# PROCESSESS FOR PRODUCING SILK DOPE

## FIELD OF THE INVENTION

The present invention relates to methods of producing silk dope comprising silk proteins with a coiled-coil structure such as honeybee silk proteins. The silk dope can be used for a variety of purposes such as in the production of personal care products, plastics, textiles, and biomedical products.

## BACKGROUND OF THE INVENTION

Silks are protein fibres produced by a wide range of insect and spider species. The silk of the domesticated silkworm, *Bombyx mori*, has been used as a suture biomaterial for centuries. Numerous efforts to clone and express silkworm or spider silks in transgenic systems have found it a Herculean task. The large sizes and highly repetitive sequences of these silk genes make them recalcitrant to expression outside specialized silk glands, and lead to low protein yields.

Although silkworm cocoons and spider webs are the best known silks, other species may produce silks better suited to transgenic production. Honeybee larvae (*Apis mellifera*) spin silk cocoons in which they pupate. Honeybee silk is encoded by four small (~30kDa each) and non-repetitive fibre genes (Sutherland et al., 2006). Homologous sets of four genes have also been found in bumblebees, bulldog ants, weaver ants, hornets and Asiatic honeybees (Sutherland et al., 2007; Sezutsu et al., 2007; Shi et al., 2008; WO 2007/038837).

Vintage x-ray fibre diffraction work demonstrated that honeybee silk contains α-helical proteins assembled into a coiled-coil conformation, most likely a tetrameric coiled-coil structure (Atkins, 1967), with the four strands likely corresponding to the four different silk proteins. Bioinformatics techniques predict that each of the honeybee silk protein sequences contains 60-68% coiled-coil (Sutherland et al., 2006).

Silk threads can be hand-drawn from the silk glands of honeybee larvae. These threads are less strong but more extensible and tougher than silkworm silk fibres (Hepburn et al., 1979).

Shi et al. (2008) recently reported recombinant production of Asiatic honeybee silk (*Apis cerana*). The four *A. cerana* silk proteins were expressed in a soluble form in *Escherichia coli* with yields of 10-60 mg per litre of ferment. A variety of experimental techniques were used to characterize the structure and interactions of the proteins at low concentration (0.03 to 0.2 wt%). These conclusively demonstrated that neither the individual proteins nor a mix of four proteins had tight tertiary packing in solution.

The proteins existed as monomers or loosely associated dimers and had predominantly random-coil conformation with little α-helical structure.

There is a need for further methods to produce silk dope from recombinantly expressed coiled-coil silk proteins which can be used to manufacture a wide variety of products.

## SUMMARY OF THE INVENTION

The present inventors have surprisingly found that surfactants and ionic liquids can be used in a process to produce silk dope comprising coiled-coil silk proteins.

In a first aspect, the present invention provides a method for producing silk dope, the method comprising

i) lysing cells producing one or more silk proteins,

ii) solubilising the silk proteins by contacting them with a surfactant or an ionic liquid, and

iii) concentrating the silk proteins to produce silk dope,

wherein the one or more silk proteins are capable of forming a tertiary structure which comprises a coiled-coil structure.

In one embodiment, the silk proteins are concentrated by

a) reducing the amount of surfactant in solution by adding a compound which precipitates the surfactant, and

b) separating the solution comprising the silk proteins from the precipitate formed in step a) to produce the silk dope.

Compounds which can be used to precipitate surfactants are known in the art and include a salt or a carbohydrate; or a combination of two or more thereof. Preferably, the salt is a potassium salt or a sodium salt. In an embodiment, the carbohydrate is α-cyclodextrin.

In another embodiment, the silk proteins are concentrated by filtration, more preferably membrane filtration, and even more preferably tangential flow filtration.

In an embodiment, the method further comprises increasing the concentration of silk proteins in the silk dope. This can be achieved by any method known in the art. For example, the silk dope is dialysed against a dehydrating solution such as solution comprising a hygroscopic polymer. Examples of hygroscopic polymers include, but are not limited to, polyethylene glycol, amylase and sericin, as well as a combination of two or more thereof.

In a preferred embodiment, the silk dope comprises at least about 0.5% w/v silk proteins. In a further embodiment, the silk dope comprises about 0.5% to about 15% silk proteins.

The cell can be any cell type, typically a recombinant cell comprising an exogenous polynucleotide(s) encoding, and capable of producing, the silk protein(s). Examples include, but are not limited, to bacterial cells, yeast cells, insect cells, plant cells or animal cells, or a combination of two or more thereof. In a preferred embodiment, the cell is a bacterial cell. In a particularly preferred embodiment, the bacterial cell is *Escherichia coli*.

In a preferred embodiment, step i) further comprises isolating inclusion bodies from the lysed cells.

The method may also comprise culturing the cells before step i).

In a preferred embodiment, the portion of the silk protein that is capable of forming a tertiary structure which comprises a coiled-coil structure comprises at least 10 copies of the heptad sequence *abcdefg*, and wherein at least 25% of the amino acids at positions *a* and *d* are alanine residues. More preferably, at least 25% of the amino acids at positions *a*, *d* and *e* are alanine residues.

In a further preferred embodiment, the silk protein comprises, more preferably consists essentially of, even more preferably consists of, a sequence selected from:

a) an amino acid sequence as provided in any one of SEQ ID NOs 1 to 8, 17 to 24, 33 to 40, 49 to 56, 65 to 72, 81 to 88, 97 or 98,

b) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 1 to 8, 17 to 24, 33 to 40, 49 to 56, 65 to 72, 81 to 88, 97 or 98, and

c) a biologically active fragment of a) or b).

In the above aspect it is preferred that as little as possible of the silk proteins are secreted from the cell. Accordingly, it is preferred that the silk proteins do not comprise an N-terminal signal sequence. Examples of silk proteins particularly useful for the above aspect include, but are not limited to, silk proteins comprising, more preferably consisting essentially of, and even more preferably consisting of, a sequence selected from:

a) an amino acid sequence as provided in any one of SEQ ID NOs 1, 3, 5, 7, 17, 19, 21, 23, 33, 35, 37, 39, 49, 51, 53, 55, 65, 67, 69, 71, 81, 83, 85, 87 or 97,

b) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 1, 3, 5, 7, 17, 19, 21, 23, 33, 35, 37, 39, 49, 51, 53, 55, 65, 67, 69, 71, 81, 83, 85, 87 or 97, and

c) a biologically active fragment of a) or b).

In an embodiment, the silk proteins can be a plurality of the same silk protein or a combination of two or more different silk proteins. In a preferred embodiment, if different silk proteins are used there are four different silk proteins.

In a further embodiment, the silk proteins comprise a first silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

a) an amino acid sequence as provided in any one of SEQ ID NOs 1, 2, 17, 18, 33, 34, 49, 50, 65, 66, 81 or 82;

b) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 1, 2, 17, 18, 33, 34, 49, 50, 65, 66, 81 or 82; and

c) a biologically active fragment of a) or b),

a second silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

d) an amino acid sequence as provided in any one of SEQ ID NOs 3, 4, 19, 20, 35, 36, 51, 52, 67, 68, 83 or 84;

e) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 3, 4, 19, 20, 35, 36, 51, 52, 67, 68, 83 or 84; and

f) a biologically active fragment of c) or d),

a third silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

g) an amino acid sequence as provided in any one of SEQ ID NOs 5, 6, 21, 22, 37, 38, 53, 54, 69, 70, 85 or 86;

h) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 5, 6, 21, 22, 37, 38, 53, 54, 69, 70, 85 or 86; and

i) a biologically active fragment of g) or h), and/or

a fourth silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

j) an amino acid sequence as provided in any one of SEQ ID NOs 7, 8, 23, 24, 39, 40, 55, 56, 71, 72, 87 or 88;

k) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs7, 8, 23, 24, 39, 40, 55, 56, 71, 72, 87 or 88; and

l) a biologically active fragment of j) or k). More preferably, in relation to above aspect the silk proteins comprise, or consist essentially of, a first silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

a) an amino acid sequence as provided in any one of SEQ ID NOs 1, 17, 33, 49, 65 or 81;

b) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 1, 17, 33, 49, 65 or 81; and

c) a biologically active fragment of a) or b),

a second silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

d) an amino acid sequence as provided in any one of SEQ ID NOs 3, 19, 35, 51, 67 or 83;

e) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 3, 19, 35, 51, 67 or 83; and

f) a biologically active fragment of d) or e),

a third silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

g) an amino acid sequence as provided in any one of SEQ ID NOs 5, 21, 37, 53, 69 or 85;

h) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 5, 21, 37, 53, 69 or 85; and

i) a biologically active fragment of g) or h), and/or

a fourth silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

j) an amino acid sequence as provided in any one of SEQ ID NOs 7, 23, 39, 55, 71 or 87;

k) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 7, 23, 39, 55, 71 or 87; and

l) a biologically active fragment of j) or k).

In an embodiment, the first silk protein, second silk protein, third silk protein and/or fourth silk protein are produced by the same cells.

In an alternate embodiment, the first silk protein, second silk protein, third silk protein and/or fourth silk protein are produced by different cells. In this embodiment, it is preferred that step ii) comprises approximate equimolar amounts of the first silk protein, the second silk protein, the third silk protein and the fourth silk protein.

At any point up until and excluding step iii) the silk proteins processed according to the invention may be prepared independently and combined. The separately prepared silk proteins can be the same or different. For example, a first silk protein as defined herein is expressed in a first cell and processed as defined in steps i) and ii), a second silk protein as defined herein is expressed in a second cell and

processed as defined in steps i) and ii), and then the two solutions combined before step iii) is performed.

The surfactant and ionic liquid solubilise precipitated protein and enables the silk protein to stay in solution whilst allowing the formation of a coiled-coil structure during later steps.

In a preferred embodiment, the surfactant is an anionic surfactant. Examples of anionic surfactants useful for the invention include, but are not limited to, sodium dodecyl sulfate (SDS), ammonium lauryl sulfate and other alkyl sulfate salts, sodium 1-octanesulfonate monohydrate, sodium lauroyl sarcosinate, sodium lauryl ether sulfate (SLES), sodium taurodeoxycholate hydrate, and alkyl benzene sulfonate; as well as a combination of two or more thereof. In a preferred embodiment, the anionic surfactant is SDS.

In an embodiment, the ionic liquid comprises

i) an anion selected from chloride, bromide, iodide, thiocyanate, acetate, $C_1$-$C_4$-alkylsulfates, methanesulfonates, tosylate, $C_1$-$C_4$-dialkylphosphates, hydrogensulfate and tetrachloroaluminate, and

ii) a cation selected from 1,3-$C_1$-$C_4$-dialkylimidazolium, 3-chloropyridinium, 4-dimethylaminopyridinium, 2-ethyl-4-aminopyridinium, 2-methylpyridinium, 2-ethylpyridinium, 2-ethyl-6-methylpyridinium, quinolinium, isoquinolinium, pyridinium, 1-$C_1$-$C_4$-alkylimidazolium, 1-methylimidazolium, 1,2-dimethylimidazolium, 1-n-butyl-imidazolium, 1,4,5-trimethylimidazolium, 1,4-dimethylimidazolium, imidazolium, 2-methylimidazolium, 1-butyl-2-methylimidazolium, 4 methylimidazolium, 1-(2'-aminoethyl)imidazolium, 1-vinylimidazolium, 2-ethylimidazolium and benzotriazolium.

In a further preferred embodiment, the method yields at least about 0.1 g, more preferably at least about 1 g, more preferably at least about 1.5 g, more preferably at least about 2 g, even more preferably at least about 2.5 g. of silk protein(s) per litre of cultured cells.

In another aspect, the present invention provides a method for producing silk dope, the method comprising

i) obtaining supernatant from cell cultures, or from a cell-free expression system, producing one or more silk proteins,

ii) solubilising the silk proteins by contacting them with a surfactant or an ionic liquid, and

iii) concentrating the silk proteins to produce the silk dope,

wherein the one or more silk proteins are capable of forming a tertiary structure which comprises a coiled-coil structure.

In this aspect, instead of the silk proteins in the cell being used to produce the silk dope, silk proteins which are secreted from the cells are used. As the skilled addressee will appreciate, step i) of the first aspect and step i) of the above aspect may be performed simultaneously or sequentially. Furthermore, at any corresponding step silk proteins derived from the cell and the supernatant could be combined and from thereon processed together. For example, step ii) of the first aspect and step ii) of the above aspect can be performed separately and the silk proteins combined for further processing including steps iii) and iv).

In a particularly preferred embodiment of the above aspect, step i) further comprises increasing the concentration of silk proteins from the supernatant. This can be achieved by any method known in the art, for example by contacting the supernatant with an agent which precipitates the silk proteins such as, but not limited to, ammonium sulfate, trichloroacetic acid, perchloric acid and acetone.

In relation to the above aspect it is preferred that as much as possible of the silk proteins are secreted from the cell. Accordingly, it is preferred that the silk proteins comprise an N-terminal signal sequence. Examples of silk proteins particularly useful for the above aspect include, but are not limited to, silk proteins comprising, more preferably consisting essentially of, and even more preferably consisting of, a sequence selected from:

a) an amino acid sequence as provided in any one of SEQ ID NOs 2, 4, 6, 8, 18, 20, 22, 24, 34, 36, 38, 40, 50, 52, 54, 56, 66, 68, 70, 72, 82, 84, 86, 88 or 98,

b) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 2, 4, 6, 8, 18, 20, 22, 24, 34, 36, 38, 40, 50, 52, 54, 56, 66, 68, 70, 72, 82, 84, 86, 88 or 98, and

c) a biologically active fragment of a) or b).

In a further embodiment, the silk proteins comprise a first silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

a) an amino acid sequence as provided in any one of SEQ ID NOs 2, 18, 34, 50, 66 or 82;

b) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 2, 18, 34, 50, 66 or 82; and

c) a biologically active fragment of a) or b),

a second silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

d) an amino acid sequence as provided in any one of SEQ ID NOs 4, 20, 36, 52, 68 or 84;

e) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 4, 20, 36, 52, 68 or 84; and

f) a biologically active fragment of c) or d),

a third silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

g) an amino acid sequence as provided in any one of SEQ ID NOs 6, 22, 38, 54, 70 or 86;

h) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 6, 22, 38, 54, 70 or 86; and

i) a biologically active fragment of g) or h), and/or

a fourth silk protein which comprises, more preferably consists essentially of, even more preferably consists of,

j) an amino acid sequence as provided in any one of SEQ ID NOs 8, 24, 40, 56, 72 or 88;

k) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 8, 24, 40, 56, 72 or 88; and

l) a biologically active fragment of j) or k).

In a further aspect, the present invention provides a method for producing a silk fibre, the method comprising extruding and/or drawing silk dope produced by a method of the invention.

In an embodiment, the extruding comprises passing the silk dope through an about 5μm to about 500μm capillary tube.

In a particularly preferred embodiment, the method comprises

i) lysing cells producing one or more silk proteins and isolating inclusion bodies from the cells,

ii) solubilising the silk proteins in the inclusion bodies by contacting them with a surfactant or an ionic liquid,

iii) concentrating the silk proteins to produce silk dope,

iv) increasing the concentration of silk proteins in the silk dope to about 2% to about 10% wt(%) silk proteins, more preferably about 3% to about 6% wt(%) silk proteins, and

vi) extruding the silk dope in a dehydrating solution.

In relation to the above embodiment, the dehydrating solution preferably comprises an alcohol such as methanol or ethanol, or a high concentration of salt such

as $MgCl_2$ or NaCl. Extruding silk fibres under these conditions is generally known in the art as wet spinning.

Preferably, the alcohol is methanol and the concentration of the methanol in the solution is about 40% to about 80% v/v, more preferably about 50% to about 70% v/v. In this embodiment, the silk dope may comprise a single type of silk polypeptide as defined herein, or two, or more different types such as four different types.

In another aspect, the present invention provides a method for producing a silk film, wherein the method comprising casting silk dope produced by a method of the invention.

In another aspect, the present invention provides silk dope produced by a method of the invention.

In a further aspect, the present invention provides a silk fibre produced by a method of the invention.

Also provided is a silk film produced by a method of the invention.

In yet another aspect, the present invention provides a product comprising a silk fibre and/or silk film of the invention.

As will be apparent, preferred features and characteristics of one aspect of the invention are applicable to many other aspects of the invention.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

The invention is hereinafter described by way of the following non-limiting Examples and with reference to the accompanying figures.

## BRIEF DESCRIPTION OF THE ACCOMPANYING DRAWINGS

**Figure 1.** SDS-PAGE of purified inclusion bodies solubilised in SDS. Lanes correspond to recombinant proteins AmelF1-4; scale is protein weight in kDa.

**Figure 2.** Fourier self deconvolution of the amide I and II regions of the infrared spectra of native honeybee silk (A) and recombinant honeybee silk (B). Assignments of bands to structures are found in Table 2.

**Figure 3.** Cross-polarized microscopy of recombinant honeybee silk fibres (A) drawn in air, and (B) drawn in air then drawn a second time in methanol.

**Figure 4.** Cross-polarized microscopy of recombinant honeybee silk fibres (A) extruded into methanol bath, and (B) air dried then drawn a second time in a methanol bath to x2 length or (C) air dried then drawn a second time in a methanol bath to x4 length.

**KEY TO THE SEQUENCE LISTING**

SEQ ID NO:1 - Honeybee silk protein termed herein Xenospira1 (also termed herein AmelF1) (minus signal peptide).

SEQ ID NO:2 - Honeybee silk protein termed herein Xenospira1.

SEQ ID NO:3 - Honeybee silk protein termed herein Xenospira2 (also termed herein AmelF2) (minus signal peptide).

SEQ ID NO:4 - Honeybee silk protein termed herein Xenospira2.

SEQ ID NO:5 - Honeybee silk protein termed herein Xenospira3 (also termed herein AmelF3) (minus signal peptide).

SEQ ID NO:6 - Honeybee silk protein termed herein Xenospira3.

SEQ ID NO:7 - Honeybee silk protein termed herein Xenospira4 (also termed herein AmelF4) (minus signal peptide).

SEQ ID NO:8 - Honeybee silk protein termed herein Xenospira4.

SEQ ID NO:9 - Nucleotide sequence encoding honeybee silk protein Xenospira1 (minus region encoding signal peptide).

SEQ ID NO:10 - Nucleotide sequence encoding honeybee silk protein Xenospira1.

SEQ ID NO:11 - Nucleotide sequence encoding honeybee silk protein Xenospira2 (minus region encoding signal peptide).

SEQ ID NO:12 - Nucleotide sequence encoding honeybee silk protein Xenospira2.

SEQ ID NO:13 - Nucleotide sequence encoding honeybee silk protein Xenospira3 (minus region encoding signal peptide).

SEQ ID NO:14 - Nucleotide sequence encoding honeybee silk protein Xenospira3.

SEQ ID NO:15 - Nucleotide sequence encoding honeybee silk protein Xenospira4 (minus region encoding signal peptide).

SEQ ID NO:16 - Nucleotide sequence encoding honeybee silk protein Xenospira4.

SEQ ID NO:17 - Bumblebee silk protein termed herein BBF1 (minus signal peptide).

SEQ ID NO:18 - Bumblebee silk protein termed herein BBF1.

SEQ ID NO:19 - Bumblebee silk protein termed herein BBF2 (minus signal peptide).

SEQ ID NO:20 - Bumblebee silk protein termed herein BBF2.

SEQ ID NO:21 - Bumblebee silk protein termed herein BBF3 (minus signal peptide).

SEQ ID NO:22 - Bumblebee silk protein termed herein BBF3.

SEQ ID NO:23 - Bumblebee silk protein termed herein BBF4 (minus signal peptide).
SEQ ID NO:24 - Bumblebee silk protein termed herein BBF4.
SEQ ID NO:25 - Nucleotide sequence encoding bumblebee silk protein BBF1 (minus region encoding signal peptide).
SEQ ID NO:26 - Nucleotide sequence encoding bumblebee silk protein BBF1.
SEQ ID NO:27 - Nucleotide sequence encoding bumblebee silk protein BBF2 (minus region encoding signal peptide).
SEQ ID NO:28 - Nucleotide sequence encoding bumblebee silk protein BBF2.
SEQ ID NO:29 - Nucleotide sequence encoding bumblebee silk protein BBF3 (minus region encoding signal peptide).
SEQ ID NO:30 - Nucleotide sequence encoding bumblebee silk protein BBF3.
SEQ ID NO:31 - Nucleotide sequence encoding bumblebee silk protein BBF4 (minus region encoding signal peptide).
SEQ ID NO:32 - Nucleotide sequence encoding bumblebee silk protein BBF4.
SEQ ID NO:33 - Bulldog ant silk protein termed herein BAF1 (minus signal peptide).
SEQ ID NO:34 - Bulldog ant silk protein termed herein BAF1.
SEQ ID NO:35 - Bulldog ant silk protein termed herein BAF2 (minus signal peptide).
SEQ ID NO:36 - Bulldog ant silk protein termed herein BAF2.
SEQ ID NO:37 - Bulldog ant silk protein termed herein BAF3 (minus signal peptide).
SEQ ID NO:38 - Bulldog ant silk protein termed herein BAF3.
SEQ ID NO:39 - Bulldog ant silk protein termed herein BAF4 (minus signal peptide).
SEQ ID NO:40 - Bulldog ant silk protein termed herein BAF4.
SEQ ID NO:41 - Nucleotide sequence encoding bulldog ant silk protein BAF1 (minus region encoding signal peptide).
SEQ ID NO:42 - Nucleotide sequence encoding bulldog ant silk protein BAF1.
SEQ ID NO:43 - Nucleotide sequence encoding bulldog ant silk protein BAF2 (minus region encoding signal peptide).
SEQ ID NO:44 - Nucleotide sequence encoding bulldog ant silk protein BAF2.
SEQ ID NO:45 - Nucleotide sequence encoding bulldog ant silk protein BAF3 (minus region encoding signal peptide).
SEQ ID NO:46 - Nucleotide sequence encoding bulldog ant silk protein BAF3.
SEQ ID NO:47 - Nucleotide sequence encoding bulldog ant silk protein BAF4 (minus region encoding signal peptide).
SEQ ID NO:48 - Nucleotide sequence encoding bulldog ant silk protein BAF4.
SEQ ID NO:49 - Weaver ant silk protein termed herein GAF1 (minus signal peptide).
SEQ ID NO:50 - Weaver ant silk protein termed herein GAF1.

SEQ ID NO:51 - Weaver ant silk protein termed herein GAF2 (minus signal peptide).

SEQ ID NO:52 - Weaver ant silk protein termed herein GAF2.

SEQ ID NO:53 - Weaver ant silk protein termed herein GAF3 (minus signal peptide).

SEQ ID NO:54 - Weaver ant silk protein termed herein GAF3.

SEQ ID NO:55 - Weaver ant silk protein termed herein GAF4 (minus signal peptide).

SEQ ID NO:56 - Weaver ant silk protein termed herein GAF4.

SEQ ID NO:57 - Nucleotide sequence encoding weaver ant silk protein GAF1 (minus region encoding signal peptide).

SEQ ID NO:58 - Nucleotide sequence encoding weaver ant silk protein GAF1.

SEQ ID NO:59 - Nucleotide sequence encoding weaver ant silk protein GAF2 (minus region encoding signal peptide).

SEQ ID NO:60 - Nucleotide sequence encoding weaver ant silk protein GAF2.

SEQ ID NO:61 - Nucleotide sequence encoding weaver ant silk protein GAF3 (minus region encoding signal peptide).

SEQ ID NO:62 - Nucleotide sequence encoding weaver ant silk protein GAF3.

SEQ ID NO:63 - Nucleotide sequence encoding weaver ant silk protein GAF4 (minus region encoding signal peptide).

SEQ ID NO:64 - Nucleotide sequence encoding weaver ant silk protein GAF4.

SEQ ID NO:65 - Hornet silk protein termed herein Vssilk3 (minus signal peptide).

SEQ ID NO:66 - Hornet silk protein termed herein Vssilk3.

SEQ ID NO:67 - Hornet silk protein termed herein Vssilk4 (minus signal peptide).

SEQ ID NO:68 - Hornet silk protein termed herein Vssilk4.

SEQ ID NO:69 - Hornet silk protein termed herein Vssilk2 (minus signal peptide).

SEQ ID NO:70 - Hornet silk protein termed herein Vssilk2.

SEQ ID NO:71 - Hornet silk protein termed herein Vssilk1 (minus signal peptide).

SEQ ID NO:72 - Hornet silk protein termed herein Vssilk1.

SEQ ID NO:73 - Nucleotide sequence encoding hornet silk protein Vssilk3 (minus region encoding signal peptide).

SEQ ID NO:74 - Nucleotide sequence encoding hornet silk protein Vssilk3.

SEQ ID NO:75 - Nucleotide sequence encoding hornet silk protein Vssilk4 (minus region encoding signal peptide).

SEQ ID NO:76 - Nucleotide sequence encoding hornet silk protein Vssilk4.

SEQ ID NO:77 - Nucleotide sequence encoding hornet silk protein Vssilk2 (minus region encoding signal peptide).

SEQ ID NO:78 - Nucleotide sequence encoding hornet silk protein Vssilk2.

SEQ ID NO:79 - Nucleotide sequence encoding hornet silk protein Vssilk1 (minus region encoding signal peptide).

SEQ ID NO:80 - Nucleotide sequence encoding hornet silk protein Vssilk1.

SEQ ID NO:81 – Asiatic honeybee silk protein termed silk protein 1 (also termed ABS1) (minus signal peptide).

SEQ ID NO:82 - Asiatic honeybee silk protein termed silk protein 1 (also termed ABS1).

SEQ ID NO:83 - Asiatic honeybee silk protein termed silk protein 2 (also termed ABS2) (minus signal peptide).

SEQ ID NO:84 - Asiatic honeybee silk protein termed silk protein 2 (also termed ABS2).

SEQ ID NO:85 - Asiatic honeybee silk protein termed silk protein 3 (also termed ABS3) (minus signal peptide).

SEQ ID NO:86 - Asiatic honeybee silk protein termed silk protein 3 (also termed ABS3).

SEQ ID NO:87 - Asiatic honeybee silk protein termed silk protein 4 (also termed ABS4) (minus signal peptide).

SEQ ID NO:88 - Asiatic honeybee silk protein termed silk protein 4 (also termed ABS4).

SEQ ID NO:89 - Nucleotide sequence encoding asiatic honeybee silk protein ABS1 (minus region encoding signal peptide).

SEQ ID NO:90 - Nucleotide sequence encoding asiatic honeybee silk protein ABS1.

SEQ ID NO:91 - Nucleotide sequence encoding asiatic honeybee silk protein ABS2 (minus region encoding signal peptide).

SEQ ID NO:92 - Nucleotide sequence encoding asiatic honeybee silk protein ABS2.

SEQ ID NO:93 - Nucleotide sequence encoding asiatic honeybee silk protein ABS3 (minus region encoding signal peptide).

SEQ ID NO:94 - Nucleotide sequence encoding asiatic honeybee silk protein ABS3.

SEQ ID NO:95 - Nucleotide sequence encoding asiatic honeybee silk protein ABS4 (minus region encoding signal peptide).

SEQ ID NO:96 - Nucleotide sequence encoding asiatic honeybee silk protein ABS4.

SEQ ID NO:97 - Lacewing silk protein termed herein MalF1 (minus signal peptide).

SEQ ID NO:98 - Lacewing silk protein termed herein MalF1.

SEQ ID NO:99 - Nucleotide sequence encoding lacewing silk protein MalF1 (minus region encoding signal peptide).

SEQ ID NO:100 - Nucleotide sequence encoding lacewing silk protein MalF1.

SEQ ID NOs 101 to 108 - Oligonucleotide primers.

## DETAILED DESCRIPTION OF THE INVENTION

General Techniques and Definitions

Unless specifically defined otherwise, all technical and scientific terms used herein shall be taken to have the same meaning as commonly understood by one of ordinary skill in the art (e.g., in cell culture, molecular genetics, silk processing, immunology, immunohistochemistry, protein chemistry, and biochemistry).

Unless otherwise indicated, the recombinant protein, cell culture, and immunological techniques utilized in the present invention are standard procedures, well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, A Practical Guide to Molecular Cloning, John Wiley and Sons (1984), J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical Approach, Volumes 1 and 2, IRL Press (1991), D.M. Glover and B.D. Hames (editors), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 and 1996), and F.M. Ausubel et al. (editors), Current Protocols in Molecular Biology, Greene Pub. Associates and Wiley-Interscience (1988, including all updates until present), Ed Harlow and David Lane (editors) Antibodies: A Laboratory Manual, Cold Spring Harbour Laboratory, (1988), and J.E. Coligan *et al.* (editors) Current Protocols in Immunology, John Wiley & Sons (including all updates until present).

The term "and/or", e.g., "X and/or Y" shall be understood to mean either "X and Y" or "X or Y" and shall be taken to provide explicit support for both meanings or for either meaning.

As used herein, the terms "silk protein" and "silk polypeptide" refer to a fibrous protein/polypeptide that can be used to produce a silk fibre, and/or a fibrous protein complex.

As used herein, the term "one or more silk proteins" refers to the process possibly using two or more different types of silk proteins such as a first silk protein, second silk protein, etc, as defined herein. Thus, in this context one silk protein means a population of identical silk protein molecules sufficient to produce silk dope.

As used herein, the term "capable of forming a tertiary structure which comprises a coiled-coil structure" refers to the ability of the proteins to form said structures under suitable conditions. For example, when processed to produce silk fibres the proteins form said structures. Furthermore, this term does not mean that the

entire protein is capable of forming a coiled-coil structure, just a portion thereof. In an embodiment, about 45% to about 90%, more preferably about 55% to about 70%, and even more preferably about 60% to about 66%, of the silk protein is capable of forming a tertiary structure which comprises a coiled-coil structure.

As used herein, the term "silk dope" refers to an aqueous solution comprising silk proteins. Preferably, the silk dope comprises at least 0.05% w/v, more preferably at least 0.1% w/v, and even more preferably at least 0.5% w/v, of a silk protein as defined herein. In an embodiment, silk dope produced by a method of the invention comprises about 0.5% to about 15% (wt%) silk protein. However, if the further step of increasing the concentration of silk proteins in the silk dope is not performed the more typical yield is about 0.5% to about 4% (wt%) silk protein. Silk dope produced using a method of the invention is amenable to extrusion for the formation of a fibre and/or film casting.

As used herein, a "silk fibre" refers to filaments comprising silk proteins which can be woven into various items such as textiles.

As used herein, the term "reducing the amount of surfactant solution", or variations thereof including reducing the amount of ionic liquid, means that the total amount of surfactant or ionic liquid is decreased. In an embodiment, the concentration of surfactant or ionic liquid following the reduction is less than about 10mM, more preferably less than 5mM and even more preferably less than 1 mM, prior to any step involving further concentration of the solution.

As used herein, a "dehydrating solution" is any solution, preferably an aqueous solution, that has a lower water concentration in solution than the silk dope that is to be concentrated.

As used herein, the term "solubilising" when referring to the silk proteins being contacted with a surfactant or ionic liquid means that the surfactant or ionic liquid associates with the silk proteins and maintains them in solution by preventing their aggregation. This is contrast to silk proteins seen in the cells, especially in the inclusion bodies.

The term "signal peptide", "N-terminal signal sequence" and variations thereof refers to an amino terminal protein/peptide preceding a secreted mature protein. The signal peptide is cleaved from and is therefore not present in the mature protein. Signal peptides have the function of directing and *trans*-locating secreted proteins across cell membranes. The signal peptide is also referred to as signal sequence, and are well known in the art.

Coiled-Coil Silk Proteins

The terms "polypeptide" and "protein" are generally used interchangeably and refer to a single polypeptide chain which may or may not be modified by addition of non-amino acid groups. The terms "proteins" and "polypeptides" as used herein also include variants, mutants, modifications, analogous and/or derivatives of the silk proteins described herein. In a preferred embodiment, a silk protein used in the invention is only comprised of naturally occurring amino acids.

The % identity of a protein is determined by GAP (Needleman and Wunsch, 1970) analysis (GCG program) with a gap creation penalty=5, and a gap extension penalty=0.3. The query sequence is at least 15 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 15 amino acids. More preferably, the query sequence is at least 50 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 50 amino acids. More preferably, the query sequence is at least 100 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 100 amino acids. Even more preferably, the query sequence is at least 250 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 250 amino acids. Even more preferably, the GAP analysis aligns the two sequences over their entire length.

As used herein a "biologically active" fragment is a portion of a protein of the invention which maintains a defined activity of the full-length protein, namely the ability to be used to produce silk. Biologically active fragments can be any size as long as they maintain the defined activity.

The term "consisting essentially of", or variations thereof, means that the defined amino acid sequence may have a few, such as one, two, three or four, additional amino acids compared to that defined. For example, when absent from the defined sequence an N-terminal methionine may be added. The term "consists of", or variations thereof, means that the defined sequence does not have additional or less amino acids when compared to the defined sequence, particularly at the N- and C-termini.

With regard to a defined protein, it will be appreciated that % identity figures higher than those provided above will encompass preferred embodiments. Thus, where applicable, in light of the minimum % identity figures, it is preferred that the protein comprises an amino acid sequence which is at least 40%, more preferably at least 45%, more preferably at least 50%, more preferably at least 55%, more preferably at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably

at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, more preferably at least 99%, more preferably at least 99.1%, more preferably at least 99.2%, more preferably at least 99.3%, more preferably at least 99.4%, more preferably at least 99.5%, more preferably at least 99.6%, more preferably at least 99.7%, more preferably at least 99.8%, and even more preferably at least 99.9% identical to the relevant nominated SEQ ID NO.

Amino acid sequence mutants of the naturally occurring silk proteins described herein can be prepared by introducing appropriate nucleotide changes into a nucleic acid encoding the silk protein, or by *in vitro* synthesis of the desired protein. Such mutants include, for example, deletions, insertions or substitutions of residues within the amino acid sequence. A combination of deletion, insertion and substitution can be made to arrive at the final construct, provided that the final protein product possesses the desired characteristics.

Mutant (altered) proteins can be prepared using any technique known in the art. For example, a polynucleotide of the invention can be subjected to *in vitro* mutagenesis. Such *in vitro* mutagenesis techniques include sub-cloning the polynucleotide into a suitable vector, transforming the vector into a "mutator" strain such as the *E. coli* XL-1 red (Stratagene) and propagating the transformed bacteria for a suitable number of generations. In another example, the polynucleotides of the invention are subjected to DNA shuffling techniques as broadly described by Harayama (1998). These DNA shuffling techniques may include genes of the invention possibly in addition to genes related to those of the present invention, such as silk genes from Hymenopteran or Neuroptean species other than the specific species characterized herein. Products derived from mutated/altered DNA can readily be screened using techniques described herein to determine if they can be used as silk proteins.

In designing amino acid sequence mutants, the location of the mutation site and the nature of the mutation will depend on characteristic(s) to be modified. The sites for mutation can be modified individually or in series, e.g., by (1) substituting first with conservative amino acid choices and then with more radical selections depending upon the results achieved, (2) deleting the target residue, or (3) inserting other residues adjacent to the located site.

Amino acid sequence deletions or insertions generally range from about 1 to 15 residues, more preferably about 1 to 10 residues and typically about 1 to 5 contiguous residues.

Substitution mutants have at least one amino acid residue in the protein molecule removed and a different residue inserted in its place. The sites of greatest interest for substitutional mutagenesis include sites identified as important for function. Other sites of interest are those in which particular residues obtained from various strains or species are identical. These positions may be important for biological activity. These sites, especially those falling within a sequence of at least three other identically conserved sites, are preferably substituted in a relatively conservative manner. Such conservative substitutions are shown in Table 1 under the heading of "exemplary substitutions".

**Table 1.** Exemplary substitutions

| **Original Residue** | **Exemplary Substitutions** |
|---|---|
| Ala (A) | val; leu; ile; gly; cys; ser; thr |
| Arg (R) | lys |
| Asn (N) | gln; his |
| Asp (D) | glu |
| Cys (C) | ser; thr; ala; gly; val |
| Gln (Q) | asn; his |
| Glu (E) | asp |
| Gly (G) | pro; ala; ser; val; thr; cyc |
| His (H) | asn; gln |
| Ile (I) | leu; val; ala; met |
| Leu (L) | ile; val; met; ala; phe |
| Lys (K) | arg |
| Met (M) | leu; phe |
| Phe (F) | leu; val; ala |
| Pro (P) | gly |
| Ser (S) | thr; ala; gly; val; gln; cys |
| Thr (T) | ser; gln; ala; cys |
| Trp (W) | tyr |
| Tyr (Y) | trp; phe |
| Val (V) | ile; leu; met; phe; ala; ser; thr |

Coiled-coil structures of silk proteins are characterized by heptad repeats represented by the consensus sequence $(abcdefg)_n$. In a preferred embodiment, the portion of the protein that has a coiled-coil structure comprises at least 10 copies of the heptad sequence *abcdefg*, and at least 25% of the amino acids at positions *a* and *d* are alanine residues.

In a preferred embodiment, the protein that has a coiled-coil structure comprises at least 12 consecutive copies, more preferably at least 15 consecutive copies, and even more preferably at least 18 consecutive copies of the heptad. In further embodiments, the protein that has a coiled-coil structure can have up to at least 28 copies of the heptad. Typically, the copies of the heptad will be tandemly repeated. However, they do not necessarily have to be perfect tandem repeats, for example, as shown in Figures 5 and 6 of WO 2007/038837 a few amino acids may be found between two heptads, or a few truncated heptads may be found (see, for example, Xenospiral in Figure 5 of WO 2007/038837).

Guidance regarding amino acid substitutions which can be made to the silk proteins which have a coiled-coil structure is provided in Figures 5 and 6, as well as Tables 6 to 10, of WO 2007/038837. Where a predicted useful amino acid substitution based on the experimental data provided herein is in anyway in conflict with the exemplary substitutions provided in Table 1 of WO 2007/038837 it is preferred that a substitution based on the experimental data is used.

Coiled-coil structures of the silk proteins have a high content of alanine residues, particularly at amino acid positions *a*, *d* and *e* of the heptad. However, positions *b*, *c*, *f* and *g* also have a high frequency of alanine residues. In a preferred embodiment, at least 15% of the amino acids at positions *a*, *d* and/or *e* of the heptads are alanine residues, more preferably at least 25%, more preferably at least 30%, more preferably at least 40%, and even more preferably at least 50%. In a further preferred embodiment, at least 25% of the amino acids at both positions *a* and *d* of the heptads are alanine residues, more preferably at least 30%, more preferably at least 40%, and even more preferably at least 50%. Furthermore, it is preferred that at least 15% of the amino acids at positions *b*, *c*, *f* and *g* of the heptads are alanine residues, more preferably at least 20%, and even more preferably at least 25%.

Typically, the heptads will not comprise any proline or histidine residues. Furthermore, the heptads will comprise few (1 or 2), if any, phenylalanine, methionine, tyrosine, cysteine, glycine or tryptophan residues. Apart from alanine, common (for example greater than 5%, more preferably greater than 10%) amino acids in the heptads include leucine (particularly at positions *b* and *d*), serine (particularly at positions *b*, *e*

and *f*), glutamic acid (particularly at positions *c*, *e* and *f*), lysine (particularly at positions *b*, *c*, *d*, *f* and *g*) as well as arginine at position *g*.

In a preferred embodiment, the heptads are determined by using the pattern recognition program MARCOIL (Delorenzi and Speed, 2002).

Proteins (and polynucleotides) useful for the methods of the invention can be purified (isolated) from a wide variety of Hymenopteran and Neuropteran species. Examples of Hymenopterans include, but are not limited to, any species of the Suborder Apocrita (bees, ants and wasps), which include the following Families of insects; Chrysididae (cuckoo wasps), Formicidae (ants), Mutillidae (velvet ants), Pompilidae (spider wasps), Scoliidae, Vespidae (paper wasps, potter wasps, hornets), Agaonidae (fig wasps), Chalcididae (chalcidids), Eucharitidae (eucharitids), Eupelmidae (eupelmids), Pteromalidae (pteromalids), Evaniidae (ensign wasps), Braconidae, Ichneumonidae (ichneumons), Megachilidae, Apidae, Colletidae, Halictidae, and Melittidae (oil collecting bees). Examples of Neuropterans include species from the following insect Families: Mantispidae, Chrysopidae (lacewings), Myrmeleontidae (antlions), and Ascalaphidae (owlflies). Such further proteins (and polynucleotides) can be characterized using the same procedures described herein for silks from *Bombus terrestris*, *Myrmecia forficata*, *Oecophylla smaragdina* and *Mallada signata*.

Also included within the scope of the invention are proteins which are differentially modified during or after synthesis, e.g., by biotinylation, benzylation, glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. These modifications may serve to increase the stability and/or bioactivity of the protein.

Polynucleotides

The term "polynucleotide" is used interchangeably herein with the term "nucleic acid".

The term "exogenous" in the context of a polynucleotide refers to the polynucleotide when present in a cell, or in a cell-free expression system, in an altered amount compared to its native state. In one embodiment, the cell is a cell that does not naturally comprise the polynucleotide. However, the cell may be a cell which comprises a non-endogenous polynucleotide resulting in an altered, preferably increased, amount of production of the encoded protein. An exogenous polynucleotide of the invention includes polynucleotides which have not been separated from other

components of the transgenic (recombinant) cell, or cell-free expression system, in which it is present, and polynucleotides produced in such cells or cell-free systems which are subsequently purified away from at least some other components.

The % identity of a polynucleotide is determined by GAP (Needleman and Wunsch, 1970) analysis (GCG program) with a gap creation penalty=5, and a gap extension penalty=0.3. Unless stated otherwise, the query sequence is at least 45 nucleotides in length, and the GAP analysis aligns the two sequences over a region of at least 45 nucleotides. Preferably, the query sequence is at least 150 nucleotides in length, and the GAP analysis aligns the two sequences over a region of at least 150 nucleotides. More preferably, the query sequence is at least 300 nucleotides in length and the GAP analysis aligns the two sequences over a region of at least 300 nucleotides. Even more preferably, the GAP analysis aligns the two sequences over their entire length.

With regard to the defined polynucleotides, it will be appreciated that % identity figures higher than those provided above will encompass preferred embodiments. Thus, where applicable, in light of the minimum % identity figures, it is preferred that a polynucleotide of the invention comprises a sequence which is at least 40%, more preferably at least 45%, more preferably at least 50%, more preferably at least 55%, more preferably at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, more preferably at least 99%, more preferably at least 99.1%, more preferably at least 99.2%, more preferably at least 99.3%, more preferably at least 99.4%, more preferably at least 99.5%, more preferably at least 99.6%, more preferably at least 99.7%, more preferably at least 99.8%, and even more preferably at least 99.9% identical to the relevant nominated SEQ ID NO.

In an embodiment, a polynucleotide which encodes a silk protein useful for the invention comprises, more preferably consists essentially of, even more preferably consists of, a sequence selected from:

a) a nucleotide sequence as provided in any one of SEQ ID NOs 9 to 16, 25 to 32, 41 to 48, 57 to 64, 73 to 80, 89 to 96, 99 or 100,

b) a nucleotide sequence which is at least 30% identical to any one or more of SEQ ID NOs 9 to 16, 25 to 32, 41 to 48, 57 to 64, 73 to 80, 89 to 96, 99 or 100, and

c) a biologically active fragment encoding portion of a) or b).

When it is preferred that as little as possible of the silk proteins is secreted from the cell, the encoded silk proteins do not comprise an N-terminal signal sequence. Examples of polynucleotides encoding such silk proteins include those comprising, more preferably consisting essentially of, even more preferably consisting of, a sequence selected from:

a) a nucleotide sequence as provided in any one of SEQ ID NOs 9, 11, 13, 15, 25, 27, 29, 31, 41, 43, 45, 47, 57, 59, 61, 63, 73, 75, 77, 79, 89, 91, 93, 95 or 97,

b) a nucleotide sequence which is at least 30% identical to any one or more of SEQ ID NOs 9, 11, 13, 15, 25, 27, 29, 31, 41, 43, 45, 47, 57, 59, 61, 63, 73, 75, 77, 79, 89, 91, 93, 95 or 97, and

c) a biologically active fragment of a) or b).

Other embodiments of the invention rely on the expression of silk proteins with an N-terminal signal sequence, and/or the co-production (in the same or different cells) of a first silk protein, second silk protein, third silk protein and/or fourth silk protein as defined herein. Based on the sequence information provided in the Sequence Listing, the skilled person could readily identifying representative polynucleotides for expression for each embodiment of the invention.

Polynucleotides for use in the methods of the present invention may possess, when compared to naturally occurring molecules, one or more mutations which are deletions, insertions, or substitutions of nucleotide residues. Mutants can be either naturally occurring (that is to say, isolated from a natural source) or synthetic (for example, by performing site-directed mutagenesis on the nucleic acid).

Polynucleotides for use in the invention can also hybridize to a silk protein encoding nucleotide sequence as provided herein, such as one or more of SEQ ID NOs 9 to 16, 25 to 32, 41 to 48, 57 to 64, 73 to 80, 89 to 96, 99 and 100, under stringent conditions. The term "stringent hybridization conditions" and the like as used herein refers to parameters with which the art is familiar, including the variation of the hybridization temperature with length of an oligonucleotide. Nucleic acid hybridization parameters may be found in references which compile such methods, Sambrook, et al. (*supra*), and Ausubel, et al. (*supra*). For example, stringent hybridization conditions, as used herein, can refer to hybridization at 65°C in hybridization buffer (3.5xSSC, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, 0.02% Bovine Serum Albumin (BSA), 2.5 mM $NaH_2PO_4$ (pH7), 0.5% SDS, 2 mM EDTA), followed by one or more washes in 0.2.xSSC, 0.01% BSA at 50°C.

Nucleic Acid Constructs

Cells for use in the methods of the invention will typically comprise a nucleic acid construct(s) encoding the silk protein(s). The construct may be integrated into the genome of the cell, or be extrachromosal such as be a recombinant vector. Such a vector contains heterologous polynucleotide sequences, that is polynucleotide sequences that are not naturally found adjacent to the polynucleotide molecule encoding the silk protein, and that preferably are derived from a species other than the species from which the polynucleotide molecule(s) are derived. The vector can be either RNA or DNA, either prokaryotic or eukaryotic, and typically is a transposon (such as described in US 5,792,294), a virus or a plasmid.

One type of recombinant vector comprises a polynucleotide molecule encoding the silk protein operatively linked to an expression vector. The phrase operatively linked refers to insertion of a polynucleotide molecule into an expression vector in a manner such that the molecule is able to be expressed when transformed into a host cell. As used herein, an expression vector is a DNA or RNA vector that is capable of transforming a host cell and of effecting expression of a specified polynucleotide molecule. Preferably, the expression vector is also capable of replicating within the host cell. Expression vectors can be either prokaryotic or eukaryotic, and are typically viruses or plasmids. Expression vectors of the present invention include any vectors that function (i.e., direct gene expression) in recombinant cells of the present invention, including in bacterial, fungal, endoparasite, arthropod, insect, animal, and plant cells. Particularly preferred expression vectors of the present invention can direct gene expression in plants cells. Vectors of the invention can also be used to produce the protein in a cell-free expression system, such systems are well known in the art.

In particular, the nucleic acid construct contains regulatory sequences such as transcription control sequences, translation control sequences, origins of replication, and other regulatory sequences that are compatible with the recombinant cell and that control the expression of the polynucleotide molecules. Transcription control sequences are sequences which control the initiation, elongation, and termination of transcription. Particularly important transcription control sequences are those which control transcription initiation, such as promoter, enhancer, operator and repressor sequences. Suitable transcription control sequences include any transcription control sequence that can function in at least one of the recombinant cells of the present invention. A variety of such transcription control sequences are known to those skilled in the art. Preferred transcription control sequences include those which function in bacterial, yeast, arthropod, plant or mammalian cells, such as, but not limited to, tac,

lac, trp, trc, oxy-pro, omp/lpp, rrnB, bacteriophage lambda, bacteriophage T7, T7lac, bacteriophage T3, bacteriophage SP6, bacteriophage SP01, metallothionein, alpha-mating factor, Pichia alcohol oxidase, alphavirus subgenomic promoters (such as Sindbis virus subgenomic promoters), antibiotic resistance gene, baculovirus, *Heliothis zea* insect virus, vaccinia virus, herpesvirus, raccoon poxvirus, other poxvirus, adenovirus, cytomegalovirus (such as intermediate early promoters), simian virus 40, retrovirus, actin, retroviral long terminal repeat, Rous sarcoma virus, heat shock, phosphate and nitrate transcription control sequences as well as other sequences capable of controlling gene expression in prokaryotic or eukaryotic cells.

As outlined above, an aspect of the invention relies on the silk protein being secreted from the cell, typically due to the presence of an N-termnal signal sequence. Examples of suitable signal segments include, but are not limited to, tissue plasminogen activator (t-PA), interferon, interleukin, growth hormone, viral envelope glycoprotein signal segments, *Nicotiana nectarin* signal peptide (US 5,939,288), tobacco extensin signal, the soy oleosin oil body binding protein signal, *Arabidopsis thaliana* vacuolar basic chitinase signal peptide, as well as native signal sequences of the silk polypeptides defined herein.

Cells

Most of the methods of the invention rely on the use of cells producing the one or more silk proteins as defined herein. Transformation of a polynucleotide molecule into a cell can be accomplished by any method by which a polynucleotide molecule can be inserted into the cell. Transformation techniques include, but are not limited to, transfection, electroporation, microinjection, lipofection, adsorption, and protoplast fusion. A recombinant cell may remain unicellular or may grow into a tissue, organ or a multicellular organism. Transformed polynucleotide molecules can remain extrachromosomal or can integrate into one or more sites within a chromosome of the transformed (i.e., recombinant) cell in such a manner that their ability to be expressed is retained.

Suitable host cells to transform include any cell that can be transformed with a polynucleotide encoding a silk polypeptide as defined herein. Host cells either can be endogenously (i.e., naturally) capable of producing the silk polypeptides or can be capable of producing such polypeptides after being transformed with at least one polynucleotide molecule as defined herein. Host cells can be any cell capable of producing at least one silk protein as defined herein, and include bacterial, fungal (including yeast), parasite, arthropod, animal and plant cells. Examples of host cells

include *Salmonella*, *Escherichia*, *Bacillus*, *Listeria*, *Saccharomyces*, *Spodoptera*, *Mycobacteria*, *Trichoplusia*, BHK (baby hamster kidney) cells, MDCK cells, CRFK cells, CV-1 cells, COS (e.g., COS-7) cells, and Vero cells. Further examples of host cells are *E. coli*, including *E. coli* K-12 derivatives; *Salmonella typhi*; *Salmonella typhimurium*, including attenuated strains; *Spodoptera frugiperda*; *Trichoplusia ni*; and non-tumorigenic mouse myoblast G8 cells (e.g., ATCC CRL 1246). Additional appropriate mammalian cell hosts include other kidney cell lines, other fibroblast cell lines (e.g., human, murine or chicken embryo fibroblast cell lines), myeloma cell lines, Chinese hamster ovary cells, mouse NIH/3T3 cells, LMTK cells and/or HeLa cells. Particularly preferred host cells are bacterial cells.

The skilled person can readily determine suitable culture conditions such as media, temperature and time for a particular cell type. For example, in an embodiment the cells are *Escherichia coli* cultered at about 30°C to about 37°C for a period of about 24h to about 48h.

Recombinant DNA technologies can be used to improve expression of a transformed polynucleotide molecule by manipulating, for example, the number of copies of the polynucleotide molecule within a host cell, the efficiency with which those polynucleotide molecules are transcribed, the efficiency with which the resultant transcripts are translated, and the efficiency of post-translational modifications. Recombinant techniques useful for increasing the expression of polynucleotide molecules include, but are not limited to, operatively linking polynucleotide molecules to high-copy number plasmids, integration of the polynucleotide molecule into one or more host cell chromosomes, addition of vector stability sequences to plasmids, substitutions or modifications of transcription control signals (e.g., promoters, operators, enhancers), substitutions or modifications of translational control signals (e.g., ribosome binding sites, Shine-Dalgarno sequences), modification of polynucleotide molecules to correspond to the codon usage of the host cell, and the deletion of sequences that destabilize transcripts.

Production of Silk Dope

The present invention relates to methods of producing silk dope which can then be used for a wide variety of applications.

One step of an aspect of the invention relates to lysing cells to liberate silk proteins produced and contained within the cells. This step can be performed by any means known in the art. For example, the cell suspension is typically centrifuged to pellet the cells and the cells resuspended into a more concentrated solution ready for

lysis. Cells can be lysed, for example, by passages through a French Press, homogenized using a Polytron (Brinkman Instruments) or sonicated on ice. Alternate methods of lysing cells, such as bacterial cells, are well known to those of skill in the art (see, e.g., Sambrook et al., *supra*). Various kits are available for cell lysis and are well known in the art, for example the Bugbuster kit (Novagen) and the ProteaPrep kit (Protea Biosciences, Inc.).

The present inventors have identified that silk proteins as defined herein expressed in bacteria form insoluble aggregates ("inclusion bodies"). In a preferred embodiment, the method includes the isolation of these inclusion bodies. Several protocols are suitable for purification of protein inclusion bodies. For example, purification of inclusion bodies typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells by the methods discussed above. In an embodiment, the cells are lysed, the cell membranes solubilised, and the insoluble fraction comprising the inclusion bodies is isolated for further processing.

An aspect of the invention relies on increasing the concentration of silk proteins from the supernatant. Again, this can be achieved by any method known in the art. In one embodiment, this is achieved by contacting the supernatant with an agent which precipitates the silk proteins such as, but not limited to, ammonium sulfate, trichloroacetic acid, perchloric acid and acetone, or commercial precipitant cocktails such as PlusOne (Amersham Biosciences), or Perfect-Focus (Geno Technology Inc.).

An optional step of the invention for producing silk dope, but nonetheless preferred in cases where the yield of silk protein is not sufficiently high, comprises increasing the concentration of silk proteins in the silk dope. Again, the can be achieved by any method known in the art for increasing the concentration of a protein an aqueous solution. In a particularly useful embodiment, the silk dope is concentrated by dialysing against a dehydrating solution such as a solution comprising a hygroscopic polymer. Examples suitable hygroscopic polymers include, but are not limited to, polyethylene glycol (PEG), amylase, and sericin, or a combination of two or more thereof. PEG molecules are available in a range of molecular sizes and the selection of the PEG will be determined by the membrane chosen for dialysis and the rate of concentration required. Preferably, the PEG is of a molecular weight of about 8,000 to about 10,000 g/mol and has a concentration of about 25% to about 50%.

*Surfactants*

In one embodiment, a step of producing silk dope as defined herein involves the use of a surfactant. The present inventors were surprised to find that surfactants, such as SDS, enables the silk proteins as defined herein to stay in solution whilst allowing the formation of a coiled-coil structure when the concentration of the surfactant is lowered.

In an embodiment, the surfactant is an anionic surfactant. Examples of anionic surfactants useful for the invention include, but are not limited to, sodium dodecyl sulfate (SDS), ammonium lauryl sulfate and other alkyl sulfate salts, sodium 1-octanesulfonate monohydrate, sodium lauroyl sarcosinate, sodium lauryl ether sulfate (SLES), sodium taurodeoxycholate hydrate, and alkyl benzene sulfonate; or a combination of two or more thereof. In a preferred embodiment, the anionic surfactant is SDS.

Any concentration of the surfactant can be used which increases the solubility of the silk proteins can be used. For example, at least about 0.1% v/v of the surfactant is used. In an embodiment, about 0.1% to about 10% v/v, more preferably, about 0.5% to about 2% v/v or about 0.5% to about 5% v/v, of the surfactant is used.

A further step of the methods of the invention for producing silk dope comprises reducing the amount of surfactant in solution by adding a compound which precipitates the surfactant to assist in the correct folding of the silk proteins. Any compound may be used which associates which, and reduces the solubility of, the surfactant. Examples include, but are not limited to, a salt or a carbohydrate such as α-cyclodextrin; or a combination of two or more thereof. Preferably, the salt is a potassium salt or a sodium salt. Preferably, the potassium salt is potassium chloride and the sodium salt is sodium acetate. Any concentration of the compound can be used which results in a reduction in the amount of surfactant in solution. For example, the compound is added to a final concentration of about 1 mM to about 1 M, more preferably about 40 mM to about 100 mM, or about 40 mM to 400 mM.

A further step of the methods of the invention for producing silk dope comprises separating the solution from the precipitate formed following the addition of the compound. This can be achieved by any method known in the art such as using centrifugation, for example at 16000g for 5 minutes, and removing the supernatant (solution) comprising (which is) the silk dope. Preferably, after this step the silk proteins constitute at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 95%, more preferably at

least 97%, more preferably at least 99%, and even more preferably 100% of the protein in solution.

*Ionic Liquids*

Generally, ionic liquids can be defined as compounds that are comprised entirely of ions and are liquids at temperatures of less than about 100°C, preferably less than about 85°C. As used in the present invention, ionic liquids generally comprise one or more anions and one or more cations. In preferred embodiments, the ionic liquids comprise organic cations created by derivatizing one or more compounds to include substituents, such as alkyl, alkenyl, alkynyl, alkoxy, alkenoxy, alkynoxy, a variety of aromatics, such as (substituted or unsubstituted) phenyl, (substituted or unsubstituted) benzyl, (substituted or unsubstituted) phenoxy, and (substituted or unsubstituted) benzoxy, and a variety of heterocyclic aromatics having one, two, or three heteroatoms in the ring portion thereof, said heterocyclics being substituted or unsubstituted. The derivatized compounds include, but are not limited to, imidazoles, pyrazoles, thiazoles, isothiazoles, azathiozoles, oxothiazoles, oxazines, oxazolines, oxazaboroles, dithiozoles, triazoles, delenozoles, oxaphospholes, pyrroles, boroles, furans, thiophenes, phospholes, pentazoles, indoles, indolines, oxazoles, isoxazoles, isotetrazoles, tetrazoles, benzofurans, dibenzofurans, benzothiophenes, dibenzothiophenes, thiadiazoles, pyridines, pyrimidines, pyrazines, pyridazines, piperazines, piperidines, morpholones, pyrans, annolines, phthalazines, quinazolines, guanidiniums, quinxalines, choline-based analogues, and combinations thereof. The basic cation structure can be singly or multiply substituted or unsubstituted.

The anionic portion of the ionic liquid can comprise an inorganic moiety, an organic moiety, or combinations thereof. In preferred embodiments, the anionic portion comprises one or more moieties selected from halogens, phosphates, alkylphosphates, alkenylphosphates, bis(trifluoromethylsulfonyl)imide ($NTf_2$), $BF_4^-$, $PF_6^-$, $AsF_6^-$, $NO_3^-$, $N(CN)_2^-$, $N(SO_3CF_3)_2^-$, amino acids, substituted or unsubstituted carboranes, perchlorates, pseudohalogens such as thiocyanate and cyanate, metal chloride-based Lewis acids (e.g., zinc chlorides and aluminum chlorides), or $C_{1\text{-}6}$ carboxylates. Pseudohalides are monovalent and have properties similar to those of halides. Examples of pseudohalides useful according to the invention include cyanides, thiocyanates, cyanates, fulminates, and azides. Exemplary carboxylates that contain 1-6 carbon atoms are formate, acetate, propionate, butyrate, hexanoate, maleate, fumarate, oxalate, lactate, pyruvate and the like.

A variety of ionic liquids can be prepared and used according to the present invention. In particular, any combination of the cations and anions noted above could be used. It is only necessary to combine one or more cations (such as those described above) with one or more anions (such as those described above) to form a material that is liquid under the conditions described herein. For example, a cation imidazolium moiety could be combined with an anionic halogen moiety to form a material that is liquid under the requisite conditions (e.g., 1-butyl-3-methyl-imidazolium chloride) and that is formed substantially completely of ionic moieties. Thus, it is clear that the present invention encompasses the use of a great diversity of ionic liquids. Specific, non-limiting examples of ionic liquids for use according to the invention include 1-butyl-3-methyl-imidazolium chloride ("BmimCl"); 1-allyl-3-methyl-imidazolium chloride ("AmimCl"); 1-ethyl-3-methyl-imidazolium chloride; 1-hydrogen-3-methyl-imidazolium chloride; 1-benzyl-3-methyl-imidazolium chloride ("BenzylmimCl"); 1-isopropyl-3-methyl-imidazolium chloride; 1-m-methoxybenzyl-3-methyl-imidazolium chloride ("MethoxyBenzylmimCl"); 1-m-methylbenzyl-3-methyl-imidazolium chloride ("MethylBenzylmimCl"); 1-benzyl-3-methyl-imidazolium chloride, and 1-methyl-3-benzyl-imidazolium dicyanamide ("BenzylmimDca").

The invention also encompasses the use of various mixtures of ionic liquids. In fact, ionic liquid mixtures can be useful for providing ionic liquids having customized properties, such as viscosity. For example, BenzylmimCl is a relatively viscous ionic liquid; however, it viscosity can be significantly reduced by mixing with AmimCl. The viscosity of the ionic liquid mixture can thus be adjusted by varying the ratio between the more viscous component and the less viscous component.

Ionic liquids for use according to the invention can be synthesized according to the literature. Preferably, the ionic liquids are dried (e.g., at 100°C) in a vacuum oven over a period of time, such as about 48 hours, prior to use. In one embodiment, the ionic liquid is formed of a material that is solid (e.g., crystalline) at ambient conditions but is liquid at increased temperature (such as greater than about 30°C, greater than about 50°C, greater than about 75°C., or greater than about 100°C). Generally, the crystalline material can be placed in an appropriate container and heated to dissolution (see, for example, Ionic Liquids in Synthesis, Wasserscheid, P. and Weldon, T. (Eds.), Wiley Pub.). Of course, the ionic liquid can also comprise a material that is liquid at ambient conditions (e.g., at a temperature around 20-25°C).

*Filtration and/or Chromatography*

The solubilised silk proteins may be concentrated and separated from impurities, such as the surfactant, ionic liquid and other cellular components based on charge, hydrophilicity, affinity, solubility or stability, or size. Non-limiting examples of separation techniques include ammonium sulfate precipitation, chromatography, and membrane-filtration (including tangential flow membrane filtration). In embodiments utilizing chromatography for separation, exemplary methods include ion-exchange (cationic or anionic), affinity chromatography, hydrophilic-interaction, hydrophobic-interaction, size-exclusion and gel-permeation (see US 6,248,570).

In some embodiments, the separation can be conducted by membrane-filtration, which includes, but is not limited to, methods such as single pass, dead-end, direct flow filtration (DFF), and crossflow or tangential flow filtration (TFF). According to the invention, filtration is based on the principle of separating molecules according to size using a semi-permeable membrane of a defined range of pore sizes. It is known to those skilled in the art that combinations of filtration methods and membrane types may be used in separation.

According to the invention, membrane-filtration is the separation of cellular components effected by polymeric or inorganic membranes. Within the art, there are four commonly accepted categories of membranes defined by the size of the material they remove from the carrier liquid. Methods of sequentially filtering through membranes from the smallest to largest pore size are Reverse Osmosis (RO), Nanofiltration (NF), Ultrafiltration (UF), and Microfiltration (MF).

Filtration with the above-mentioned membranes separates molecules according to their molecular weight by using membranes with specific pore sizes. For example, separation with RO membranes that have pore sizes less than 0.001 micrometers is intended to separate molecules that have a molecular weight less than 200 Daltons. Filtration with NF membranes that have pore sizes from 0.001-0.008 micrometers, inclusive, is intended to separate molecules that have a molecular weight from 200 Daltons to 15 kilodaltons (kDa) inclusive. Filtration with UF membranes that have pore sizes from 0.005-0.1 micrometers, inclusive, is intended to separate molecules that have a molecular weight from 5 kDa-300 kDa, inclusive. Filtration with microfiltration membranes that have pore sizes from 0.05-3.0 micrometers, inclusive, is intended to separate molecules that have a molecular weight from 100 kDa-3000 kDa and larger.

According to this invention, membrane-filtration can separate the solubilised silk proteins from other components based on size exclusion by utilizing membranes that have a particular Molecular Weight Cut-Off (MQWCO) that is determined by the

pore size of the membrane. The MWCO, also called Nominal Molecular Weight Limit (NMWL) or Nominal Molecular Weight Cut-Off (NMWCO), is the kilodalton size designation for the filtration by membranes. The MWCO is defined as the molecular weight of the molecule that is 90% retained by the membrane. Because, for example, molecules of the same molecular weight can have significantly different shapes, the MWCO is not an exact metric, but is nevertheless a useful metric and is commonly employed by filter manufacturers. Both hydrophobic as well as hydrophilic membranes may be used in the present invention. Such membranes may be used as flat sheets or in a spirally wound configuration. Hollow fibers may also be used. In relation to compositions of UF membranes, any number of potential membrane materials may be used including, but not limited to, regenerated cellulose, polyether sulfone (which may or may not be modified to alter its inherent hydrophobicity), polyvinylidene fluoride, and ceramic and metal oxide aggregates. Many polyether sulfone UF membranes can withstand a pH range of 0.5-13, and temperatures ranging up to 85°C. Materials for MF membranes include everything used for UF membranes, as well as polycarbonate, polypropylene, polyethylene and PTFE (TEFLON™).

In some embodiments, the solubilised silk proteins can be filtered for the separation of large cellular debris from smaller cellular components to prevent the cellular debris from interfering with the proceeding separation and purification steps that involve the use of membranes or chromatography. In these embodiments, the permeate comprises the silk proteins and is recovered.

In some embodiments, a membrane can be used in a separation step having a suitable MWCO. For example, typical silks proteins used in the methods of the invention have a MW of about 30kDa. In these embodiments, the retentate comprises the silk protein and can be recovered.

In a preferred embodiment, tangential flow filtration acts to both diafilter and concentrate the silk proteins. In TFF, typically, the solution flows parallel to the filter membrane. A pressure differential across the membrane causes fluid and filterable solutes (whose molecular weight is smaller than that of the membranes or behaves like so, such as globular proteins) to flow through the filter. In HPTFF (high performance tangential flow filtration) the membrane is charged, therefore using both size and charge of molecules to separate contaminants (see US 20030229212). According to the invention, diafiltration can be either discontinuous or continuous diafiltration. In discontinuous diafiltration, the solution is concentrated, and the lost volume is replaced by a new buffer. In continuous diafiltration, the solution volume is maintained by the inflow of new buffer solution while the old buffer solution is removed. In some

embodiments, the separation and purification of the silk proteins can be performed by tangential flow filtration methods using ultrafiltration membranes.

Uses

The silk dope produced using the methods of the invention can be used for a broad and diverse array of medical, military, industrial and commercial applications. For example, the silk dope is used to produce silk fibres which in turn can be used in the manufacture of medical devices such as sutures, skin grafts, cellular growth matrices, replacement ligaments, and surgical mesh, and in a wide range of industrial and commercial products, such as, for example, cable, rope, netting, fishing line, clothing fabric, bullet-proof vest lining, container fabric, backpacks, knapsacks, bag or purse straps, adhesive binding material, non-adhesive binding material, strapping material, tent fabric, tarpaulins, pool covers, vehicle covers, fencing material, sealant, construction material, weatherproofing material, flexible partition material, sports equipment; and, in fact, in nearly any use of fibre or fabric for which high tensile strength and elasticity are desired characteristics. The silk dope also have applications for use in the production of compositions for personal care products such as cosmetics, skin care, hair care and hair colouring; and in coating of particles, such as pigments.

The silks may be used in their native form or they may be modified to form derivatives, which provide a more beneficial effect. For example, the silks may be modified by conjugation to a polymer to reduce allergenicity as described in US 5,981,718 and US 5,856,451. Suitable modifying polymers include, but are not limited to, polyalkylene oxides, polyvinyl alcohol, poly-carboxylates, poly(vinylpyrolidone), and dextrans. In another example, the silks may be modified by selective digestion and splicing of other protein modifiers. For example, the silk proteins may be cleaved into smaller peptide units by treatment with acid at an elevated temperature of about 60°C. The useful acids include, but are not limited to, dilute hydrochloric, sulfuric or phosphoric acids. Alternatively, digestion of the silk proteins may be done by treatment with a base, such as sodium hydroxide, or enzymatic digestion using a suitable protease may be used.

The proteins may be further modified to provide performance characteristics that are beneficial in specific applications for personal care products. The modification of proteins for use in personal care products is well known in the art. For example, commonly used methods are described in US 6,303,752, US 6,284,246, and US 6,358,501. Examples of modifications include, but are not limited to, ethoxylation to promote water-oil emulsion enhancement, siloxylation to provide lipophilic

compatibility, and esterification to aid in compatibility with soap and detergent compositions. Additionally, the silk proteins may be derivatized with functional groups including, but not limited to, amines, oxiranes, cyanates, carboxylic acid esters, silicone copolyols, siloxane esters, quaternized amine aliphatics, urethanes, polyacrylamides, dicarboxylic acid esters, and halogenated esters. The silk proteins may also be derivatized by reaction with diimines and by the formation of metal salts.

Consistent with the above definitions of "polypeptide" (and "protein"), such derivatized and/or modified molecules are also referred to herein broadly as "polypeptides" and "proteins".

The silk dope can be spun together and/or bundled or braided with other fibre types. Examples include, but are not limited to, polymeric fibres (e.g., polypropylene, nylon, polyester), fibres and silks of other plant and animal sources (e.g., cotton, wool, *Bombyx mori* or spider silk), and glass fibres. A preferred embodiment is silk fibre braided with 10% polypropylene fibre. The present invention contemplates that the production of such combinations of fibres can be readily practiced to enhance any desired characteristics, e.g., appearance, softness, weight, durability, water-repellant properties, improved cost-of-manufacture, that may be generally sought in the manufacture and production of fibres for medical, industrial, or commercial applications.

*Personal Care Products*

Cosmetic and skin care compositions may be anhydrous compositions comprising an effective amount of silk in a cosmetically acceptable medium. The uses of these compositions include, but are not limited to, skin care, skin cleansing, make-up, and anti-wrinkle products. An effective amount of a silk for cosmetic and skin care compositions is herein defined as a proportion of from about $10^{-4}$ to about 30% by weight, but preferably from about $10^{-3}$ to 15% by weight, relative to the total weight of the composition. This proportion may vary as a function of the type of cosmetic or skin care composition. Suitable compositions for a cosmetically acceptable medium are described in US 6,280,747. For example, the cosmetically acceptable medium may contain a fatty substance in a proportion generally of from about 10 to about 90% by weight relative to the total weight of the composition, where the fatty phase containing at least one liquid, solid or semi-solid fatty substance. The fatty substance includes, but is not limited to, oils, waxes, gums, and so-called pasty fatty substances. Alternatively, the compositions may be in the form of a stable dispersion such as a water-in-oil or oil-in-water emulsion. Additionally, the compositions may contain one or more

conventional cosmetic or dermatological additives or adjuvants, including but not limited to, antioxidants, preserving agents, fillers, surfactants, UVA and/or UVB sunscreens, fragrances, thickeners, wetting agents and anionic, nonionic or amphoteric polymers, and dyes or pigments.

Emulsified cosmetics and quasi drugs which are producible with the use of emulsified materials comprising silk produced by a method of the invention, for example, cleansing cosmetics (beauty soap, facial wash, shampoo, rinse, and the like), hair care products (hair dye, hair cosmetics, and the like), basic cosmetics (general cream, emulsion, shaving cream, conditioner, cologne, shaving lotion, cosmetic oil, facial mask, and the like), make-up cosmetics (foundation, eyebrow pencil, eye cream, eye shadow, mascara, and the like), aromatic cosmetics (perfume and the like), tanning and sunscreen cosmetics (tanning and sunscreen cream, tanning and sunscreen lotion, tanning and sunscreen oil, and the like), nail cosmetics (nail cream and the like), eyeliner cosmetics (eyeliner and the like), lip cosmetics (lipstick, lip cream, and the like), oral care products (tooth paste and the like) bath cosmetics (bath products and the like), and the like.

The cosmetic composition may also be in the form of products for nail care, such as a nail varnish. Nail varnishes are herein defined as compositions for the treatment and colouring of nails, comprising an effective amount of silk in a cosmetically acceptable medium. An effective amount of a silk for use in a nail varnish composition is herein defined as a proportion of from about $10^{-4}$ to about 30% by weight relative to the total weight of the varnish. Components of a cosmetically acceptable medium for nail varnishes are described in US 6,280,747. The nail varnish typically contains a solvent and a film forming substance, such as cellulose derivatives, polyvinyl derivatives, acrylic polymers or copolymers, vinyl copolymers and polyester polymers. The composition may also contain an organic or inorganic pigment.

Hair care compositions are herein defined as compositions for the treatment of hair, including but not limited to shampoos, conditioners, lotions, aerosols, gels, and mousses, comprising an effective amount of silk in a cosmetically acceptable medium. An effective amount of a silk for use in a hair care composition is herein defined as a proportion of from about $10^{-2}$ to about 90% by weight relative to the total weight of the composition. Components of a cosmetically acceptable medium for hair care compositions are described in US 2004/0170590, US 6,280,747, US 6,139,851, and US 6,013,250. For example, these hair care compositions can be aqueous, alcoholic or aqueous-alcoholic solutions, the alcohol preferably being ethanol or isopropanol, in a proportion of from about 1 to about 75% by weight relative to the total weight, for the

aqueous-alcoholic solutions. Additionally, the hair care compositions may contain one or more conventional cosmetic or dermatological additives or adjuvants, as given above.

Hair colouring compositions are herein defined as compositions for the colouring, dyeing, or bleaching of hair, comprising an effective amount of silk in a cosmetically acceptable medium. An effective amount of a silk for use in a hair colouring composition is herein defined as a proportion of from about $10^{-4}$ to about 60% by weight relative to the total weight of the composition. Components of a cosmetically acceptable medium for hair colouring compositions are described in US 2004/0170590, US 6,398,821 and US 6,129,770. For example, hair colouring compositions generally contain a mixture of inorganic peroxygen-based dye oxidizing agent and an oxidizable coloring agent. The peroxygen-based dye oxidizing agent is most commonly hydrogen peroxide. The oxidative hair coloring agents are formed by oxidative coupling of primary intermediates (for example p-phenylenediamines, p-aminophenols, p-diaminopyridines, hydroxyindoles, aminoindoles, aminothymidines, or cyanophenols) with secondary intermediates (for example phenols, resorcinols, m-aminophenols, m-phenylenediamines, naphthols, pyrazolones, hydroxyindoles, catechols or pyrazoles). Additionally, hair colouring compositions may contain oxidizing acids, sequestrants, stabilizers, thickeners, buffers carriers, surfactants, solvents, antioxidants, polymers, non-oxidative dyes and conditioners.

The silks can also be used to coat pigments and cosmetic particles in order to improve dispersibility of the particles for use in cosmetics and coating compositions. Cosmetic particles are herein defined as particulate materials such as pigments or inert particles that are used in cosmetic compositions. Suitable pigments and cosmetic particles, include, but are not limited to, inorganic color pigments, organic pigments, and inert particles. The inorganic color pigments include, but are not limited to, titanium dioxide, zinc oxide, and oxides of iron, magnesium, cobalt, and aluminium. Organic pigments include, but are not limited to, D&C Red No. 36, D&C Orange No. 17, the calcium lakes of D&C Red Nos. 7, 11, 31 and 34, the barium lake of D&C Red No. 12, the strontium lake D&C Red No. 13, the aluminium lake of FD&C Yellow No. 5 and carbon black particles. Inert particles include, but are not limited to, calcium carbonate, aluminium silicate, calcium silicate, magnesium silicate, mica, talc, barium sulfate, calcium sulfate, powdered Nylon™, perfluorinated alkanes, and other inert plastics.

The silks may also be used in dental floss (see, for example, US 2005/0161058). The floss may be monofilament yarn or multifilament yarn, and the fibres may or may

not be twisted. The dental floss may be packaged as individual pieces or in a roll with a cutter for cutting pieces to any desired length. The dental floss may be provided in a variety of shapes other than filaments, such as but not limited to, strips and sheets and the like. The floss may be coated with different materials, such as but not limited to, wax, polytetrafluoroethylene monofilament yarn for floss.

The silks may also be used in soap (see, for example, US 2005/0130857).

*Pigment and Cosmetic Particle Coating*

The effective amount of a silk for use in pigment and cosmetic particle coating is herein defined as a proportion of from about $10^{-4}$ to about 50%, but preferably from about 0.25 to about 15% by weight relative to the dry weight of particle. The optimum amount of the silk to be used depends on the type of pigment or cosmetic particle being coated. For example, the amount of silk used with inorganic color pigments is preferably between about 0.01% and 20% by weight. In the case of organic pigments, the preferred amount of silk is between about 1% to about 15% by weight, while for inert particles, the preferred amount is between about 0.25% to about 3% by weight. Methods for the preparation of coated pigments and particles are described in US 5,643,672. These methods include: adding an aqueous solution of the silk to the particles while tumbling or mixing, forming a slurry of the silk and the particles and drying, spray drying a solution of the silk onto the particles or lyophilizing a slurry of the silk and the particles. These coated pigments and cosmetic particles may be used in cosmetic formulations, paints, inks and the like.

*Biomedical*

The silks may be used as a coating on a bandage to promote wound healing. For this application, the bandage material is coated with an effective amount of the silk. For the purpose of a wound-healing bandage, an effective amount of silk is herein defined as a proportion of from about $10^{-4}$ to about 30% by weight relative to the weight of the bandage material. The material to be coated may be any soft, biologically inert, porous cloth or fibre. Examples include, but are not limited to, cotton, silk, rayon, acetate, acrylic, polyethylene, polyester, and combinations thereof. The coating of the cloth or fibre may be accomplished by a number of methods known in the art. For example, the material to be coated may be dipped into an aqueous solution containing the silk. Alternatively, the solution containing the silk may be sprayed onto the surface of the material to be coated using a spray gun. Additionally, the solution containing the silk may be coated onto the surface using a roller coat printing process.

The wound bandage may include other additives including, but not limited to, disinfectants such as iodine, potassium iodide, povidon iodine, acrinol, hydrogen peroxide, benzalkonium chloride, and chlorohexidine; cure accelerating agents such as allantoin, dibucaine hydrochloride, and chlorophenylamine malate; vasoconstrictor agents such as naphazoline hydrochloride; astringent agents such as zinc oxide; and crust generating agents such as boric acid.

The silk dope may also be used in the form of a film as a wound dressing material. The use of silk, in the form of an amorphous film, as a wound dressing material is described in US 6,175,053. The amorphous film comprises a dense and nonporous film of a crystallinity below 10% which contains an effective amount of silk. For a film for wound care, an effective amount of silk is herein defined as between about 1 to 99% by weight. The film may also contain other components including but not limited to other proteins such as sericin, and disinfectants, cure accelerating agents, vasoconstrictor agents, astringent agents, and crust generating agents, as described above. Other proteins such as sericin may comprise 1 to 99% by weight of the composition. The amount of the other ingredients listed is preferably below a total of about 30% by weight, more preferably between about 0.5 to 20% by weight of the composition. The wound dressing film may be prepared by dissolving the above mentioned materials in an aqueous solution, removing insolubles by filtration or centrifugation, and casting the solution on a smooth solid surface such as an acrylic plate, followed by drying.

The silk dope may also be used to produce sutures (see, for example, US 2005/0055051). Such sutures can feature a braided jacket made of ultrahigh molecular weight fibres and silk fibres. The polyethylene provides strength. Polyester fibres may be woven with the high molecular weight polyethylene to provide improved tie down properties. The silk may be provided in a contrasting color to provide a trace for improved suture recognition and identification. Silk also is more tissue compliant than other fibres, allowing the ends to be cut close to the knot without concern for deleterious interaction between the ends of the suture and surrounding tissue. Handling properties of the high strength suture also can be enhanced using various materials to coat the suture. The suture advantageously has the strength of Ethibond No. 5 suture, yet has the diameter, feel and tie-ability of No. 2 suture. As a result, the suture is ideal for most orthopedic procedures such as rotator cuff repair, Achilles tendon repair, patellar tendon repair, ACL/PCL reconstruction, hip and shoulder reconstruction procedures, and replacement for suture used in or with suture anchors. The suture can be uncoated, or coated with wax (beeswax, petroleum wax, polyethylene wax, or

others), silicone (Dow Corning silicone fluid 202A or others), silicone rubbers, PBA (polybutylate acid), ethyl cellulose (Filodel) or other coatings, to improve lubricity of the braid, knot security, or abrasion resistance, for example.

The silk dope may also be used to produce stents (see, for example, US 2004/0199241). For example, a stent graft is provided that includes an endoluminal stent and a graft, wherein the stent graft includes silk. The silk induces a response in a host who receives the stent graft, where the response can lead to enhanced adhesion between the silk stent graft and the host's tissue that is adjacent to the silk of the silk stent graft. The silk may be attached to the graft by any of various means, e.g., by interweaving the silk into the graft or by adhering the silk to the graft (e.g., by means of an adhesive or by means of suture). The silk may be in the form of a thread, a braid, a sheet, powder, etc. As for the location of the silk on the stent graft, the silk may be attached only the exterior of the stent, and/or the silk may be attached to distal regions of the stent graft, in order to assist in securing those distal regions to neighbouring tissue in the host. A wide variety of stent grafts may be utilized within the context of the present invention, depending on the site and nature of treatment desired. Stent grafts may be, for example, bifurcated or tube grafts, cylindrical or tapered, self-expandable or balloon-expandable, unibody or, modular, etc.

In addition to silk, the stent graft may contain a coating on some or all of the silk, where the coating degrades upon insertion of the stent graft into a host, the coating thereby delaying contact between the silk and the host. Suitable coatings include, without limitation, gelatin, degradable polyesters (e.g., PLGA, PLA, MePEG-PLGA, PLGA-PEG-PLGA, and copolymers and blends thereof), cellulose and cellulose derivatives (e.g., hydroxypropyl cellulose), polysaccharides (e.g., hyaluronic acid, dextran, dextran sulfate, chitosan), lipids, fatty acids, sugar esters, nucleic acid esters, polyanhydrides, polyorthoesters and polyvinylalcohol (PVA). The silk-containing stent grafts may contain a biologically active agent (drug), where the agent is released from the stent graft and then induces an enhanced cellular response (e.g., cellular or extracellular matrix deposition) and/or fibrotic response in a host into which the stent graft has been inserted.

The silk dope may also be used to produce a matrix for producing ligaments and tendons *ex vivo* (see, for example, US 2005/0089552). A silk-fibre-based matrix can be seeded with pluripotent cells, such as bone marrow stromal cells (BMSCs). The bioengineered ligament or tendon is advantageously characterized by a cellular orientation and/or matrix crimp pattern in the direction of applied mechanical forces, and also by the production of ligament and tendon specific markers including collagen

type I, collagen type III, and fibronectin proteins along the axis of mechanical load produced by the mechanical forces or stimulation, if such forces are applied. In a preferred embodiment, the ligament or tendon is characterized by the presence of fibre bundles which are arranged into a helical organization. Some examples of ligaments or tendons that can be produced include anterior cruciate ligament, posterior cruciate ligament, rotator cuff tendons, medial collateral ligament of the elbow and knee, flexor tendons of the hand, lateral ligaments of the ankle and tendons and ligaments of the jaw or temporomandibular joint. Other tissues that may be produced by methods of the present invention include cartilage (both articular and meniscal), bone, muscle, skin and blood vessels.

The silk dope may also be used to produce hydrogels (see, for example, US 2005/0266992). Silk fibroin hydrogels can be characterized by an open pore structure which allows their use as tissue engineering scaffolds, substrate for cell culture, wound and burn dressing, soft tissue substitutes, bone filler, and as well as support for pharmaceutical or biologically active compounds.

The silk dope may also be used to produce dermatological compositions (see, for example, US 2005/0019297). Furthermore, the dope may also be used to produce sustained release compositions (see, for example, US 2004/0005363).

*Textiles*

The silk dope may also be used to produce a coating for the surface of fibres for subsequent use in textiles. This provides a monolayer of the protein film on the fibre, resulting in a smooth finish. US 6,416,558 and US 5,232,611 describe the addition of a finishing coat to fibres. The methods described in these disclosures provide examples of the versatility of finishing the fibre to provide a good feel and a smooth surface. For this application, the fibre is coated with an effective amount of the silk. For the purpose of fibre coating for use in textiles, an effective amount of silk is herein defined as a proportion of from about 1 to about 99% by weight relative to the weight of the fibre material. The fibre materials include, but are not limited to textile fibres of cotton, polyesters such as rayon and Lycra™, nylon, wool, and other natural fibres including native silk. Compositions suitable for applying the silk onto the fibre may include co-solvents such as ethanol, isopropanol, hexafluoranols, isothiocyanouranates, and other polar solvents that can be mixed with water to form solutions or microemulsions. The silk containing solution may be sprayed onto the fibre or the fibre may be dipped into the solution. While not necessary, flash drying of the coated material is preferred. An alternative protocol is to apply the silk composition onto

woven fibres. An ideal embodiment of this application is the use of silks to coat stretchable weaves such as used for stockings.

*Composite Materials*

Silk fibres can be added to polyurethane, other resins or thermoplastic fillers to prepare panel boards and other construction material or as moulded furniture and benchtops that replace wood and particle board. The composites can be also be used in building and automotive construction especially rooftops and door panels. The silk fibres re-enforce the resin making the material much stronger and allowing lighterweight construction which is of equal or superior strength to other particle boards and composite materials. Silk fibres may be isolated and added to a synthetic composite-forming resin or be used in combination with plant-derived proteins, starch and oils to produce a biologically-based composite materials. Processes for the production of such materials are described in JP 2004284246, US 2005175825, US 4,515,737, JP 47020312 and WO 2005/017004.

*Paper Additives*

The fibre properties of the silk can add strength and quality texture to paper making. Silk papers are made by mottling silk threads in cotton pulp to prepare extra smooth handmade papers is used for gift wrapping, notebook covers, carry bags. Processes for production of paper products from silk dope are generally described in JP 2000139755.

*Advanced Materials*

Silks produced from silk dope of the invention have considerable toughness and stands out among other silks in maintaining these properties when wet (Hepburn et al., 1979).

Areas of substantial growth in the clothing textile industry are the technical and intelligent textiles. There is a rising demand for healthy, high value functional, environmentally friendly and personalized textile products. Fibres, such as those of the invention, that do not change properties when wet and in particular maintain their strength and extensibility are useful for functional clothing for sports and leisure wear as well as work wear and protective clothing.

Developments in the weapons and surveillance technologies are prompting innovations in individual protection equipments and battle-field related systems and structures. Besides conventional requirements such as material durability to prolonged

exposure, heavy wear and protection from external environment, silk textiles produced from silk dope of the invention can be processed to resist ballistic projectiles, fire and chemicals. Processes for the production of such materials are described in WO 2005/045122 and US 2005268443.

**EXAMPLES**

**EXAMPLE 1 - Recombinant production and purification of honeybee silk proteins**

To create recombinant expression constructs, the four honeybee silk gene sequences (Genbank Accession Nos: FJ235088, FJ235089, FJ235090, FJ235091) without signal peptides were amplified by PCR from the cDNA clones described in Sutherland et al. (2006) using the following oligonucleotide primer sets:

AmelF1: GGAATT C**TC ATG A***GT TTG GAG GGG CCG GGC AAC TCG* (SEQ ID NO:101) and CGGC **GGATCC** TTA TTA *AAA TAC GTT GCT CTT CAA GT* (SEQ ID NO:102);

AmelF2: GGAATT C**TC ATG A**GC *CGC GTG ATT AAT CAC GAG TCC CTG* (SEQ ID NO:103) and CGGC **GGATCC** TTA TTA *TTC CAA CTT TGC TAC ATG TAT TTT C* (SEQ ID NO:104);

AmelF3: GGAATT C**CC ATG *G***GC *GTC GAG GAA TTC AAG TCC TCG* (SEQ ID NO:105) and CGGC **AGATCT** TTA TTA *AAA TTT TTT ATC CTC AAT A* (SEQ ID NO:106);

AmelF4: GGAATT C**CC ATG *G***CA *AGG GAA GAG GTG GAG ACA CGG* (SEQ ID NO:107) and CGGC **GGATCC** TTA TTA *CTT CAC CTC CCA TTC TTC ATT C* (SEQ ID NO:108) (cloning restriction enzyme sites are underlined and in bold and sequences that match the cDNA sequence are shown in italics).

The PCR amplicons were cloned into restriction enzyme sites (AmelF1 and AmelF2: BspH1 and Bam HI ; AmelF3: Nco1 and Bgl II; AmelF4: Nco1 and Bam HI) of the pET14b expression vector (Novagen) and the sequences verified by DNA sequencing before expression.

The constructs were transformed into Rosetta 2 (DE3) competent cells (Novagen) and the silk proteins were initially expressed in 50 mL overnight express instant TB medium (Novagen) in shake flasks. The four honeybee silk proteins, AmelF1-4, were synthesized in *E. coli* cells in the soluble form at 20°C and insoluble form at 30°C and 37°C. The highest yields of protein, as judged by comparative protein band intensity after SDS-PAGE (Figure 1), were obtained when expression was conducted for extended periods (24-36 h) at temperatures ≥30°C with the proteins

recovered from the inclusion bodies. Quantitative gel band intensity analysis, with protein identities confirmed by mass spectroscopy, indicated that protein recovered from the inclusion bodies was essentially pure (>95%) silk protein. Subsequent analysis found that proteins solubilised from inclusion bodies self assembled into native-like structure. Thus, all subsequent expression was conducted under conditions such that recombinant proteins were recovered from the inclusion bodies.

In order to increase protein yield a large-scale batch fed fermentation process was developed and optimised for AmelF3. Fermentations were carried out in 2-litre Biostat B culture vessels (Sartorius Stedim, Melsungen, Germany) using minimal medium (starting volume 1.6 litres). Glucose was used as initial carbon source, switching to a glycerol feed following induction of silk protein expression with IPTG. The initial medium contained (per litre): $KH_2PO_4$, 13.3 g; $(NH_4)_2HPO_4$, 4 g and citric acid 1.7 g. The pH of the medium was adjusted to a final value of 7.0 using 2 M NaOH. The following components were sterilized separately, then added (per litre of final medium): 40 ml of 50% (w/v) glucose; 5 ml of 1M $MgSO_4$; 130 µl of 0.1M thiamine hydrochloride; 1 ml of 100 mg/ml ampicillin and 5 ml of a vitamin/trace metal solution containing (per litre of solution): biotin, 0.2 g; $CuSO_4.5H_2O$, 2.0 g; NaI, 0.08 g; $MnSO_4.H_2O$, 3.0 g; $Na_2MoO_4.2H_2O$, 0.2g; boric acid 0.02 g; $CoCl_2.6H_2O$, 0.5 g; $ZnCl_2$, 7.0 g; $FeSO_4.7H_2O$, 22.0 g; $CaSO_4.2H_2O$, 0.5 g and $H_2SO_4$, 1 ml.

The fermentation inocula for the four strains were cultured for 20 h at 37 °C in the same medium as used in the fermenter. Once the inocula were transferred to the fermentation vessels, the pH of the medium was adjusted and controlled at 7.0 through the addition of 10% (w/v) $NH_4OH$ or 10% (w/v) $H_3PO_4$. The temperature was controlled at 37 °C, and dissolved oxygen (DO) concentration was maintained above 40% air saturation by manipulating the agitation speed up to 1100 rpm and enriching the air supply with pure oxygen when required.

When the cultures had grown to an optical density at 600 nm ($OD_{600nm}$) value of approximately 20 (~10 h after inoculation), IPTG was added to a final concentration of 1 mM. The fermenters were operated in batch mode until all of the glucose was consumed, as indicated by a sharp rise in DO (~12 h after inoculation). In the glycerol fed-batch phase, 400 ml of 62% (v/v) glycerol solution was fed into the fermenter at a rate of 50 ml/h. Cultures were grown for 24 h, after which the cells were harvested by centrifugation and stored at -80 °C. Under these conditions the $OD_{600nm}$ value of the ferment was 34 and the yield of purified recombinant AmelF3 after solubilisation was approximately 2.5 grams per litre of ferment. The same fermentation conditions were used to express the other honeybee silk proteins. The strains expressing silk proteins

AmelF1, 2 and 4 grew to $OD_{600nm}$ values of 30, 67, and 57 respectively. The yields of purified recombinant proteins AmelF1, 2 and 4 after solubilisation were approximately 0.2, 1.5 and 1.9 grams per litre of ferment respectively.

The yield of 2.5 g/L purified protein from the optimised batch fed fermentation system is by far the highest reported expression levels for any recombinant silk protein. Factors contributing to this high yield include the size and nature of the honeybee fibroin genes and the structural properties of the silk proteins. In contrast to the large size (>10kbp) and highly repetitive nature of the genes encoding the well-studied dragline silk of spiders and cocoon silk of silkworms, the honeybee silk genes are small (approximately 1kbp) with far less repetition in their DNA sequences (Sutherland et al., 2006). The smaller size and reduced level of repetition means that the honeybee genes are not prone to the genetic instabilities including premature translation termination and truncation that result from transgenic expression of highly repetitive nucleotide sequences.

The silk proteins in inclusion bodies were purified from the *E. coli* cells after repeatedly treating with BugBuster Master Mix (Novagen), according to manufacturer's protocol for soluble or inclusion body preparation. Protein solutions were analyzed by SDS polyacrylamide gel electrophoresis (SDS-PAGE) with 4-12% gradient (Invitrogen). Recombinant silk protein identification was verified by tandem mass spectrometry as previously described (Sutherland et al., 2006).

Silk proteins in inclusion bodies were solubilised in 3% sodium dodecyl sulfate (SDS) with 2 h incubation at 60°C. Protein concentration in solution was measured using a QuantiPro BCA assay kit (Sigma). Where required, solutions of the each of the four recombinant honeybee silk proteins were mixed at equimolar ratios. Excess SDS was removed from protein solutions by dialysis against 5g/L KCl solution causing KDS precipitation. The precipitate was removed by centrifugation at 16000*g* for 5 min.

**EXAMPLE 2 - Fourier transform infrared spectroscopy (FTIR)**

Fourier transform infrared spectroscopy was used to compare the protein structure of native and recombinant honeybee silks. Native honeybee silk sheets were obtained from a commercial hive, washed extensively in chloroform to remove wax and washed extensively in warm water to remove water-soluble contaminants. Solutions of each of the four recombinant honeybee silk proteins were mixed at equimolar ratios, cast and dried into a film. Infrared spectra from these samples were obtained in transmission mode using a Perkin-Elmer System 2000 Fourier transform spectrometer fitted with an i-series imaging infrared microscope accessory. Spectra

were collected using Spectrum software (version 5.3.1) and represent the average of 256 scans collected at a resolution of $4cm^{-1}$. Post-collection data manipulation and analysis was carried out using Grams/AI software v5.05. The deconvolution of the amide I region for each silk spectrum is shown in Figure 2. A summary of the results and the component secondary structure assignments is presented in Table 2.

**Table 2.** FTIR curve fitting summary.

| **Native honeybee silk** | | **Recombinant silk** | | |
|---|---|---|---|---|
| **Frequency ($cm^{-1}$)** | **Amide I area (%)** | **Frequency ($cm^{-1}$)** | **Amide I area (%)** | **Assignment** |
| 1717 | | 1710 | | COOH |
| 1693 | | 1693 | | side chains |
| 1670 | 25 | 1673 | 30 | β-sheet and side chains |
| 1651 | 23 | 1655 | 26 | coiled-coil |
| 1642 | 13 | 1643 | 7 | coiled-coil |
| 1628 | 29 | 1632 | 26 | coiled-coil |
| 1604 | 10 | 1610 | 11 | β-sheet |

The FTIR results suggest that the native honeybee silk contains approximately 65% coiled-coil structure, which is consistent with previous sequence–based predictions (Sutherland et al., 2006). The recombinant silk spectrum is very similar to the native silk spectrum, and the recombinant silk is estimated to contain 59% coiled-coil structure.

**EXAMPLE 3 – Dry spinning**

The inclusion bodies were solubilised in 3% SDS to give soluble honeybee fibroin solutions of generally between 0.5 – 2 wt% protein and up to 3 wt% protein. The excess SDS was removed from silk protein solutions by KCl precipitation. The potassium precipitation removed up to 95%, such as 70-80%, of SDS (by weighing the precipitate) but <10% of protein (by measuring protein concentrations in solution). The silk solutions were concentrated by extended dialysis against 20 wt% polyethylene glycol (PEG, MW 8000, Sigma) or Slide-A-Lyzer concentrating solution (Pierce), until a honey-like viscosity was obtained (around 10-15 wt% protein). A droplet of

concentrated silk dope was suspended between the prongs of a pair of tweezers in air and the tweezers were opened to form a fine thread (Figure 3A). These single-drawn threads were stable in air but dissolved in water. Fibres were then submerged in a 90% methanol 10% water bath, drawn a second time to approximately 2x length, and air-dried Figure 3B). The double-drawn threads were not soluble in water. Single-drawn and double-drawn fibres were examined by a light microscope with polarizing lenses, and by a Zeiss EVO LS15 environmental scanning electron microscope.

Recombinant silk threads imaged by ESEM (not shown) were circular in cross-section and fairly uniform in diameter along their length. Single-drawn fibres had small bodies adhering to their surface that could be salt crystals, however double-drawn fibres had smooth surfaces. Polarized light microscopy showed that single-drawn fibres are not birefringent, but that double-drawn fibres are strongly birefringent (Figure 3).

Single-drawn and double-drawn fibres and recombinant silk films were analysed by wide-angle x-ray scattering at the SAXS/WAXS beamline of the Australian Synchrotron. A wavelength of 0.886 Å and camera length of 0.558 m provided a $q$-range of approximately 0.07 to 1.4 $Å^{-1}$, which was calibrated using a silver behenate standard. The WAXS patterns for the film and for single-drawn fibres were dominated by a strong signal from SDS crystals, but this was not detectable in the double-drawn fibres. The present inventors therefore calculate that double-drawn threads contain <0.1% of the SDS crystals per unit length found in the single-drawn threads. The protein scattering patterns from recombinant silk could not be analysed due to either the strong SDS diffraction limiting the sensitivity of the technique, or to low signal-to-noise ratio in the case of the very fine double-drawn fibres.

The strength and extensibility of recombinant honeybee silk threads were measured on an Instron Tensile Tester model 4501 at a rate of 2.5 mm/min. Tests were conducted in air at 21°C and 65% relative humidity. Prior to testing each fibre was placed across a 3mm slot in a plastic frame and fixed with epoxy glue. The gauge length ($L_0$) and diameter of each fibre were measured on an optical microscope. Table 3 compares the mechanical properties of recombinant silk fibres to the properties of native fibres drawn from the honeybee silk gland.

**Table 3.** Tensile properties of recombinant honeybee silk fibres compared to native fibres.

| | Diameter (μm) | Breaking stress (MPa) | Breaking strain (%) | True breaking stress (MPa) |
|---|---|---|---|---|
| Single-drawn fibres | 30 ± 5 | 15 ± 3 | 225 ± 10 | 50 ± 12 |
| Double-drawn fibres | 13 ± 7 | 150 ± 39 | 47 ± 26 | 213 ± 63 |
| Native fibres (Hepburn, 1979) | 9 | 132 | 204 | 400 |

**EXAMPLE 4 – Wet spinning**

Silk proteins were generally prepared as described in Example 1. Generally protein concentration after SDS solubilisation was around 3% silk protein. If protein solutions had lower concentration they were concentrated by extended dialysis against 20 wt% polyethylene glycol (PEG, MW 8000, Sigma) or Slide-A-Lyzer concentrating solution (Pierce), until solutions were 3 – 6% silk protein.

The concentrated protein solutions of either equimolar mixtures of AmelF1-4 or AmelF3 alone were extruded through 10cm 100μm capillary tubing at a rate of 10m/min into methanol solution (50-90% methanol) which caused a fine and continuous thread to form. The threads were dried in the air and examined by a light microscope with polarizing lenses. The threads showed significant birefringence indicating that the proteins within the threads were directionally aligned (Figure 4A). Air dried fibres were submerged in a 90% methanol 10% water bath and drawn a second time to approximately 2x length (Figure 4B) or 4x length (Figure 4C), and air-dried. The strength and extensibility of recombinant honeybee silk threads were measured on an Instron Tensile Tester model 4501 at a rate of 2.5 mm/min. Tests were conducted in air at 21°C and 65% relative humidity. Prior to testing each fibre was placed across a 3mm slot in a plastic frame and fixed with epoxy glue. The gauge length ($L_0$) and diameter of each fibre were measured on an optical microscope.

Tables 4 and 5 describes the mechanical properties of undrawn recombinant silk fibres. Drawing resulted in threads that were stronger and insoluble in water, and highly birefringent.

**Table 4.** Tensile properties of recombinant honeybee silk fibres after extrusion of concentrated silk protein dope into methanol with and without drawing.

| | Diameter (μm) | Breaking stress (MPa) | Breaking strain (%) |
|---|---|---|---|
| Undrawn fibres (equimolar mixture of 4 proteins) | 22-39 | 70-78 | 200-250 |
| Fibres drawn X2 length (equimolar mixture of 4 proteins) | 17-22 | 50-92 | 80-160 |
| Fibres drawn x4 length (equimolar mixture of 4 proteins) | 19-21 | 150-161 | 38-91 |
| Undrawn fibres (AmelF3) | 39-41 | 39-53 | 256-275 |
| Fibres drawn X2 length (AmelF3) | 32-42 | 68-80 | 88-211 |
| Fibres drawn X4 length (AmelF3) | 28-30 | 99-117 | 131-154 |

**Table 5**. Mechanical properties of honeybee silk fibers.

| Method of fabrication | Constituent proteins | Diameter (μm) | Breaking stress (MPa) | Breaking strain (%) | Toughness (MPa) |
|---|---|---|---|---|---|
| Extruded into 70% MeOH | AmelF1-4 | 31 ± 2 | 70 ± 4 | 190 ± 11 | 91 ± 7 |
| | AmelF3 | 45 ± 2 | 50 ± 3 | 243 ± 10 | 105 ± 6 |
| Extruded into 70% MeOH then drawn ≈100% in 90% MeOH | AmelF1-4 | 21 ± 1 | 133 ± 11 | 94 ± 11 | 85 ± 9 |
| | AmelF3 | 34 ± 2 | 97 ± 7 | 129 ± 15 | 97 ± 10 |
| Extruded into 70% MeOH then drawn ≈300% in 90% MeOH | AmelF1-4 | 17 ± 1 | 203 ± 10 | 51 ± 5 | 70 ± 8 |
| | AmelF3 | 23 ± 1 | 178 ± 20 | 68 ± 9 | 85 ± 18 |
| *Natural*[3] | *Native* | *9* | *132* | *204* | NG |

[1]Calculated as $(d_0/d_1)^2$ where $d_0$ and $d_1$ are the diameters of the initial and drawn fibres

**EXAMPLE 5 – Circular dichroism (CD)**

The AmelF3 honeybee silk protein was expressed into the inclusion bodies of *E. coli*. AmelF3 inclusion bodies were unfolded using equivalent dry weight of the detergent sodium dodecyl sulphate (SDS) to generate 2-4% monomeric protein solutions. Dynamic light scattering (DLS) measured the hydrodynamic diameter of particles in the protein-detergent solution diluted ten-fold in 100mM NaCl as 9.2 +/- 0.1 nm (peak containing 98.2% of particle volume). The diameter of SDS micelles in 3% SDS solutions without protein under the same experimental conditions was a single peak at 5.5 +/- 0.2 nm. No SDS micelles were detected in the SDS–protein solutions confirming that the majority of the SDS was bound to the protein.

Proteins were refolded by removing SDS using KCl. Potassium dodecyl sulfate has significantly lower solubility than SDS and precipitates out of solution where it can be removed by centrifugation. The protein solutions were dialysed against water to reduce salt levels then concentrated by dialysis against PEG8000, resulting in 3-4% protein and 0.2-0.4% SDS concentration. When AmelF3 solutions were diluted ten-fold in 100 mM NaCl (comparable to physiological salt levels) the particle diameter increased to 20.3 +/- 0.7 nm (peak containing 86.8% of particle volume), in agreement with the approximate particle diameter calculated for an AmelF3 coiled coil.

CD spectra of honeybee AmelF3 solutions (0.12%) held in 0.01 mm path length sandwich quartz cell (Nova Biotech, El Cajon, CA) were collected using a AVIV Model 410 spectrophotometer (AVIV Biomedical, Inc., Lakewood, NJ) with a temperature controller. All samples were scanned at 25°C) with a 1 nm bandwidth from 260 nm to 180 nm, and the results were averaged from four repeated experiments. The CD spectra of AmelF3 solutions showed strong spectral minima at 220 and 209 nm and a 220 nm/209 nm ratio of 1.02 supporting a coiled coil structure. A 220 nm/209 nm ratio of one or more is indicative of coiled coils whereas a ratio of less than 0.86 is indicative of isolated helices. DLS measurements indicated that after addition of SDS back to the AmelF3 solutions the hydrodynamic particle diameter was reduced to the size observed in the original monomeric solutions, confirming that the removal of most SDS is a prerequisite for the protein to fold into a native-like silk protein conformation. In contrast to the AmelF3 protein, His-tagged recombinant versions of the homologous protein from *Apis cerana* remained monomeric and predominantly random coil at comparable concentrations (Shi et al., 2008). This result shows that AmelF3 alone, when prepared in the presence of low levels of SDS, folds to adopt a native like silk molecular structure.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

The present application claims priority from US 61/237,156 filed 26 August 2009, and US 61/315,812 filed 19 March 2010, the entire contents of both of which are incorporated herein by reference.

All publications discussed and/or referenced herein are incorporated herein in their entirety.

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed before the priority date of each claim of this application.

## REFERENCES

Atkins (1967) J. Mol. Biol. 24: 139-41.

Delorenzi and Speed (2002) Bioinformatics 18:617-625.

Harayama (1998) Trends Biotech. 16: 76-82.

Hepburn et al. (1979) Insect Biochem. 9: 69-77.

Needleman and Wunsch (1970) J. Mol. Biol. 48: 443-453.

Sezutzu et al. (2007) Biosci. Biotechnol. Biochem. 71: 2725-34.

Shi et al. (2008) Biomaterials 29: 2820-8.

Sutherland et al. (2006) Genome Res 16: 1414-21.

Sutherland et al. (2007) Mol Biol Evol 24: 2424-32.

## CLAIMS

1. A method for producing silk dope comprising silk proteins which have a coiled coil tertiary structure, the method comprising

i) lysing cells producing one or more silk proteins capable of forming a coiled coil tertiary structure,

ii) solubilising the silk proteins by contacting them with an amount of anionic surfactant sufficient to solubilize the silk proteins, and

iii) concentrating the silk proteins while reducing the amount of the anionic surfactant to produce the silk dope, wherein the silk dope comprises the anionic surfactant and silk proteins which have a coiled coil tertiary structure.

2. The method of claim 1, wherein the silk proteins are concentrated by

a) reducing the amount of surfactant in solution by adding a compound which precipitates the surfactant, and

b) separating the solution comprising the silk proteins from the precipitate formed in step a) to produce the silk dope.

3. The method of claim 1, wherein the compound which precipitates the surfactant is a salt or a carbohydrate; or a combination of two or more thereof.

4. The method of claim 3, wherein the salt is a potassium salt or a sodium salt.

5. The method of claim 1, wherein the silk proteins are concentrated by filtration.

6. The method of claim 1 which further comprises increasing the concentration of silk proteins in the silk dope.

7. The method of claim 6 which comprises dialysing the silk dope against a dehydrating solution.

8. The method of claim 7, wherein the dehydrating solution comprises a hygroscopic polymer.

9. The method of claim 1, wherein the silk dope comprises at least about 0.5% w/v silk proteins.

10. The method of claim 9, wherein the silk dope comprises about 0.5% to about 15% w/v silk proteins.

11. The method of claim 1, wherein the cells are bacterial cells, yeast cells, insect cells, plant cells or animal cells, or a combination of two or more thereof.

12. The method of claim 11, wherein the cells are bacterial cells.

13. The method of claim 1, wherein step i) further comprises isolating inclusion bodies from the lysed cells.

14. The method of claim 1 which further comprises culturing the cells before step i).

15. The method of claim 1, wherein the portion of the silk protein that is capable of forming a tertiary structure which comprises a coiled-coil structure comprises at least 10 copies of the heptad sequence *abcdefg,* and wherein at least 25% of the amino acids at positions *a* and *d* in the at least 10 copies of the heptad are alanine residues.

16. The method of claim 1, wherein the anionic surfactant is sodium dodecyl sulfate (SDS), ammonium lauryl sulfate or other alkyl sulfate salts, sodium 1-octanesulfonate monohydrate, sodium lauroyl sarcosinate, sodium lauryl ether sulfate (SLES), sodium taurodeoxycholate hydrate, alkyl benzene sulfonate; or a combination of two or more thereof.

17. A method for producing silk dope comprising silk proteins which have a coiled coil tertiary structure, the method comprising

i) obtaining supernatant from cell cultures, or from a cell-free expression system, producing one or more silk proteins capable of forming a coiled coil tertiary structure,

ii) solubilising the silk proteins by contacting them with an amount of anionic surfactant

sufficient to solubilize the silk proteins, and

iii) concentrating the silk proteins while reducing the amount of the anionic surfactant to produce the silk dope, wherein the silk dope comprises the anionic surfactant and silk proteins which have a coiled coil tertiary structure.

18. A method for producing a silk fibre, the method comprising extruding and/or drawing silk dope produced by the method of claim 1.

19. A method for producing a silk film, the method comprising casting silk dope produced by the method of claim 1.

20. The method of claim 16, wherein the anionic surfactant is sodium dodecyl sulfate (SDS).

21. The method of claim 17, wherein the anionic surfactant is sodium dodecyl sulfate (SDS).