Santos-Pinto et al., 2014). Phosphorylation sites found in MaSp1 isoforms A and B reside almost exclusively in the GGX region of the repeat domain (Santos-Pinto et al., 2015). For MaSp2, phosphorylation sites are found both in the repeat domain and C-terminus. The phosphorylation sites of MaSp2 repeat domain reside at identical positions within each repeat domain (Santos-Pinto et al., 2016). The reason why only certain locations in the repeat domain are phosphorylated is unknown. We speculate that these special locations may reflect the inter- and intra-molecular interaction in the 3D structure after protein folding, through which a hierarchical network can be formed and eventually lead to the fiber assembly.

*N. plumipes* dragline silk displays an arginine hydrogen bonding with the amorphous regions, and hydroxyproline is detected through Dynamic Nuclear Polarization (DNP) NMR spectroscopy (Craig et al., 2019). Peptide glycosylation is detected in the solubilized dragline silk of *N. clavipes*, although the oligosaccharide molecule is still unknown (Guehrs et al., 2008). While PTM may play a critical role in fiber assembly, the mechanisms of PTM of spidroins in spider ampullate gland are poorly understood. The recent proteomic data of the gland may shed some light on the mechanism involved in PTM of spidroins (Santos-Pinto et al., 2019).

### Molecular Basis of Major Ampullate Spidroin Biomaterials Assembly

The transition from liquid crystalline folding of spidroins to solid fibers in the spider duct is a complex process that is far from being understood (Vollrath, 2016). The current model on spidroin fiber assembly is shown in **Figure 1**. Progress has been made to capture the fiber assembly initiator structures such as transient oligomers (Landreh et al., 2017) or oligomeric micelles (Bauer and Scheibel, 2017), providing evidence that fiber assembly may involve formation of initial “seed” scaffolds for further nucleation of additional spidroins. The transition requires 300–700 MPa shear force and at least 6 poly alanine modules (Giesa et al., 2016). In the solid fiber state, 40% of alanine forms oriented  $\beta$ -sheet stacked crystallites and about 60% of alanine folds into poorly oriented  $\beta$ -sheet (Simmons et al., 1996; Ene et al., 2009). The ordered and amorphous regions provide the molecular basis that allows physical strength and extensibility in one fiber (Patil et al., 2014; Xu et al., 2014). This two-phased structural arrangement also contributes to phonon propagation bandgap along the fiber axis (Schneider et al., 2016).

## Plant-Derived Spidroins

Recombinant spidroin analogs have provided useful testing materials for *in vitro* assessment of issues such as protein solubility and/or premature aggregation, translational pause, and post-translational modification. Host system selection, however, impacts the quality and quantity of the recombinant protein production. Each host system offers promising possibilities but possesses also some limitations.

To attain the desired mechanical properties, recombinant major ampullate spidroin proteins (rMaSp1 and rMaSp2) are expected to contain large numbers of repeat domains, ideally equivalent to the native protein (~100 repeat domains) and produce a large protein with molecular weight greater than 300 kDa. Engineering ~100 repeat domains is a challenge for cloning, protein solubility during purification, and host transcriptional regulation (Guerette et al., 1996; Lewis et al., 1996; Fukushima, 1998; Xia et al., 2010; Hauptmann et al., 2013). The frequent demand for alanine and glycine depletes the correspondent tRNA pools quickly and causes translational pause (Candelas et al., 1990). With the advances in function elucidation of the non-repetitive terminal domains, many efforts are devoted to incorporation of native N- and C-termini in recombinant spidroins to enhance solubility, protein alignment, and fibril assembly (Ittah et al., 2006; Barroso da Silva et al., 2016). Efforts to functionalize rMaSp-based biomaterials to have desired property include additives or linkers that recognize cell binding site, or are targeted to the affinity binding domains (Bini et al., 2006; Widhe et al., 2013; Jansson et al., 2014) that improve physical properties (Teulé et al., 2012), and that control tertiary structures through phosphorylation and dephosphorylation (Winkler et al., 2000).

Although not specifically discussed in the literature, other potential challenges may include: 1. the accumulation of negatively charged rMaSp1 that may cause toxicity to the host cell, 2. the post-translational modification system in the host may not match that of the spiders, 3. lack of economic feasibility to obtain a desired yield of rMaSp, 4. challenges with optimization of the process of fiber assembly and biomaterials development.

Non-plant hosts that have been used to produce major ampullate spidroins include *E. coli* and *Salmonella* (Lewis et al., 1996; Arcidiacono et al., 1998; Widmaier et al., 2009; Xia et al., 2010; Edlund et al., 2018), yeast *Pichia* (Fahnestock and Bedzyk, 1997; Gaines and Marcotte, 2011; Liu et al., 2018), protozoa *Leishmania* (Lyda et al., 2017), mammalian cell lines (Lazaris et al., 2002), silkworm transformed with fusion protein or CRISPR/Cas9 site specific exchange (Miao et al., 2006; Teulé et al., 2011; Zhang et al., 2019), and mice or goats (Service, 2002; Xu et al., 2007). Here we focus on the efforts and strategies of recombinant spidroin production from plants.

### Targeted Recombinant Spidroin Accumulation

Recombinant production localized at the endoplasmic reticulum (ER), using the ER retention signal KDEL at the C-terminus, can improve the yield of recombinant proteins as evidenced by the results summarized in **Figure 2**. Thermally stable MaSp1 repeat domain analogs (up to 100 kDa) with a C-terminal KDEL ER retention signal accumulated up to 2% of total soluble protein (TSP) in tobacco and potato leaf (Scheller et al., 2001), which

demonstrated for the first time that rMaSp can be produced from plant host. The yield of rMaSp targeted to ER is four times higher as compared to the MaSp1 analogs 1f5 and 1f9 (fusion protein with a tetramer or octamer of the repeat domain) that accumulated in tobacco leaf without an ER target signal (0.5% of TSP) (Piruzian et al., 2003). MaSp2 and MaSp1/MaSp2 analog rADF-3 (containing the repeat motifs ASAAAAA, GPGGQGPYGP, GGYGPGS, and (GPGQQ)<sub>n</sub>) were also targeted to the ER (using KDEL) and produced in tobacco leaves (Menassa et al., 2004). The maximum production yield for ER-directed MaSp1 and MaSp2 analogs were 0.68 and 3.05 mg/kg fresh leaf tissue, respectively. The production of spidroin-like proteins was successfully retained when stably-transformed tobacco plants were transferred into the fields.

MaSp1 synthetic analogues (8 or 16 copies of DP1B monomer sequence GQGGYGLGSQAGRGLGGQAGA<sub>7</sub>GGA) were transformed into *Arabidopsis* using the cauliflower mosaic virus (CaMV) 35S promoter for leaf expression and the β-conglycinin α' subunit promoter for *Arabidopsis* seed and somatic soybean embryo expression. Although *Arabidopsis* seeds showed better recombinant spidroin recovery, soybean embryos produced lesser amounts of the 127 kDa DP1B 16-mer (Barr et al., 2004), indicating that the length of the recombinant spidroin repeat domain affects tissue specific production yield. Yang and colleagues compared the yield of the same synthetic MaSp1 analogs targeted to apoplast, ER lumen, and vacuole in *Arabidopsis* leaf or seeds using sporamin-targeting determinant peptides and the ER retention KDEL peptide. Transgenic plants with seed-specific ER targeting constructs produced the highest amount of recombinant spidroins (18% TSP) (Yang et al., 2005).

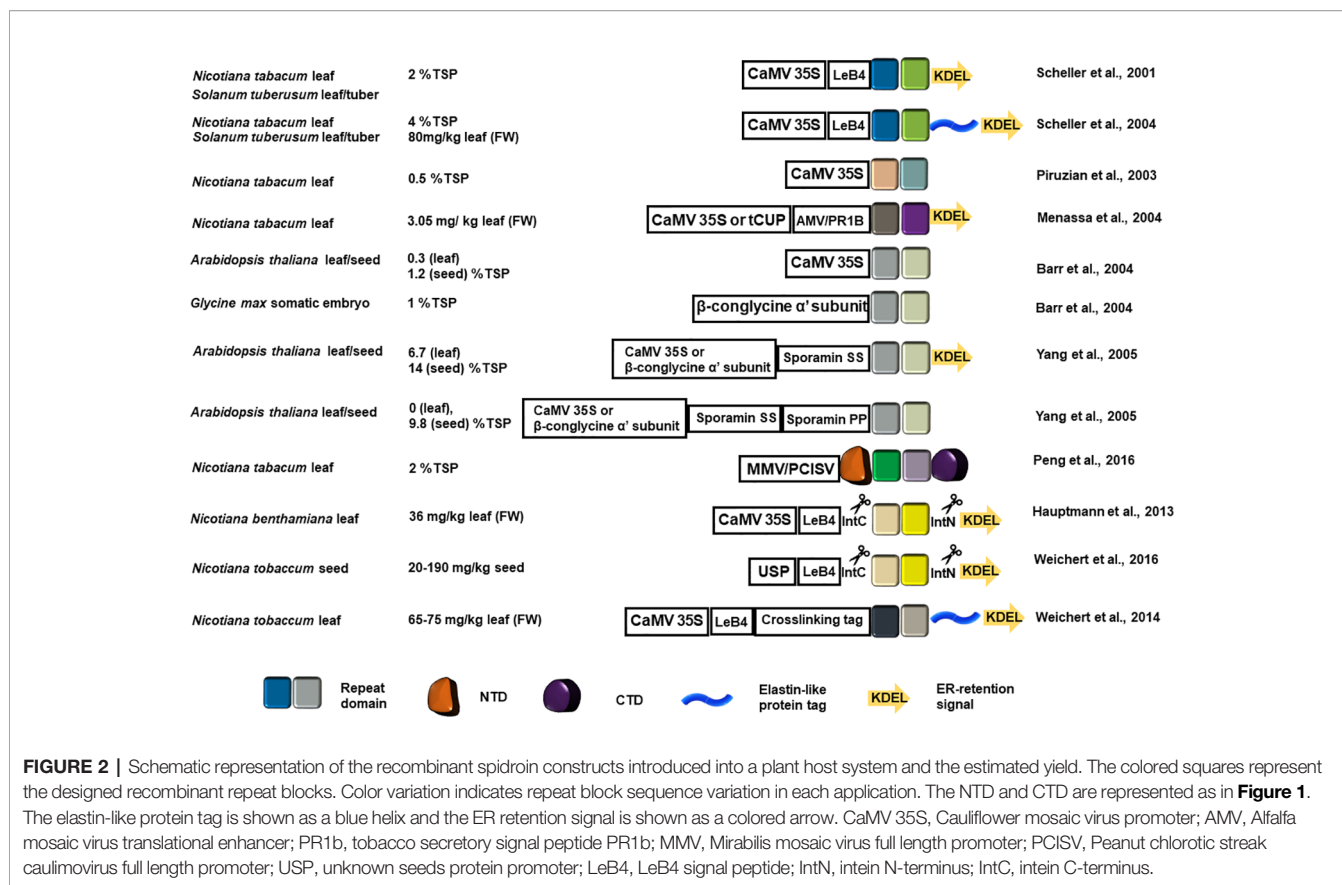
### Spidroin Mimics Design

The sequences of the rMaSp in the aforementioned efforts are derived from MaSp repeat domains. The importance of non-repetitive domains at the N-, and C-terminus in spidroin solubility and fiber assembly has propelled the inclusion of the terminal domains in synthetic constructs (Huemmerich et al., 2004a; Ittah et al., 2006; Hagn et al., 2011). In this context, mimetic spidroins with native N- and C-terminal domains flanking various numbers of consensus block repeat domains of MaSp1 or MaSp2 were produced in tobacco leaves (Peng et al., 2016). For MaSp1, constructs with 8 and 16 copies of repeat domains with flanking terminal domains were detected from tobacco leaf extracts and the yield for rMaSp1R8 (8 copies of repeats) was 0.7% TSP. For MaSp2, 8, 16, and 32 copies of the repeat domains with flanking terminal domains were detected, and the yield of rMaSp2R8 (8 copies of repeats) was 2% TSP. Each rMaSp was represented as an intact full-length protein from the crude leaf extracts despite some autonomous removal of the C-terminal intein tag of rMaSp2. The N- and C-terminal domains substantially increased the solubility of the rMaSp and allowed retention of the concentrated rMaSp in a liquid state even after freeze-drying (Peng et al., 2016).

### Elastin-Like Polypeptide Tags That Facilitate Protein Purification and Protein Body Formation

Fusing the MaSp1 repeat domain mimic protein with 100 copies of VPGXG (X= G, V, or A) elastin-like polypeptide (ELP) tag led





to a high yield (80 mg/kg leaf) when the fusion protein was directed to the ER of tobacco and potato leaves (Scheller et al., 2004). The solubility of ELPs is temperature-dependent and at temperatures below the melting temperature rMaSp1-100xELP fusion protein is insoluble and, therefore, isolable from other soluble proteins through precipitation. Through an inverse transition cycling to increase the temperature, the rMaSp1-100xELP fusion can be reconstituted into the buffer (Meyer and Chilkoti, 1999). ELP also induces protein bodies that sheath the recombinant proteins, leading to increased recombinant protein accumulation (Conley et al., 2009a).

### Post-Translational Modification (PTM) and *In Vitro* Modification

Plants, as eukaryotic hosts, can provide PTM as recombinant proteins mature *in vivo*. However, the PTMs equivalent to those described in silk glands have not been reported from spidroins produced from plants. This may result from either the lack of PTM analysis being performed or the PTM of plant-derived spidroin is undetectable due to low efficiency of the equivalent plant enzymes. With the variation of number of domains designed and the sequence variation within domains, the PTM patterns may be different from each study.

Plant hosts have been shown to process an N-terminal precursor peptide, the legumin type B precursor signal peptide (LeB4), from MaSp1 repeat domain analogs produced in tobacco leaves and potato leaves (Scheller et al., 2001).

Intein-mediated PTM has facilitated repeat domain elongation and produced rMaSps with large numbers of repeat domains, some reaching the size of native spidroins. A flagelliform (FLAG) spidroin flanked by intein self-splicing elements can be directed to the ER in tobacco leaves. After translation, intein self-splicing, and end joining, FLAG proteins of various sizes were produced (Hauptmann et al., 2013). Synthetic FLAG spidroin sequence, flanked by intein splicing elements and directed to the ER in tobacco seeds, produced multimers of up to 450 kDa with an estimated yield of 20–190 mg per kg seed (fresh weight). The recombinant FLAG proteins produced in tobacco seeds are stable at 15°C for one year (Weichert et al., 2016). Another example of *in vitro* PTM is multimerization of rMaSp1-100x ELP fusion protein tagged with lysine or glutamine by transglutaminase, which can produce near native sized spidroin-like recombinant protein (Weichert et al., 2014).

### Plant-Derived Spidroin-Based Biomaterials and Their Applications

Spidroin analogue-based materials are hypoallergenic, non-toxic, non-hemolytic, and minimally induce inflammatory reactions (Gomes et al., 2011; Dams-Kozłowska et al., 2013; Thurber et al., 2015; Hauptmann et al., 2015; Kuhbier et al., 2017). Therefore, their potential use in biomedical application is propitious and emerging. To be applied biomedically, spidroin analogue-based fibers, or materials can be autoclaved in water. The conformation or cytotoxicity of the materials does not change by sterilization

(Lucke et al., 2015). Spidroin analogue-based materials can also be sterilized through ultra-violet radiation (Hafner et al., 2017), although UV may reduce the strength of the fibers (Lai and Goh, 2015). Here, we describe biomaterial formation from plant-derived spidroin mimics and their potentials in regenerative medicine and other applications.

### Regenerative Medicine

Coatings derived from rMaSp1 (repeats)-100x ELP fusion protein produced in tobacco leaves led to increased human chondrocyte cell growth and cellular mass. The overall stimulatory effect was equivalent to that observed with collagen coatings and twice as much as the non-coated control plate (Scheller et al., 2004). Collagen coating also induced long fibroblastoid morphology, which is disfavored in chondrocyte culturing. On rMaSp1-100x ELP coating, however, the chondrocyte cells are round-shaped similar to cells found *in vivo*. Coatings of similar rMaSp1 (repeats)-100x ELP fusion protein produced from tobacco leaves formed polymer films. Coatings of these proteins significantly stimulated murine embryonic fibroblast cells after 24 hours of incubation, an indication of cell proliferation (Hauptmann et al., 2015). In addition, hydrogels made from tobacco-derived rMaSp1R8 (N- and C-termini flanking eight copies of repeats) and rMaSp2R8 (N- and C-termini flanking 8 copies of repeats) promoted human dental pulp stem cell attachment and proliferation (Hafner et al., 2017).

### Diabetes

Recombinant spidroin-like proteins are stable under normal seed storage conditions (Weichert et al., 2016). An economic advantage of transgenic crop seeds producing recombinant protein for medicinal usage is that subjects can be fed directly thereby bypassing cumbersome protein purification. Transgenic rice producing *Araneus ventricosus* MaSp repeats and C-terminus was used to feed diabetic BKS.Cg-m<sup>+/+</sup>Lepr<sup>db</sup> mice. The transgenic rMaSp-producing rice lowered the blood glucose of the diabetic mice. While the mechanism(s) by which *A. ventricosus* rMaSp lowers blood glucose in this animal model is in question, the study reports differences in cellular localization of receptor substrate 1 (IRS1), six-transmembrane protein of prostate 2 (STAMP2), and adenosine monophosphate-activated protein kinase (AMPK) (Park et al., 2019).

## COLLAGEN-BASED SCAFFOLD PROTEINS PRODUCED FROM PLANTS

### Collagen Sequence, Structure, and Folding

Collagen is a primary extracellular matrix scaffold protein of many tissues including connective tissues, basement membrane, skin, vascular tissue, brain, and spinal cord (Miller and Gay, 1982). Collagen biosynthesis is initiated from its precursor protein, procollagen. Variation in sequence, chain length, and chain combination leads to 29 different human collagen molecules (Sorushanova et al., 2019). Post-translational modification of procollagen chain includes lysyl and prolyl hydroxylation by lysyl

hydroxylase (LH) and prolyl 4-hydroxylase (P4H) (Hutton et al., 1967; Kivirikko et al., 1992; Myllyharju, 2003). The collagenous region of the procollagen chain is essentially composed of repetitive motif, Gly-Xaa-Yaa, where Xaa is proline and Yaa is hydroxyproline. The procollagen molecule has flanking non-collagenous N- and C-terminal propeptides, which play critical roles in chain registration and helix formation through a zipper-like propagation starting from the C-propeptide trimerization of disulfide bond (Figure 3; Kirk et al., 1987; Engel and Prockop, 1991).

Three collagen monomer chains assemble into a triple alpha helical fibril, as a homotrimer or a heterotrimer (Brodsky and Ramshaw, 1997). The flanking N- and C- propeptides are cleaved by the procollagen N- and C-proteinases and the newly formed N- C-terminus are denoted as the N- and C-telopeptides. The resulting tropocollagen self-assembles into collagen microfibrils through crosslinking events mediated by lysyl oxidase and transglutaminase or through non-enzymatic glycation (Koide and Nagata, 2005). The collagen microfibrils assemble into collagen fiber with a characteristic D-periodic striation structure (Shoulders and Raines, 2009). Hydroxyproline composition of a non-recombinant human type I collagen homotrimer is 10.8%, and 10.3% for non-recombinant human type I collagen heterotrimer (Nokelainen et al., 2001). Hydroxyproline is critical for triple helix stability (Kotch et al., 2008).

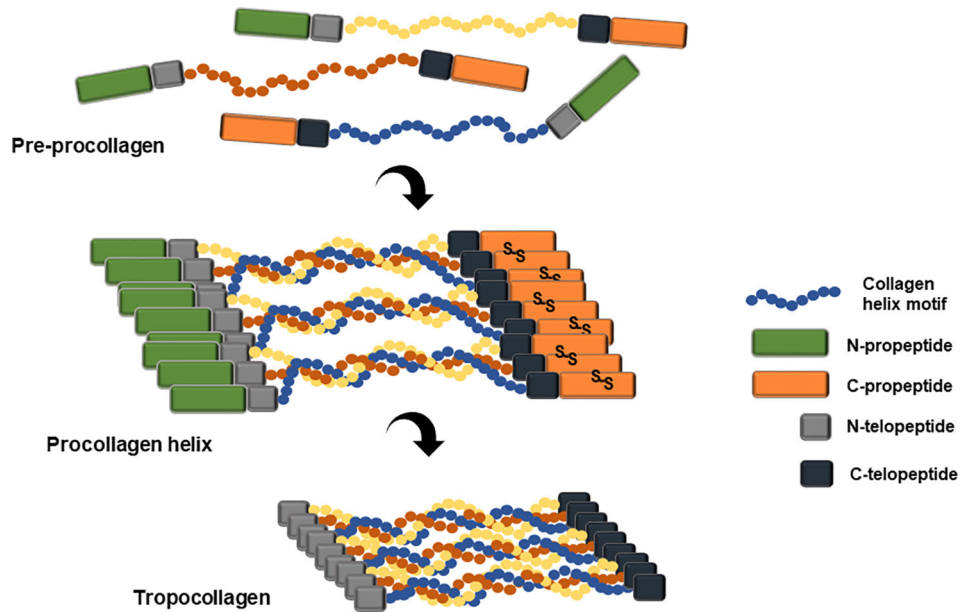
### Plant-Derived Collagen

Although collagen is ubiquitously present in Animalia, the amino acid sequences and the physical properties vary among species (Stover and Verrelli, 2010). Collagen can be obtained from animal tissues, such as bovine, fish, or mouse but collagen from animal sources causes allergic reaction and sometimes may contain pathogen contamination (Mullins et al., 1996).

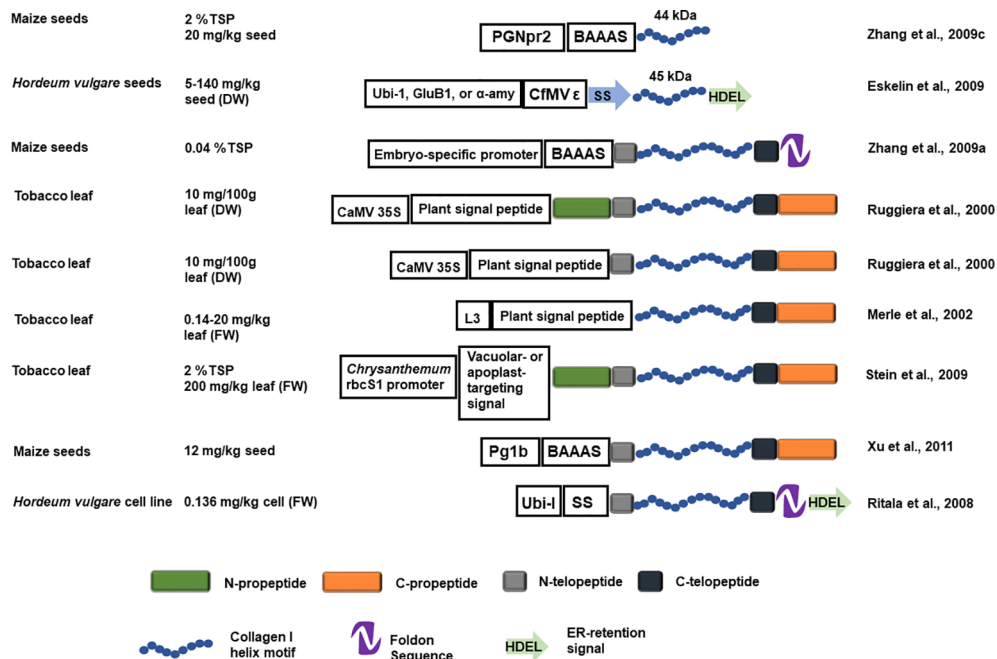
Recombinant collagen has been produced in the yeast *Pichia pastoris* (Nokelainen et al., 2001; Olsen et al., 2005), *Saccharomyces cerevisiae* (Toman et al., 2000), mammalian cells (Geddis and Prockop, 1993; Fukuda et al., 1997; Frischholz et al., 1998), mammals (mouse, John et al., 1999), and silkworm (Tomita et al., 2003) with some success but yield is often limited and/or the recombinant proteins lack post-translational prolyl and lysyl hydroxylation. Although plant hosts produce an indigenous prolyl 4-hydroxylase (P4H), its specificity and affinity to collagen Xaa-Yaa-Gly motif are different than the animal P4H (Tanaka et al., 1981; Hieta and Myllyharju, 2002). For stable triple helix assembly of plant-derived collagen, co-transformation of human P4H is required.

### ER-Targeted Recombinant Collagen Accumulation

A fusion construct of a human procollagen chain helical region, with flanking N- and C-telopeptide, the bacteriophage T4 fibrin foldon (a fusion polypeptide that is expected to facilitate helix assembly), and an ER-targeting signal, collectively called hCl $\alpha$ 1, was transformed into barley. The stably transformed barley cells produced an ~130 kDa unprolylhydroxylated peptide (rhCl $\alpha$ 1) up to 0.136 mg/kg (fresh weight). The rhCl $\alpha$ 1, however, was unstable with a low melting temperature and no PTM phosphorylation was detected (Figure 4; Ritala et al., 2008). ER-targeted full-length



**FIGURE 3 |** Schematic representation of pre-procollagen structure and the formation of tropocollagen. Pre-procollagen is translated and post-translationally modified by various enzymes including the prolyl hydroxylase and lysyl hydroxylase. The mature procollagen initiates triple helix formation through C-propeptide alignment and disulfide bond formation and twists along the helix motifs reach the N-propeptide. Several triple helix chains clusters align with each other and the N- and C- propeptide are cleaved by the N- and C- proteinase, either intracellularly or extracellularly, resulting in newly formed N- and C-termini, named N- and C-telopeptide. This tropocollagen helix stacks and packs into staggered collagen fibrils and alternately forms collagen fibers through multiple crosslinking reactions.



**FIGURE 4 |** Schematic representation of the recombinant collagen constructs introduced into a plant host system and the estimated yield. The blue helix represents the designed recombinant collagen helix block, the N- and C- propeptide is represented by green and orange rectangles. The N- and C- telopeptide is represented with grey and black rectangles. The NTD and CTD are represented as in **Figure 1**. The foldon sequence is represented by a purple hairpin shape and a green arrow represents the ER-retention signal. GluB1, rice glutelin B1 promoter;  $\alpha$ -amy, barley  $\alpha$ -amylase fusion promoter 46 /4-6; BAAAS, Barley  $\alpha$ -amylase signal sequence; Ubi-I, maize ubiquitin promoter and first intron; SS, Arabidopsis basic chitinase signal sequence;  $\epsilon$ , the 5'UTR of Cocksfoot mottle virus; Pg1b, maize embryo specific globulin-1 promoter; PGNpr2, embryo-specific maize globulin-1 promoter; L3, L3 promoter.

rhCI $\alpha$ 1 chain helical region, with flanking N- and C-telopeptides, is barely detectable in transgenic barley seeds (Eskelin et al., 2009). However, a 45 kDa abbreviated pro $\alpha$ 1(I) chain helical region, with flanking N- and C-telopeptides, can be produced in barley seeds. When expressed under the rice glutelin B1 promoter, the 45 kDa rhCI $\alpha$ 1 accumulated to 140 mg/kg in T<sub>1</sub> seeds, whereas the barley  $\alpha$ -amylase fusion promoter 46/4-6 only directed expression of 2 mg/kg in T<sub>1</sub> seeds. Doubled haploid (DH) progeny were generated and the 45-kDa rhCI $\alpha$ 1 production achieved from the best DH lines was 13 mg/kg dry seeds under the ubiquitin promoter and 45 mg/kg dry seeds under the glutelin promoter. Purified 45 kDa rhCI $\alpha$ 1 displayed only low levels of hydroxylated proline (2.8%, Eskelin et al., 2009).

Full-length pro $\alpha$ 1(I) chain, with the flanking N- and C-telopeptides and foldon, has also been produced in transgenic corn seeds after *Agrobacterium*-mediated transformation into corn immature embryos. The rhCI $\alpha$ 1 yield is estimated to be 0.04% of TSP, with a low prolylhydroxylation rate (1.23%) and low melting temperature (26–27°C). A portion of the rhCI $\alpha$ 1 is able to form triple alpha helices. Sequencing revealed that the N-terminal telopeptide had been removed but it remains unknown whether the removal occurred during *in vivo* protein maturation or during the protein purification steps (Zhang et al., 2009a). Protein purification optimization using this low yield rhCI $\alpha$ 1-producing corn seeds was assessed after spiking with *Pichia*-produced pro $\alpha$ 1(I) chain with or without foldon. A combination of high purity without too much compromise on yield is achieved when the foldon is cleaved after the fusion protein is precipitated from the crude extract (Setina et al., 2015). While there is no difference in the rhCI $\alpha$ 1 recovered from dry milling or wet milling of the transgenic corn seeds, dry milling may be preferred for practical reasons at the farm (Zhang et al., 2009b).

### Collagen Mimics Design

Recombinant human procollagen pro $\alpha$ 1(I) chain (~120 kDa) with the flanking N- and C-terminal telopeptides plus propeptides, or the C-terminal telopeptide plus propeptide alone have been stably transformed into tobacco. Absence of the N-terminal propeptide led to production of truncated rhCI $\alpha$ 1 protein of various lengths indicating it is required to maintain the integrity of full-length rhCI $\alpha$ 1 during translation and protein extraction. Interestingly, the N-terminal propeptide of rhCI $\alpha$ 1 was intact in plantlet extracts but removed from rhCI $\alpha$ 1 found in mature tobacco leaf extracts. The C-terminal propeptide forms a disulfide bond in the precursor protein but C-terminal propeptides are cleaved both in plantlets and mature plants. The rhCI $\alpha$ 1 is minimally prolylhydroxylated and unstable. However, alpha helices are formed and can be detected through circular dichroism spectra. Triple helix chain assembly may be facilitated by protein disulfide isomerase (Ruggiero et al., 2000). The above tobacco leaf-derived unhydroxylated collagen I (without N-terminal telopeptides) was used to determine the function of hydroxyproline in collagen folding and fibril formation. Hydroxyproline stabilizes the triple helix chain and assists a faster reassembly after denaturation. Unprolylhydroxylated collagen adopts more flexible conformation as compared to prolylhydroxylated

collagen under the same temperature. Unprolylhydroxylated collagen fails to form the D-periodic striation in the fibrils at physiological ionic strength conditions but can form the D-periodic striation at lower ionic strength (Perret et al., 2001).

### Post-Translational Modification (PTM) and *In Vitro* Modification

Human procollagen pro $\alpha$ 1(I) chain (with N- and C- telopeptides but without N- propeptide) has been produced in tobacco leaves through transient or stable transformation along with a chimeric proline-4-hydroxylase (P4H) consisting of a *Caenorhabditis elegans*  $\alpha$  subunit and mouse P4H  $\beta$  subunit. Stable co-expression in this system yielded approximately 0.14–20 mg rhCI $\alpha$ 1 per kg of leaf and 8.41% prolylhydroxylation. Transient co-expression led to 0.5–1 mg rhCI $\alpha$ 1 per kg of leaf material and 6.84% prolylhydroxylation. The helical structure was assembled and the melting temperature (37°C) agreed with the level of prolylhydroxylation (Merle et al., 2002).

Transgenic plants containing human procollagen pro $\alpha$ 1(I) and pro $\alpha$ 2(I) chains constructs targeted to the vacuole, apoplast, or untargeted were obtained through co-transformation into tobacco. Only plants directing vacuole-targeted procollagen expression produced suitable protein products. An additional transgenic line created by co-transformation of human p4H  $\alpha$  subunit, human p4H  $\beta$  subunit, and human lysyl hydroxylase 3 (LH3), that were also targeted to tobacco vacuole, was crossed with the procollagen-expressing line. The resulting heteromeric recombinant collagen-expressing line displayed PTM prolyl hydroxylation and lysyl hydroxylation, which was comparable to native human collagen. The harvested leaves yielded ~2% of TSP of rhCI $\alpha$ 1, stable triple helical fibrils resistant to pepsin digestion were formed, and the characteristic D-periodic striation structures were observed (Stein et al., 2009).

Human procollagen pro $\alpha$ 1(I) coding sequence with or without the P4H  $\alpha$  subunit and P4H  $\beta$  subunit fusion construct was transformed into corn. The production of the P4H  $\alpha$  and  $\beta$  subunit caused the reduction of overall full-length rhCI $\alpha$ 1 yield but resulted in 18.11% of hydroxyproline modification in recombinant collagen, which formed stable triple helices (Xu et al., 2011). Despite prolyl hydroxylation and lysyl hydroxylation, no phosphorylation was detected (Merle et al., 2002) and no glycosylation was detected of a 44 kDa rhCI $\alpha$ 1 fragment from corn seeds (Zhang et al., 2009c).

### Plant-Derived Collagen-Based Biomaterials and Their Applications

Collagen from an animal source such as bovine collagen transplant is frequently associated with dermatomyositis (Cukier et al., 1993). Recombinant collagens are alternatives to be used in wound healing (Davison-Kotler et al., 2019), drug delivery (Olsen et al., 2003), and regenerative medicine (Werkmeister and Ramshaw, 2012). The rhCI $\alpha$ 1 obtained from the patented production from transgenic tobacco formed stable triple helical fibrils and supported peripheral blood mononuclear cells (PBMNCs) proliferation, comparable to human non-recombinant collagen type I (Stein et al., 2009). The same tobacco-derived collagens promote the attachment



and the multiplication of endothelial, fibroblast, and keratinocyte cells (Willard et al., 2013). The fibroblast infiltration, epidermal differentiation, and cellular metabolism of engineered skin based on tobacco-derived collagen are equivalent to engineered skin supported by human and animal collagen source. However, the interleukin 1 beta (IL-1 $\beta$ ) accumulation of the macrophage cells (THP-1) is significantly less evident on engineered skin containing tobacco-derived collagen, indicating the hypoimmunogenicity of the plant-derived collagen matrix (Willard et al., 2013).

## ELASTIN-BASED SCAFFOLD PROTEIN PRODUCED FROM PLANTS

Elastin is another fundamental extracellular matrix protein component that is found in connective tissues, vascular tissues, and basement membrane (Halper and Kjaer, 2014). As it is named, elastin is a protein elastomer with superior mechanical strain and has frequent demand in tissue regeneration practices (Urry et al., 2002; Daamen et al., 2007). This has sparked considerable interest in recombinant production of the major precursor protein tropoelastin. Tropoelastin is a secreted extracellular matrix protein with predicted alternating hydrophilic and hydrophobic regions (Wise and Weiss, 2009). The conserved C-terminal motif GRKRK binds to integrin  $\alpha_v\beta_3$ .

The tandem repeats of elastin-like peptides VPGXG (X= G, V, or A) have been used as a fusion tag to assist recombinant protein purification and protein body induction in applications in recombinant spider silk protein production (Scheller et al., 2004; Conley et al., 2009b). Recombinant production of full-length tropoelastin has been just recently explored. A 2175 bp-ELN *orf* of tropoelastin was synthesized and transformed into tobacco plants through transient infiltration. A band corresponding to ~70 kDa was detected during electrophoresis of fresh leaf extracts using an antibody against tropoelastin. However, recombinant tropoelastin degraded into two smaller molecular weight peptides when leaf extracts were frozen (Abdelghani et al., 2015). As the instability of protein going through freeze thaw of the leaf tissues will create a problem for potential harvesting and storage, recombinant tropoelastin directed to seeds may be more suitable, since recombinant protein can be stably stored in the seeds (Weichert et al., 2016).

## OSTEOPONTIN-BASED PROTEIN PRODUCED FROM PLANTS

Osteopontin is a negatively charged glycoprotein that is secreted to extracellular matrices from bone cells and mesenchymal stem cells (Reinholt et al., 1990; Scatena et al., 2007). The myriad of functions of osteopontin in the extracellular matrices include integrin binding through the RGD site (Yokosaki et al., 2005), binding to fibronectin (Mukherjee et al., 1995) and collagen (Chen et al., 1992), inducing cytokine and chronic immune response (Lund et al., 2009), and regulatory roles in diabetes and obesity (Kahles et al., 2014). A human osteopontin (OPN)

construct driven by the CaMV 35S promoter was transformed into *Nicotiana benthamiana* leaves and allowed plant-based production of a ~50 kDa osteopontin recognized by anti-osteopontin antibody. Recombinant OPN accumulated up to ~100  $\mu$ g rhOPN per kg leaf mass. Tobacco-derived OPN promoted proliferation of human periodontal ligament stem cells and transcriptional increase of osteogenic differentiation and bone mineralization related genes, OSX, DMP1, and Wnt3, indicating that the plant-derived product had retained its biological activity (Rattanapit et al., 2017).

## EMERGING TOOLS FOR RECOMBINANT SCAFFOLD PROTEIN PRODUCTION IN PLANTS

The considerable success of laboratory scale production of scaffold proteins using plants as a host system has been possible thanks to advances in molecular biology tools. However, any individual approach may not produce the same result for each extraction strategy due to the specific characteristics of the recombinant proteins being produced, diversity among plant hosts, and variation in host tissue types. For example, directing recombinant proteins into the plant vacuole can be a useful tool to store recombinant protein secluded from the cytoplasmic enzymes (Marin Viegas et al., 2016). While this method increases recombinant collagen yield (Stein et al., 2009), it results in no detectable yield of recombinant spider silk proteins (Yang et al., 2005). Therefore, there is no universal tool that fits all recombinant protein types, host systems, and host tissues. The selection is mostly empirical, and the outcomes are largely unpredictable.

### Protein Body Tag

Fusion tags like the Zein protein tag Zera<sup>®</sup>, (PPPVHL)<sub>8</sub> (Llop-Tous et al., 2010), hydrophobin-I (Reuter et al., 2016), and the elastin-like polypeptide (ELP), VPGXG (X= G, V, or A), can induce the formation of protein body-like organelles (PBs) in transgenic plants. In these PBs, membrane surrounds the protein body protecting the recombinant protein from proteolytic degradation. Thus, protein body tag constitutes a promising tool to enhance recombinant protein accumulation in plant tissues. The ability to purify recombinant ELP fusion through inverse transition cycling (ITC) ameliorates the expensive chromatography steps (Meyer and Chilkoti, 1999). A total of 30 VPGXG repeats is sufficient for effective protein recovery through ITC (Conley et al., 2009b). ELP induces large protein bodies with ER-derived membrane secluding the recombinant protein, which remains within the ER instead of being permanently directed to the vacuolar storage. Interestingly, the fusion tags are not a strict requirement for PB formation and when the ER-directed recombinant protein concentration reaches 0.2% TSP, PBs can be formed (Saberianfar et al., 2015). The ability to induce PB formation has been used to increase expression of desirable recombinant proteins such as erythropoietin (EPO) and the human cytokine interleukin-10 (hIL-10) as these co-expressed recombinant proteins are passively sequestered to the interior of the PBs.

## GAANTRY System

Recombinant production of scaffold proteins is hindered by a lack of tools to assemble large numbers of tandem repeats and by difficulties in co-transformation of PTM enzymes to achieve the optimal mechanical property and physiological function of the product. The recently developed GAANTRY system (Collier et al., 2018) uses recombinase-facilitated excision of plasmid backbone sequence and allows gene stacking of up to 10 different modular sequences from different donor plasmids. The modularity can be different domains of the same gene, or enzymes with catalytic functions in the same pathway. To produce tobacco-derived recombinant collagens with the appropriate PTM, five different cloning processes, two co-transformations, and a subsequent breeding were needed (Stein et al., 2009). The GAANTRY system may minimize the number of transformations, the number of antibiotics needed for the selection, and could reduce the subsequent plant breeding steps. The cloning strategies recently developed for long repeat sequences (Riet et al., 2017) and the recursive directional ligation approach (Dinjaski et al., 2018) may be used to create donor plasmids for the GAANTRY system. Combinations of different sequence modules from the donor plasmid will essentially allow an array of molecules, each with a unique combination of domains.

## POTENTIAL APPLICATION OF SCAFFOLD PROTEINS FOR DRUG DELIVERY AND DRUG DISCOVERY

Recombinant scaffold proteins produced from plant hosts promoted cell attachment and proliferation (Stein et al., 2009; Hauptmann et al., 2015). With the known affinity of spider silk analogues for liposome binding and encapsulation of low molecular weight compounds (Antonenko et al., 2010; Hardy et al., 2013; Agostini et al., 2015), plant derived spider silk proteins may also be used for controllable intake and release of therapeutic molecules (Doblhofer and Scheibel, 2015). The special affinity of spider silk mimics to HER2 expressing cells (Florczak et al., 2014) and the tumor-homing peptides (Numata et al., 2011) provides opportunities for the spider silk-based materials to be used in targeted melanoma cell delivery.

In addition to providing a supporting scaffold, extracellular matrix proteins (fibronectin, laminin, and collagen) serve as protection barriers and mediators of signal transduction (Singh et al., 2012). Exposed or partially degraded ECM proteins are known for attracting pathogens and providing scaffold for pathogen proliferation (Steukers et al., 2012). Agglutinin-like adhesin from common fungal pathogen *Candida albicans* specifically binds to fibronectin, laminin, and collagen IV (Gaur and Klotz, 1997). The direct connection of fungal pathogenicity with extracellular matrix protein interaction is postulated based on pathogenic fungi *Paracoccidioides brasiliensis*; when cells were pre-coated with ECM protein laminin, the infection consisting of granulomas became more severe (Vicentini et al., 1994). Although extracellular matrix proteins have gained attention in drug discovery, no antifungal agents are designed targeting the

interaction of fungal pathogen with host extracellular matrix proteins (Järveläinen et al., 2009). Currently, no known fabricated extracellular matrices have assisted the screen of antifungals. Fabrication of extracellular matrices from plant-derived collagen, tropoelastin, and osteopontin is possible, and have a potential in future antifungal discovery.

## CONCLUDING REMARKS

### Economical Feasibility

Despite the success in proof of concept demonstration of a variety of scaffold proteins produced in plants, the commercialization of plant-derived scaffold proteins has been limited (Schillberg et al., 2019). The feasibility and the success of utilizing biotechnology tools vary depending on specific plant type and plant tissues and is rather unpredictable. Transient infiltration production flow can be improved to reduce the cost using ready-to-use cryo-stock of *Agrobacterium tumefaciens* (Spiegel et al., 2019). The yield based on current methods ranges from microgram to about 200 milligram of recombinant protein per kilogram of leaf or seed tissues. The pilot scale purification of MaSp1-ELP fusion proteins from stable transformed tobacco leaves using heat and acetone precipitation followed by centrifugal inverse transition cycling, achieved 80 mg MaSp1-ELP per kg of leaves (Heppner et al., 2016). This yield still does not meet the commercially acceptable level (1–5 g recombinant protein/kg of plant tissue) (Zhang et al., 2009c). Revolutionary recombinant protein production methods and strategies are essential to allow industrial scale production. Other than prolyl hydroxylation detection of recombinant collagen, other PTMs and the regulation of the plant-derived scaffold proteins have not been explored.

### Biomaterial Assembly and Application

Biomaterials assembled from recombinant fibrous protein or extracellular matrix proteins have positively enhanced biomedical applications such as boosting pancreatic islets survival and promote human fibroblast and human dermal microvascular endothelial cell adherence and multiplication (Annabi et al., 2013; Johansson et al., 2015; Widhe et al., 2016; Pereira et al., 2017; Tasiopoulos et al., 2018). Fusion of scaffold proteins with antimicrobial peptides can inhibit bacterial infection (Gomes et al., 2011). Recent *N. clavipes* genome assembly revealed novel spidroin motifs, which can be engineered potentially and add to the complexity of the plethora of recombinant spidroin molecules. Recombinant collagen production can also be expanded to the other types of collagen, especially collagen type IV, which is the major component of basement membrane. The combination of spider silk analogues with the motifs from the extracellular matrix proteins collagen, fibronectin, and laminin may provide unlimited opportunities for functionalized biomaterials. Crosslinking methods such as click chemistry (Harvey et al., 2017) and photocrosslinking (Johansson et al., 2015) are also applicable to plant-derived scaffold proteins. These scaffold proteins have potentials to be developed into bioink for precise

and consistent biomaterial fabrication (Xiao et al., 2011; DeSimone et al., 2017; Chawla et al., 2018).

## AUTHOR CONTRIBUTIONS

CP and WM wrote the manuscript. LK provided helpful suggestions, provided critical comments, and helped edit the manuscript.

## REFERENCES

- Abdelghani, M., El-Heba, G. A. A., Abdelhadi, A. A., and Abdallah, N. A. (2015). Expression of synthetic human tropoelastin (hTE) protein in *Nicotiana tabacum*. *GM Crops Food* 6 (1), 54–62. doi: 10.1080/21645698.2015.1026524
- Agostini, E., Winter, G., and Engert, J. (2015). Water-based preparation of spider silk films as drug delivery matrices. *J. Control Release* 213, 134–141. doi: 10.1016/j.jconrel.2015.06.025
- Altman, G. H., Diaz, F., Jakuba, C., Calabro, T., Horan, R. L., Chen, J., et al. (2003). Silk-based biomaterials. *Biomaterials* 24 (3), 401–416. doi: 10.1016/S0142-9612(02)00353-8
- Andersson, M., Holm, L., Ridderstrale, Y., Johansson, J., and Rising, A. (2013). Morphology and composition of the spider major ampullate gland and dragline silk. *Biomacromolecules* 14 (8), 2945–2952. doi: 10.1021/bm400898t
- Andersson, M., Chen, G., Otkovs, M., Landreh, M., Nordling, K., Kronqvist, N., et al. (2014). Carbonic anhydrase generates CO<sub>2</sub> and H<sup>+</sup> that drive spider silk formation via opposite effects on the terminal domains. *PLoS Biol.* 12 (8), e1001921–e1001921. doi: 10.1371/journal.pbio.1001921
- Andersson, M., Johansson, J., and Rising, A. (2016). Silk Spinning in Silkworms and Spiders. *Int. J. Mol. Sci.* 17 (8), 1290. doi: 10.3390/ijms17081290
- Annabi, N., Mithieux, S. M., Zorlutuna, P., Camci-Unal, G., Weiss, A. S., and Khademhosseini, A. (2013). Engineered cell-laden human protein-based elastomer. *Biomaterials* 34 (22), 5496–5505. doi: 10.1016/j.biomaterials.2013.03.076
- Antonenko, Y. N., Perevoshchikova, I. V., Davydova, L. I., Agapov, I. A., and Bogush, V. G. (2010). Interaction of recombinant analogs of spider silk proteins 1F9 and 2E12 with phospholipid membranes. *Biochim. Biophys. Acta* 1798 (6), 1172–1178. doi: 10.1016/j.bbame.2010.03.006
- Arcidiacono, S., Mello, C., Kaplan, D., Cheley, S., and Bayley, H. (1998). Purification and characterization of recombinant spider silk expressed in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 49 (1), 31–38. doi: 10.1007/s002530051133
- Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A., et al. (2010). Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* 465 (7295), 236–238. doi: 10.1038/nature08962
- Atkinson, J. H., Parnham, S., Marcotte, W. R. J., and Olsen, S. K. (2016). Crystal structure of the nephila clavipes major ampullate spidroin 1A N-terminal domain reveals plasticity at the dimer interface. *J. Biol. Chem.* 291 (36), 19006–19017. doi: 10.1074/jbc.M116.736710
- Babb, P. L., Lahens, N. F., Correa-Garhwal, S. M., Nicholson, D. N., Kim, E. J., Hogenesch, J. B., et al. (2017). The Nephila clavipes genome highlights the diversity of spider silk genes and their complex expression. *Nat. Genet.* 49 (6), 895–903. doi: 10.1038/ng.3852
- Barr, L., Fahnestock, S., and Yang, J. (2004). Production and purification of recombinant DP1B silk-like protein in plants. *Mol. Breed.* 13 (4), 345–356. doi: 10.1023/B:MOLB.0000034089.92263.ff
- Barroso da Silva, F. L., Pasquali, S., Derreumaux, P., and Dias, L. G. (2016). Electrostatics analysis of the mutational and pH effects of the N-terminal domain self-association of the major ampullate spidroin. *Soft Matter* 12 (25), 5600–5612. doi: 10.1039/c6sm00860g
- Bauer, J., and Scheibel, T. (2017). Conformational stability and interplay of helical N- and C-Terminal domains with implications on major ampullate spidroin assembly. *Biomacromolecules* 18 (3), 835–845. doi: 10.1021/acs.biomac.6b01713
- Beckwith, R., and Arcidiacono, S. (1994). Sequence conservation in the C-terminal region of spider silk proteins (Spidroin) from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentenarius* (Araneidae). *J. Biol. Chem.* 269 (9), 6661–6663.

## FUNDING

This work is supported by the Clemson University Libraries Open Access Publishing Fund. LK and CP are supported by the NIH grant # 1P20GM109094-01A1. This project was also funded in part by Clemson University's R-Initiative Program to WM.

- Bini, E., Knight, D. P., and Kaplan, D. L. (2004). Mapping domain structures in silks from insects and spiders related to protein assembly. *J. Mol. Biol.* 335 (1), 27–40. doi: 10.1016/j.jmb.2003.10.043
- Bini, E., Foo, C. W., Huang, J., Karageorgiou, V., Kitchel, B., and Kaplan, D. L. (2006). RGD-functionalized bioengineered spider dragline silk biomaterial. *Biomacromolecules* 7 (11), 3139–3145. doi: 10.1021/bm0607877
- Bourzac, K. (2015). Spiders: web of intrigue. *Nature* 519 (7544), S4–S6. doi: 10.1038/519S4a
- Brodsky, B., and Ramshaw, J. A. (1997). The collagen triple-helix structure. *Matrix Biology: J. Int. Soc. Matrix Biol.* 15 (8–9), 545–554. doi: 10.1016/s0945-053x(97)90030-5
- Candelas, G. C., Arroyo, G., Carrasco, C., and Dompencil, R. (1990). Spider silk glands contain a tissue-specific alanine tRNA that accumulates *in vitro* in response to the stimulus for silk protein synthesis. *Dev. Biol.* 140 (1), 215–220. doi: 10.1016/0012-1606(90)90069-U
- Chawla, S., Midha, S., Sharma, A., and Ghosh, S. (2018). Silk-Based bioinks for 3D bioprinting. *Adv. Healthc. Mat.* 7 (8), e1701204. doi: 10.1002/adhm.201701204
- Chen, Y., Bal, B. S., and Gorski, J. P. (1992). Calcium and collagen binding properties of osteopontin, bone sialoprotein, and bone acidic glycoprotein-75 from bone. *J. Biol. Chem.* 267 (34), 24871–24878.
- Collier, R., Thomson, J. G., and Thilmony, R. (2018). A versatile and robust Agrobacterium-based gene stacking system generates high-quality transgenic Arabidopsis plants. *Plant J. Cell Mol. Biol.* 95 (4), 573–583. doi: 10.1111/tpj.13992
- Collin, M. A., Clarke, T. H.3rd, Ayoub, N. A., and Hayashi, C. Y. (2018). Genomic perspectives of spider silk genes through target capture sequencing: conservation of stabilization mechanisms and homology-based structural models of spidroin terminal regions. *Int. J. Biol. Macromol.* 113, 829–840. doi: 10.1016/j.ijbiomac.2018.02.032
- Conley, A. J., Joensuu, J. J., Menassa, R., and Brandle, J. E. (2009a). Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol.* 7, 48. doi: 10.1186/1741-7007-7-48
- Conley, A. J., Joensuu, J. J., Jevnikar, A. M., Menassa, R., and Brandle, J. E. (2009b). Optimization of elastin-like polypeptide fusions for expression and purification of recombinant proteins in plants. *Biotechnol. Bioeng.* 103 (3), 562–573. doi: 10.1002/bit.22278
- Craig, H. C., Blamires, S. J., Sani, M.-A., Kasumovic, M. M., Rawal, A., and Hook, J. M. (2019). DNP NMR spectroscopy reveals new structures, residues and interactions in wild spider silks. *Chem. Commun. (Cambridge England)* 55 (32), 4687–4690. doi: 10.1039/c9cc01045a
- Cukier, J., Beauchamp, R. A., Spindler, J. S., Spindler, S., Lorenzo, C., and Trentham, D. E. (1993). Association between bovine collagen dermal implants and a dermatomyositis or a polymyositis-like syndrome. *Ann. Internal Med.* 118 (12), 920–928. doi: 10.7326/0003-4819-118-12-199306150-00002
- Daamen, W. F., Veerkamp, J. H., van Hest, J. C. M., and van Kuppevelt, T. H. (2007). Elastin as a biomaterial for tissue engineering. *Biomaterials* 28 (30), 4378–4398. doi: 10.1016/j.biomaterials.2007.06.025
- Dams-Kozłowska, H., Majer, A., Tomasiwicz, P., Lozinska, J., Kaplan, D. L., and Mackiewicz, A. (2013). Purification and cytotoxicity of tag-free bioengineered spider silk proteins. *J. Biomed. Mater. Res. A* 101 (2), 456–464. doi: 10.1002/jbm.a.34353
- Davison-Kotler, E., Marshall, W. S., and Garcia-Gareta, E. (2019). Sources of collagen for biomaterials in skin wound healing. *Bioeng. (Basel Switzerland)* 6 (3), 56. doi: 10.3390/bioengineering6030056
- DeSimone, E., Schacht, K., Pellert, A., and Scheibel, T. (2017). Recombinant spider silk-based bioinks. *Biofabrication* 9 (4), 44104. doi: 10.1088/1758-5090/aa90db



- Dicko, C., Kenney, J. M., and Vollrath, F. (2006). Beta-silks: enhancing and controlling aggregation. *Adv. Protein Chem.* 73, 17–53. doi: 10.1016/S0065-3233(06)73002-9
- Dinjaski, N., Huang, W., and Kaplan, D. L. (2018). Recursive Directional Ligation Approach for Cloning Recombinant Spider Silks. *Methods Mol. Biol.* 1777, 181–192. doi: 10.1007/978-1-4939-7811-3\_10
- Doblhofer, E., and Scheibel, T. (2015). Engineering of recombinant spider silk proteins allows defined uptake and release of substances. *J. Pharm. Sci.* 104 (3), 988–994. doi: 10.1002/jps.24300
- Edlund, A. M., Jones, J., Lewis, R., and Quinn, J. C. (2018). Economic feasibility and environmental impact of synthetic spider silk production from *Escherichia coli*. *New Biotechnol.* 42, 12–18. doi: 10.1016/j.nbt.2017.12.006
- Eisoldt, L., Hardy, J. G., Heim, M., and Scheibel, T. R. (2010). The role of salt and shear on the storage and assembly of spider silk proteins. *J. Struct. Biol.* 170 (2), 413–419. doi: 10.1016/j.jsb.2009.12.027
- Ene, R., Papadopoulos, P., and Kremer, F. (2009). Combined structural model of spider dragline silk. *Soft Matter* 5 (22), 4568–4574. doi: 10.1039/B911159J
- Engel, J., and Prockop, D. J. (1991). The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper. *Annu. Rev. Biophys. Biophys. Chem.* 20 (1), 137–152. doi: 10.1146/annurev.bb.20.060191.001033
- Eskelin, K., Ritala, A., Suntio, T., Blumer, S., Holkeri, H., Wahlstrom, E. H., et al. (2009). Production of a recombinant full-length collagen type I alpha-1 and of a 45-kDa collagen type I alpha-1 fragment in barley seeds. *Plant Biotechnol. J.* 7, 657–672. doi: 10.1111/j.1467-7652.2009.00432.x
- Fahnestock, S. R., and Bedzyk, L. A. (1997). Production of synthetic spider dragline silk protein in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* 47 (1), 33–39. doi: 10.1007/s002530050884
- Florczak, A., Mackiewicz, A., and Dams-Kozłowska, H. (2014). Functionalized spider silk spheres as drug carriers for targeted cancer therapy. *Biomacromolecules* 15 (8), 2971–2981. doi: 10.1021/bm500591p
- Frischholz, S., Beier, F., Girkontaite, I., Wagner, K., Poschl, E., Turnay, J., et al. (1998). Characterization of human type X procollagen and its NC-1 domain expressed as recombinant proteins in HEK293 cells. *J. Biol. Chem.* 273 (8), 4547–4555. doi: 10.1074/jbc.273.8.4547
- Fukuda, K., Hori, H., Utani, A., Burbelo, P. D., and Yamada, Y. (1997). Formation of recombinant triple-helical [alpha 1(IV)]2 alpha 2(IV) collagen molecules in CHO cells. *Biochem. Biophys. Res. Commun.* 231 (1), 178–182. doi: 10.1006/bbrc.1997.6069
- Fukushima, Y. (1998). Genetically engineered syntheses of tandem repetitive polypeptides consisting of glycine-rich sequence of spider dragline silk. *Biopolymers* 45 (4), 269–279. doi: 10.1002/(SICI)1097-0282(19980405)45:4<269::AID-BIP1>3.0.CO;2-J
- Gaines, W. A., and Marcotte, W. R. Jr. (2011). Recombinant Dragline Silk-Like Proteins-Expression and Purification. *AATCC Rev.* 11 (2), 75–79.
- Gaines, W. A., Sehorn, M. G., and Marcotte, W. R. Jr. (2010). Spidroin N-terminal domain promotes a pH-dependent association of silk proteins during self-assembly. *J. Biol. Chem.* 285 (52), 40745–40753. doi: 10.1074/jbc.M110.163121
- Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., and Lewis, R. (2001). Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science* 291 (5513), 2603–2605. doi: 10.1126/science.1057561291/5513/2603
- Gaur, N. K., and Klotz, S. A. (1997). Expression, cloning, and characterization of a *Candida albicans* gene, ALA1, that confers adherence properties upon *Saccharomyces cerevisiae* for extracellular matrix proteins. *Infect. Immun.* 65 (12), 5289–5294. doi: 10.1128/IAI.65.12.5289-5294.1997
- Geddis, A. E., and Prockop, D. J. (1993). Expression of human COL1A1 gene in stably transfected HT1080 cells: the production of a thermostable homotrimer of type I collagen in a recombinant system. *Matrix* 13 (5), 399–405. doi: 10.1016/S0934-8832(11)80045-4
- Giesa, T., Perry, C. C., and Buehler, M. J. (2016). Secondary Structure Transition and Critical Stress for a Model of Spider Silk Assembly. *Biomacromolecules* 17 (2), 427–436. doi: 10.1021/acs.biomac.5b01246
- Gomes, S., Leonor, I. B., Mano, J. F., Reis, R. L., and Kaplan, D. L. (2011). Spider silk-bone sialoprotein fusion proteins for bone tissue engineering. *Soft Matter* 7 (10), 4964–4973. doi: 10.1039/C1SM05024A
- Gosline, J. M., Guerette, P. A., Ortlepp, C. S., and Savage, K. N. (1999). The mechanical design of spider silks: from fibroin sequence to mechanical function. *J. Exp. Biol.* 202 (Pt 23), 3295–3303.
- Gould, J. (2015). Learning from nature's best. *Nature* 519 (7544), S2–S3. doi: 10.1038/519S2a
- Guehrs, K. H., Schlott, B., Grosse, F., and Weisshart, K. (2008). Environmental conditions impinge on dragline silk protein composition. *Insect Mol. Biol.* 17 (5), 553–564. doi: 10.1111/j.1365-2583.2008.00826.x
- Guerette, P. A., Ginzinger, D. G., Weber, B. H., and Gosline, J. M. (1996). Silk properties determined by gland-specific expression of a spider fibroin gene family. *Science* 272 (5258), 112–115. doi: 10.1126/science.272.5258.112
- Hafner, K., Montag, D., Maeser, H., Peng, C., Marcotte, W. R. J., Dean, D., et al. (2017). Evaluating adhesion and alignment of dental pulp stem cells to a spider silk substrate for tissue engineering applications. *Mat. Sci. Eng. C. Mat. Biol. Appl.* 81, 104–112. doi: 10.1016/j.msec.2017.07.019
- Hagn, F., Eisoldt, L., Hardy, J. G., Vendrely, C., Coles, M., Scheibel, T., et al. (2010). A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* 465 (7295), 239–242. doi: 10.1038/nature08936
- Hagn, F., Thamm, C., Scheibel, T., and Kessler, H. (2011). pH-dependent dimerization and salt-dependent stabilization of the N-terminal domain of spider dragline silk—implications for fiber formation. *Angew Chem. Int. Ed Engl.* 50 (1), 310–313. doi: 10.1002/anie.201003795
- Halper, J., and Kjaer, M. (2014). “Basic Components of Connective Tissues and Extracellular Matrix: Elastin, Fibrillin, Fibulins, Fibrinogen, Fibronectin, Laminin, Tenascins and Thrombospondins,” in *Progress in Heritable Soft Connective Tissue Diseases*. (Springer Dordrecht) Ed. J. Halper, 31–47. doi: 10.1007/978-94-007-7893-1\_3
- Hardy, J. G., Leal-Egana, A., and Scheibel, T. R. (2013). Engineered spider silk protein-based composites for drug delivery. *Macromol. Biosci.* 13 (10), 1431–1437. doi: 10.1002/mabi.201300233
- Harvey, D., Bardelang, P., Goodacre, S. L., Cockayne, A., and Thomas, N. R. (2017). Antibiotic spider silk: site-specific functionalization of recombinant spider silk using “click” chemistry. *Adv. Mat.* 29 (10), 1604245. doi: 10.1002/adma.201604245
- Hauptmann, V., Weichert, N., Menzel, M., Knoch, D., Paegle, N., Scheller, J., et al. (2013). Native-sized spider silk proteins synthesized in planta via intein-based multimerization. *Transgenic Res.* 22 (2), 369–377. doi: 10.1007/s11248-012-9655-6
- Hauptmann, V., Menzel, M., Weichert, N., Reimers, K., Spohn, U., and Conrad, U. (2015). In planta production of ELPylated spidroin-based proteins results in non-cytotoxic biopolymers. *BMC Biotechnol.* 15, 9. doi: 10.1186/s12896-015-0123-2
- Hayashi, C. Y., Shipley, N. H., and Lewis, R. V. (1999). Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *Int. J. Biol. Macromol.* 24 (2–3), 271–275. doi: 10.1016/S0141-8130(98)00089-0
- Hedhammar, M., Rising, A., Grip, S., Martinez, A. S., Nordling, K., Casals, C., et al. (2008). Structural properties of recombinant nonrepetitive and repetitive parts of major ampullate spidroin 1 from *Euprostenops australis*: implications for fiber formation. *Biochemistry* 47 (11), 3407–3417. doi: 10.1021/bi702432y
- Heiby, J. C., Rajab, S., Rat, C., Johnson, C. M., and Neuweiler, H. (2017). Conservation of folding and association within a family of spidroin N-terminal domains. *Sci. Rep.* 7 (1), 16789. doi: 10.1038/s41598-017-16881-6
- Heppner, R., Weichert, N., Schierhorn, A., Conrad, U., and Pietzsch, M. (2016). Low-tech, pilot scale purification of a recombinant spider silk protein analog from tobacco leaves. *Int. J. Mol. Sci.* 17 (10), 1687. doi: 10.3390/ijms17101687
- Hieta, R., and Myllyharju, J. (2002). Cloning and characterization of a low molecular weight prolyl 4-hydroxylase from *Arabidopsis thaliana*. effective hydroxylation of proline-rich, collagen-like, and hypoxia-inducible transcription factor alpha-like peptides. *J. Biol. Chem.* 277 (26), 23965–23971. doi: 10.1074/jbc.M201865200
- Hinman, M., Dong, Z., Xu, M., and Lewis, R. V. (1992). Spider silk: a mystery starting to unravel. *Results Probl. Cell Differ.* 19, 227–254. doi: 10.1007/978-3-540-47207-0\_8
- Holland, C., Numata, K., Rnjak-Kovacic, J., and Seib, F. P. (2019). The biomedical use of silk: past, present, future. *Adv. Healthc. Mat.* 8 (1), e1800465. doi: 10.1002/adhm.201800465
- Huang, W., Ling, S., Li, C., Omenetto, F. G., and Kaplan, D. L. (2018). Silkworm silk-based materials and devices generated using bio-nanotechnology. *Chem. Soc. Rev.* 47 (17), 6486–6504. doi: 10.1039/C8CS00187A



- Huby, N., Vie, V., Renault, A., Beaufile, S., Lefevre, T., Paquet-Mercier, F., et al. (2013). Native spider silk as a biological optical fiber. *Appl. Phys. Lett.* 102 (12), 123702. doi: 10.1063/1.4798552
- Huemmerich, D., Helsen, C. W., Quedzuweit, S., Oschmann, J., Rudolph, R., and Scheibel, T. (2004a). Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry* 43, 13604–13612. doi: 10.1021/bi048983q
- Huemmerich, D., Scheibel, T., Vollrath, F., Cohen, S., Gat, U., and Ittah, S. (2004b). Novel assembly properties of recombinant spider dragline silk proteins. *Curr. Biol.* 14, 2070–2207. doi: 10.1016/j.cub.2004.11.005
- Hutton, J. J. Jr, Kaplan, A., and Udenfriend, S. (1967). Conversion of the amino acid sequence gly-pro-pro in protein to gly-pro-hyp by collagen proline hydroxylase. *Arch. Biochem. Biophys.* 121, 384–391. doi: 10.1016/0003-9861(67)90091-4
- Ittah, S., Cohen, S., Garty, S., Cohn, D., and Gat, U. (2006). An essential role for the C-terminal domain of a dragline spider silk protein in directing fiber formation. *Biomacromolecules* 7 (6), 1790–1795. doi: 10.1021/bm060120k
- Järveläinen, H., Sainio, A., Koulu, M., Wight, T. N., and Penttinen, R. (2009). Extracellular matrix molecules: potential targets in pharmacotherapy. *Pharmacol. Rev.* 61 (2), 198–223. doi: 10.1124/pr.109.001289
- Jansson, R., Thatikonda, N., Lindberg, D., Rising, A., Johansson, J., Nygren, P. A., et al. (2014). Recombinant spider silk genetically functionalized with affinity domains. *Biomacromolecules* 15 (5), 1696–1706. doi: 10.1021/bm500114e
- Johansson, U., Ria, M., Avall, K., Dekki Shalaly, N., Zaitsev, S. V., Berggren, P. O., et al. (2015). Pancreatic Islet Survival and Engraftment Is Promoted by Culture on Functionalized Spider Silk Matrices. *PLoS One* 10 (6), e0130169. doi: 10.1371/journal.pone.0130169
- John, D. C., Watson, R., Kind, A. J., Scott, A. R., Kadler, K. E., and Bulleid, N. J. (1999). Expression of an engineered form of recombinant procollagen in mouse milk. *Nat. Biotechnol.* 17 (4), 385–389. doi: 10.1038/7945
- Kahles, F., Findeisen, H. M., and Brummer, D. (2014). Osteopontin: A novel regulator at the cross roads of inflammation, obesity and diabetes. *Mol. Metab.* 3 (4), 384–393. doi: 10.1016/j.molmet.2014.03.004
- Kirk, T. Z., Evans, J. S., and Veis, A. (1987). Biosynthesis of type I procollagen. Characterization of the distribution of chain sizes and extent of hydroxylation of polysome-associated pro- $\alpha$ -chains. *J. Biol. Chem.* 262 (12), 5540–5545. Retrieved from <http://www.jbc.org/content/262/12/5540>.
- Kivirikko, K. I., Myllylä, R., and Pihlajaniemi, T. (1992). "Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins," in *Post-translational modifications of proteins*. Eds. J. J. Harding and M.J.C. Crabbe (Boca Raton: CRC Press), 1–52.
- Koide, T., and Nagata, K. (2005). "Collagen Biosynthesis," in *Collagen. Topics in Current Chemistry*, vol. 247. Eds. J. Brinckmann, H. Notbohm and P. K. Müller (Springer, Berlin, Heidelberg).
- Kotch, F. W., Guzei, I. A., and Raines, R. T. (2008). Stabilization of the collagen triple helix by O-methylation of hydroxyproline residues. *J. Am. Chem. Soc.* 130, 2952–2953. doi: 10.1021/ja800225k
- Kuhbier, J. W., Cogger, V., Mueller, J., Liebsch, C., Schlottmann, F., Bucan, V., et al. (2017). Influence of direct or indirect contact for the cytotoxicity and blood compatibility of spider silk. *J. Mat. Sci. Mat. Med.* 28 (8), 127. doi: 10.1007/s10856-017-5936-1
- Lai, W. L., and Goh, K. L. (2015). Consequences of ultra-violet irradiation on the mechanical properties of spider silk. *J. Funct. Biomater.* 6 (3), 901–916. doi: 10.3390/jfb6030901
- Landreh, M., Andersson, M., Marklund, E. G., Jia, Q., Meng, Q., Johansson, J., et al. (2017). Mass spectrometry captures structural intermediates in protein fiber self-assembly. *Chem. Commun. (Camb.)* 53 (23), 3319–3322. doi: 10.1039/c7cc00307b
- Lavine, M., Frisk, M., and Pennisi, E. (2012). Biomaterials. *Science* 338 (6109), 899. doi: 10.1126/science.338.6109.899
- Lazaris, A., Arcidiacono, S., Huang, Y., Zhou, J.-F., Duguay, F., Chretien, N., et al. (2002). Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. *Science* 295 (5554), 472–476. doi: 10.1126/science.1065780
- Lee, M., Kwon, J., and Na, S. (2016). Mechanical behavior comparison of spider and silkworm silks using molecular dynamics at atomic scale. *Phys. Chem. Chem. Phys.* 18 (6), 4814–4821. doi: 10.1039/c5cp06809f
- Lewis, R. V., Hinman, M., Kothakota, S., and Fournier, M. J. (1996). Expression and purification of a spider silk protein: a new strategy for producing repetitive proteins. *Protein Expr. Purif.* 7 (4), 400–406. doi: 10.1006/prep.1996.0060
- Liu, B., Wang, T., Xiao, L., Zhang, G., Li, G., Luo, J., et al. (2018). A directed self-assembly quasi-spider silk protein expressed in *Pichia pastoris*. *Biotechnol. Biotechnol. Equip.* 32 (2), 451–461. doi: 10.1080/13102818.2017.1327823
- Llop-Tous, I., Madurga, S., Giral, E., Marzabal, P., Torrent, M., and Ludevid, M. D. (2010). Relevant elements of a Maize  $\gamma$ -zein domain involved in protein body biogenesis. *J. Biol. Chem.* 285, 35633–35644. doi: 10.1074/jbc.M110.116285
- Lucke, M., Winter, G., and Engert, J. (2015). The effect of steam sterilization on recombinant spider silk particles. *Int. J. Pharm.* 481 (1–2), 125–131. doi: 10.1016/j.ijpharm.2015.01.024
- Lund, S. A., Giachelli, C. M., and Scatena, M. (2009). The role of osteopontin in inflammatory processes. *J. Cell Commun. Signaling* 3 (3–4), 311–322. doi: 10.1007/s12079-009-0068-0
- Lyda, T. A., Wagner, E. L., Bourg, A. X., Peng, C., Tomaraei, G. N., Dean, D., et al. (2017). A Leishmania secretion system for the expression of major amputate spidroin mimics. *PLoS One* 12 (5), e0178201. e017820110.1371/journal.pone.0178201
- Marhabaie, M., Leeper, T. C., and Blackledge, T. A. (2014). Protein composition correlates with the mechanical properties of spider (*Argiope trifasciata*) dragline silk. *Biomacromolecules* 15 (1), 20–29. doi: 10.1021/bm401110b
- Marin Viegas, V. S., Ocampo, C. G., and Petrucci, S. (2016). Vacuolar deposition of recombinant proteins in plant vegetative organs as a strategy to increase yields. *Bioengineered* 8 (3), 203–211. doi: 10.1080/21655979.2016.1222994
- Menassa, R., Zhu, H., Karatzas, C. N., Lazaris, A., Richman, A., and Brandle, J. (2004). Spider dragline silk proteins in transgenic tobacco leaves: accumulation and field production. *Plant Biotechnol. J.* 2 (5), 431–438. doi: 10.1111/j.1467-7652.2004.00087.x
- Merle, C., Perret, S., Lacour, T., Jonval, V., Hudaverdian, S., Garrone, R., et al. (2002). Hydroxylated human homotrimeric collagen I in Agrobacterium tumefaciens-mediated transient expression and in transgenic tobacco plant. *FEBS Lett.* 515 (1–3), 114–118. doi: 10.1016/S0014-5793(02)02452-3
- Meyer, D. E., and Chilkoti, A. (1999). Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nat. Biotechnol.* 17 (11), 1112–1115. doi: 10.1038/15100
- Miao, Y., Zhang, Y., Nakagaki, K., Zhao, T., Zhao, A., Meng, Y., et al. (2006). Expression of spider flagelliform silk protein in *Bombyx mori* cell line by a novel Bac-to-Bac/BmNPV baculovirus expression system. *Appl. Microbiol. Biotechnol.* 71 (2), 192–199. doi: 10.1007/s00253-005-0127-2
- Michal, C. A., Simmons, A. H., Chew, B. G., Zax, D. B., and Jelinski, L. W. (1996). Presence of phosphorus in Nephila clavipes dragline silk. *Biophys. J.* 70 (1), 489–493. doi: 10.1016/S0006-3495(96)79594-2
- Miller, E. J., and Gay, S. (1982). Collagen: an overview. *Methods Enzymol.* 82 (Pt A), 3–32. doi: 10.1016/0076-6879(82)82058-2
- Mithieux, S. M., and Weiss, A. S. B. T.-A. (2005). "Elastin," in *Fibrous Proteins: Coiled-Coils, Collagen and Elastomers*, (Amsterdam, Netherlands: Elsevier Academic Press) vol. 70, 437–461. doi: 10.1016/S0065-3233(05)70013-9
- Motriuk-Smith, D., Smith, A., Hayashi, C. Y., and Lewis, R. V. (2005). Analysis of the conserved N-terminal domains in major amputate spider silk proteins. *Biomacromolecules* 6 (6), 3152–3159. doi: 10.1021/bm050472b
- Mukherjee, B. B., Nemir, M., Beninati, S., Cordella-Miele, E., Singh, K., Chackalaparampil, I., et al. (1995). Interaction of osteopontin with fibronectin and other extracellular matrix molecules. *Ann. New York Acad. Sci.* 760, 201–212. doi: 10.1111/j.1749-6632.1995.tb44631.x
- Mullins, R. J., Richards, C., and Walker, T. (1996). Allergic reactions to oral, surgical and topical bovine collagen. Anaphylactic risk for surgeons. *Aust. New Z. J. Ophthalmol.* 24 (3), 257–260. doi: 10.1111/j.1442-9071.1996.tb01589.x
- Myllyharju, J. (2003). Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biol.* 22, 15–24. doi: 10.1016/S0945-053X(03)00006-4
- W. Nentwig (Ed.) (2013). *Spider ecophysiology* (Dordrecht London: Springer Heidelberg New York).
- Nokelainen, M., Tu, H., Vuorela, A., Notbohm, H., Kivirikko, K. I., and Myllyharju, J. (2001). High-level production of human type I collagen in the yeast *Pichia pastoris*. *Yeast* 18 (9), 797–806. doi: 10.1002/yea.730

- Numata, K., Reagan, M. R., Goldstein, R. H., Rosenblatt, M., and Kaplan, D. L. (2011). Spider silk-based gene carriers for tumor cell-specific delivery. *Bioconjug. Chem.* 22 (8), 1605–1610. doi: 10.1021/bc200170u
- Olsen, D., Yang, C., Bodo, M., Chang, R., Leigh, S., Baez, J., et al. (2003). Recombinant collagen and gelatin for drug delivery. *Adv. Drug Deliv. Rev.* 55 (12), 1547–1567. doi: 10.1016/j.addr.2003.08.008
- Olsen, D., Jiang, J., Chang, R., Duffy, R., Sakaguchi, M., Leigh, S., et al. (2005). Expression and characterization of a low molecular weight recombinant human gelatin: development of a substitute for animal-derived gelatin with superior features. *Protein Expr. Purif.* 40 (2), 346–357. doi: 10.1016/j.pep.2004.11.016
- Park, J.-E., Jeong, Y. J., Park, J. B., Kim, H. Y., Yoo, Y. H., Lee, K. S., et al. (2019). Dietary exposure to transgenic rice expressing the spider silk protein fibroin reduces blood glucose levels in diabetic mice: the potential role of insulin receptor substrate-1 phosphorylation in adipocytes. *Dev. Reprod.* 23 (3), 223–229. doi: 10.12717/DR.2019.23.3.223
- Parry, D. A. D., David, A. D. P., and John, M. S. (2005). “Structural and Functional Implications of Sequence Repeats in Fibrous Proteins,” in *Adv. Protein Chem.*, vol. 70, (Amsterdam, Netherlands: Elsevier Academic Press) 11–35. doi: 10.1016/S0065-3233(05)70002-4
- Patil, S. P., Xiao, S., Gkagkas, K., Markert, B., and Grater, F. (2014). Viscous friction between crystalline and amorphous phase of dragline silk. *PLoS One* 9 (8), e104832. doi: 10.1371/journal.pone.0104832
- Peng, C. A., Russo, J., Gravgard, C., McCartney, H., Gaines, W., and Marcotte, W. R. Jr. (2016). Spider silk-like proteins derived from transgenic *Nicotiana tabacum*. *Transgenic Res.* 25 (4), 517–526. doi: 10.1007/s11248-016-9949-1
- Pereira, A. M., Machado, R., da Costa, A., Ribeiro, A., Collins, T., Gomes, A. C., et al. (2017). Silk-based biomaterials functionalized with fibronectin type II promotes cell adhesion. *Acta Biomater.* 47, 50–59. doi: 10.1016/j.actbio.2016.10.002
- Perret, S., Merle, C., Bernocco, S., Berland, P., Garrone, R., Hulmes, D. J. S., et al. (2001). Unhydroxylated triple helical collagen I produced in transgenic plants provides new clues on the role of hydroxyproline in collagen folding and fibril formation. *J. Biol. Chem.* 276 (47), 43693–43698. doi: 10.1074/jbc.M105507200
- Piruzian, E. S., Bogush, V. G., Sidoruk, K. V., Goldenkova, I. V., Musiichuk, K. A., and Debabov, V. G. (2003). [Construction of the synthetic genes for protein analogs of spider silk carcass spidroin I and their expression in tobacco plants]. *Mol. Biol. (Mosk)* 37 (4), 654–662.
- Rattanapit, K., Abdulheem, S., Chaikeawkaew, D., Kubera, A., Mason, H. S., Ma, J. K.-C., et al. (2017). Recombinant human osteopontin expressed in *Nicotiana benthamiana* stimulates osteogenesis related genes in human periodontal ligament cells. *Sci. Rep.* 7 (1), 17358. doi: 10.1038/s41598-017-17666-7
- Reinholt, F. P., Hulthenby, K., Oldberg, A., and Heinegard, D. (1990). Osteopontin e a possible anchor of osteoclasts to bone. *Proc. Natl. Acad. Sci. U. S. A.* 87, 4473e4475. doi: 10.1073/pnas.87.12.4473
- Reuter, L., Ritala, A., Linder, M., and Joensuu, J. (2016). Novel hydrophobin fusion tags for plant-produced fusion proteins. *PLoS One* 11 (10), e0164032. doi: 10.1371/journal.pone.0164032
- Ries, J., Schwarze, S., Johnson, C. M., and Neuweiler, H. (2014). Microsecond folding and domain motions of a spider silk protein structural switch. *J. Am. Chem. Soc.* 136 (49), 17136–17144. doi: 10.1021/ja508760a
- Riet, J., Ramos, L. R. V., Lewis, R. V., and Marins, L. F. (2017). Improving the PCR protocol to amplify a repetitive DNA sequence. *Genet. Mol. Res.: GMR* 16 (3), gmr16039796. doi: 10.4238/gmr16039796
- Rising, A., Hjalms, G., Engstrom, W., and Johansson, J. (2006). N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindrical spider silk proteins. *Biomacromolecules* 7 (11), 3120–3124. doi: 10.1021/bm060693x
- Ritala, A., Wahlstrom, E. H., Holkeri, H., Hafren, A., Makelainen, K., Baez, J., et al. (2008). Production of a recombinant industrial protein using barley cell cultures. *Protein Expr. Purif.* 59, 274–281. doi: 10.1016/j.pep.2008.02.013
- Ruggiero, F., Exposito, J. Y., Bournat, P., Gruber, V., Perret, S., Comte, J., et al. (2000). Triple helix assembly and processing of human collagen produced in transgenic tobacco plants. *FEBS Lett.* 469 (1), 132–136. doi: 10.1016/S0014-5793(00)01259-X
- Saberianfar, R., Joensuu, J. J., Conley, A. J., and Menassa, R. (2015). Protein body formation in leaves of *Nicotiana benthamiana*: a concentration-dependent mechanism influenced by the presence of fusion tags. *Plant Biotechnol. J.* 13, 927–937. doi: 10.1111/pbi.12329
- Santos-Pinto, J. R. A., Dos Lamprecht, G., Chen, W. Q., Heo, S., Hardy, J. G., Priewalder, H., et al. (2014). Structure and post-translational modifications of the web silk protein spidroin-1 from *Nephila* spiders. *J. Proteomics* 105, 174–185. doi: 10.1016/j.jprot.2014.01.002
- Santos-Pinto, J. R. A., Dos Arcuri, H. A., Priewalder, H., Salles, H. C., Palma, M. S., and Lubec, G. (2015). Structural Model for the Spider Silk Protein Spidroin-1. *J. Proteome Res.* 14 (9), 3859–3870. doi: 10.1021/acs.jproteome.5b00243
- Santos-Pinto, J. R. A., Dos Arcuri, H. A., Lubec, G., and Palma, M. S. (2016). Structural characterization of the major ampullate silk spidroin-2 protein produced by the spider *Nephila clavipes*. *Biochim. Biophys. Acta* 1864 (10), 1444–1454. doi: 10.1016/j.bbapap.2016.05.007
- Santos-Pinto, J. R. A., Dos Esteves, F. G., Sialana, F. J., Ferro, M., Smidak, R., Rares, L. C., et al. (2019). A proteotranscriptomic study of silk-producing glands from the orb-weaving spiders. *Mol. Omics* 15 (4), 256–270. doi: 10.1039/c9mo00087a
- Scatena, M., Liaw, L., and Giachelli, C. M. (2007). Osteopontin: a multifunctional molecule regulating chronic inflammation and vascular disease. *Arterioscler. Thromb. Vasc. Biol.* 27, 2302e2309. doi: 10.1161/ATVBAHA.107.144824
- Scheller, J., and Conrad, U. (2005). “Spider Silk Proteins from Transgenic Plants,” in *Biopolymers Online* (Hoboken, New Jersey: Wiley-Blackwell). doi: 10.1002/3527600035.bpol8004
- Scheller, J., Guhrs, K. H., Grosse, F., and Conrad, U. (2001). Production of spider silk proteins in tobacco and potato. *Nat. Biotechnol.* 19 (6), 573–577. doi: 10.1038/89335
- Scheller, J., Henggeler, D., Viviani, A., and Conrad, U. (2004). Purification of Spider Silk-elastin from Transgenic Plants and Application for Human Chondrocyte Proliferation. *Transgenic Res.* 13 (1), 51–57. doi: 10.1023/B:TRAG.0000017175.78809.7a
- Schillberg, S., Raven, N., Spiegel, H., Rasche, S., and Buntru, M. (2019). Critical analysis of the commercial potential of plants for the production of recombinant proteins. *Front. Plant Sci.* 10, 720. doi: 10.3389/fpls.2019.00720
- Schneider, D., Gomopoulos, N., Koh, C. Y., Papadopoulos, P., Kremer, F., Thomas, E. L., et al. (2016). Nonlinear control of high-frequency phonons in spider silk. *Nat. Mat.* 15 (10), 1079–1083. doi: 10.1038/nmat4697
- Service, R. F. (2002). *Materials science. Mammalian cells spin a spidery new yarn.* *Science* 295 (5554), 419–421. doi: 10.1126/science.295.5554.419b
- Setina, C. M., Haase, J. P., and Glatz, C. E. (2015). Process integration for recovery of recombinant collagen type I alpha 1 from corn seed. *Biotechnol. Prog.* 32 (1), 98–107. doi: 10.1002/btpr.2191
- Shoseyov, O., Posen, Y., and Grynspan, F. (2014). Human collagen produced in plants: more than just another molecule. *Bioengineered* 5 (1), 49–52. doi: 10.4161/bioe.26002
- Shoulders, M. D., and Raines, R. T. (2009). Collagen structure and stability. *Annu. Rev. Biochem.* 78, 929–958. doi: 10.1146/annurev.biochem.77.032207.120833
- Simmons, A. H., Michal, C. A., and Jelinski, L. W. (1996). Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science* 271 (5245), 84–87.
- Singh, B., Fleury, C., Jalalvand, F., and Riesbeck, K. (2012). Human pathogens utilize host extracellular matrix proteins laminin and collagen for adhesion and invasion of the host. *FEMS Microbiol. Reviews* 36, 1122–1180. doi: 10.1126/science.271.5245.84
- Sorushanova, A., Delgado, L. M., Wu, Z., Shologu, N., Kshirsagar, A., Raghunath, R., et al. (2019). The collagen suprafamily: from biosynthesis to advanced biomaterial development. *Adv. Mat.* 31 (1), 1801651. doi: 10.1002/adma.201801651
- Spiegel, H., Boes, A., Perales Morales, C., Rademacher, T., and Buyel, J. F. (2019). Ready-to-use stocks of agrobacterium tumefaciens can simplify process development for the production of recombinant proteins by transient expression in plants. *Biotechnol. J.* 14 (10), 1900113. doi: 10.1002/biot.201900113
- Sponner, A., Schlott, B., Vollrath, F., Unger, E., Grosse, F., and Weisshart, K. (2005a). Characterization of the protein components of *Nephila clavipes* dragline silk. *Biochemistry* 44 (12), 4727–4736. doi: 10.1021/bi047671k
- Sponner, A., Vater, W., Rommerskirch, W., Vollrath, F., Unger, E., Grosse, F., et al. (2005b). The conserved C-termini contribute to the properties of spider silk fibroins. *Biochem. Biophys. Res. Commun.* 338, 897–902. doi: 10.1016/j.bbrc.2005.10.048
- Stein, H., Wilensky, M., Tsafir, Y., Rosenthal, M., Amir, R., Avraham, T., et al. (2009). Production of bioactive, post-translationally modified, heterotrimeric,

- human recombinant type-I collagen in transgenic tobacco. *Biomacromolecules* 10 (9), 2640–2645. doi: 10.1021/bm900571b
- Steuers, L., Glorieux, S., Vandekerckhove, A. P., Favoreel, H. W., and Nauwynck, H. J. (2012). Diverse microbial interactions with the basement membrane barrier. *Trends Microbiol.* 20 (3), 147–155. doi: 10.1016/j.tim.2012.01.001
- Stover, D. A., and Verrelli, B. C. (2010). Comparative vertebrate evolutionary analyses of type I collagen: potential of COL1A1 gene structure and intron variation for common bone-related diseases. *Mol. Biol. Evol.* 28 (1), 533–542. doi: 10.1093/molbev/msq221
- Tanaka, M., Sato, K., and Uchida, T. (1981). Plant prolyl hydroxylase recognizes poly(L-proline) II helix. *J. Biol. Chem.* 256 (22), 11397–11400.
- Tasiopoulos, C. P., Widhe, M., and Hedhammar, M. (2018). Recombinant spider silk functionalized with a motif from fibronectin mediates cell adhesion and growth on polymeric substrates by entrapping cells during self-assembly. *ACS Appl. Mat. Interfaces* 10 (17), 14531–14539. doi: 10.1021/acsami.8b02647
- Teulé, F., Miao, Y.-G., Sohn, B.-H., Kim, Y.-S., Hull, J. J., Fraser, M. J., et al. (2011). Silkworms transformed with chimeric silkworm/spider silk genes spin composite silk fibers with improved mechanical properties. *Proc. Natl. Acad. Sci. U. S. A.* 109 (3), 923–928. doi: 10.1073/pnas.1109420109
- Teulé, F., Addison, B., Cooper, A. R., Ayon, J., Henning, R. W., Benmore, C. J., et al. (2012). Combining flagelliform and dragline spider silk motifs to produce tunable synthetic biopolymer fibers. *Biopolymers* 97 (6), 418–431. doi: 10.1002/bip.21724
- Thurber, A. E., Omenetto, F. G., and Kaplan, D. L. (2015). In vivo bioresponses to silk proteins. *Biomaterials* 71, 145–157. doi: 10.1016/j.biomaterials.2015.08.039
- Toman, P. D., Chisholm, G., McMullin, H., Giere, L. M., Olsen, D. R., Kovach, R. J., et al. (2000). Production of recombinant human type I procollagen trimers using a four-gene expression system in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* 275, 23303–23309. doi: 10.1074/jbc.M002284200
- Tomita, M., Munetsuna, H., Sato, T., Adachi, T., Hino, R., Hayashi, M., et al. (2003). Transgenic silkworms produce recombinant human type III procollagen in cocoons. *Nat. Biotechnol.* 21 (1), 52–56. doi: 10.1038/nbt771
- Urry, D. W., Hugel, T., Seitz, M., Gaub, H. E., Sheiba, L., Dea, J., et al. (2002). Elastin: a representative ideal protein elastomer. *Philos. Trans. R. Soc. Lond. Ser. B. Biol. Sci.* 357 (1418), 169–184. doi: 10.1098/rstb.2001.1023
- Vicentini, A., Gesztes, J., and Franco, M. (1994). Binding of *Paracoccidioides brasiliensis* to laminin through surface glycoprotein gp43. *Infect. Immun.* 62 (4), 1465–1469.
- Vollrath, F. (2016). The complexity of silk under the spotlight of synthetic biology. *Biochem. Soc. Trans.* 44 (4), 1151–1157. doi: 10.1042/BST20160058
- Weichert, N., Hauptmann, V., Menzel, M., Schallau, K., Gunkel, P., Hertel, T. C., et al. (2014). Transglutamination allows production and characterization of native-sized ELPylated spider silk proteins from transgenic plants. *Plant Biotechnol. J.* 12 (2), 265–275. doi: 10.1111/pbi.12135
- Weichert, N., Hauptmann, V., Helmold, C., and Conrad, U. (2016). Seed-specific expression of spider silk protein multimers causes long-term stability. *Front. Plant Sci.* 7, 6. doi: 10.3389/fpls.2016.00006
- Werkmeister, J. A., and Ramshaw, J. A. M. (2012). Recombinant protein scaffolds for tissue engineering. *Biomed. Mat.* 7 (1), 12002. doi: 10.1088/1748-6041/7/1/012002
- Widhe, M., Johansson, U., Hillerdahl, C. O., and Hedhammar, M. (2013). Recombinant spider silk with cell binding motifs for specific adherence of cells. *Biomaterials* 74 (33), 8223–8234. doi: 10.1016/j.biomaterials.2013.07.058
- Widhe, M., Shalaly, N. D., and Hedhammar, M. (2016). A fibronectin mimetic motif improves integrin mediated cell binding to recombinant spider silk matrices. *Biomaterials* 74, 256–266. doi: 10.1016/j.biomaterials.2015.10.013
- Widmaier, D. M., Tullman-Ercek, D., Mirsky, E. A., Hill, R., Govindarajan, S., Minshall, J., et al. (2009). Engineering the *Salmonella* type III secretion system to export spider silk monomers. *Mol. Syst. Biol.* 309. doi: 10.1038/msb.2009.62
- Willard, J. J., Drexler, J. W., Das, A., Roy, S., Shilo, S., Shoseyov, O., et al. (2013). Plant-derived human collagen scaffolds for skin tissue engineering. *Tissue Eng. Part A* 19 (13–14), 1507–1518. doi: 10.1089/ten.TEA.2012.0338
- Winkler, S., Wilson, D., and Kaplan, D. L. (2000). Controlling beta-sheet assembly in genetically engineered silk by enzymatic phosphorylation/dephosphorylation. *Biochemistry* 39 (41), 12739–12746. doi: 10.1021/bi001335w
- Wise, S. G., and Weiss, A. S. (2009). Tropoelastin. *Int. J. Biochem. Cell Biol.* 41 (3), 494–497. doi: 10.1016/j.biocel.2008.03.017
- Wong Po Foo, C., and Kaplan, D. L. (2002). Genetic engineering of fibrous proteins: spider dragline silk and collagen. *Adv. Drug Deliv. Rev.* 54 (8), 1131–1143. doi: 10.1016/s0169-409x(02)00061-3
- Xia, X. X., Qian, Z. G., Ki, C. S., Park, Y. H., Kaplan, D. L., and Lee, S. Y. (2010). 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 (32), 14059–14063. doi: 10.1073/pnas.1003366107
- Xiao, W., He, J., Nichol, J. W., Wang, L., Hutson, C. B., Wang, B., et al. (2011). Synthesis and characterization of photocrosslinkable gelatin and silk fibroin interpenetrating polymer network hydrogels. *Acta Biomater.* 7, 2384–2393. doi: 10.1016/j.actbio.2011.01.016
- Xu, M., and Lewis, R. V. (1990). Structure of a protein superfiber: spider dragline silk. *Proc. Natl. Acad. Sci. U. S. A.* 87 (18), 7120–7124. doi: 10.1073/pnas.87.18.7120
- Xu, H. T., Fan, B. L., Yu, S. Y., Huang, Y. H., Zhao, Z. H., Lian, Z. X., et al. (2007). Construct synthetic gene encoding artificial spider dragline silk protein and its expression in milk of transgenic mice. *Anim. Biotechnol.* 18 (1), 1–12. doi: 10.1080/10495390601091024
- Xu, X., Gan, Q., Clough, R. C., Pappu, K. M., Howard, J. A., Baez, J. A., et al. (2011). Hydroxylation of recombinant human collagen type I alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase. *BMC Biotechnol.* 11 (1), 69. doi: 10.1186/1472-6750-11-69
- Xu, G., Gong, L., Yang, Z., and Liu, X. Y. (2014). What makes spider silk fibers so strong? From molecular-crystallite network to hierarchical network structures. *Soft Matter* 10 (13), 2116–2123. doi: 10.1039/c3sm52845f
- Yang, J., Barr, L. A., Fahnestock, S. R., and Liu, Z.-B. (2005). High yield recombinant silk-like protein production in transgenic plants through protein targeting. *Transgenic Res.* 14 (3), 313–324. doi: 10.1007/s11248-005-0272-5
- Yokosaki, Y., Tanaka, K., Higashikawa, F., Yamashita, K., and Eboshida, A. (2005). Distinct structural requirements for binding of the integrins  $\alpha$ 6 $\beta$ 6,  $\alpha$ 6 $\beta$ 3,  $\alpha$ 5 $\beta$ 5,  $\alpha$ 5 $\beta$ 1 and  $\alpha$ 9 $\beta$ 1 to osteopontin. *Matrix Biol.* 24 (6), 418–427. doi: 10.1016/j.matbio.2005.05.005
- Zhang, C., Baez, J., Pappu, K. M., and Glatz, C. E. (2009a). Purification and characterization of a transgenic corn grain-derived recombinant collagen type I alpha 1. *Biotechnol. Prog.* 25 (6), 1660–1668. doi: 10.1002/btpr.257
- Zhang, C., Glatz, C. E., Fox, S. R., and Johnson, L. A. (2009b). Fractionation of transgenic corn seed by dry and wet milling to recover recombinant collagen-related proteins. *Biotechnol. Prog.* 25, 1396–1401. doi: 10.1002/btpr.220
- Zhang, C., Baez, J., and Glatz, C. E. (2009c). Purification and characterization of a 44-kDa recombinant collagen I alpha 1 fragment from corn grain. *J. Agric. Food Chem.* 57, 880–887. doi: 10.1021/jf8026205
- Zhang, X., Xia, L., Day, B. A., Harris, T. I., Oliveira, P., Knittel, C., et al. (2019). CRISPR/Cas9 initiated transgenic silkworms as a natural spinner of spider silk. *Biomacromolecules* 20 (6), 2252–2264. doi: 10.1021/acs.biomac.9b00193

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Peng, Kozubowski and Marcotte. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.