(12) STANDARD PATENT (11) Application No. **AU 2013239320 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
Silk polypeptides

(51) International Patent Classification(s)
C12N 15/12 (2006.01) ***C07K 14/435*** (2006.01)
A61K 38/17 (2006.01)

(21) Application No: **2013239320** (22) Date of Filing: **2013.03.26**

(87) WIPO No: **WO13/142901**

(30) Priority Data

(31) Number	(32) Date	(33) Country
61/615,745	**2012.03.26**	**US**

(43) Publication Date: **2013.10.03**
(44) Accepted Journal Date: **2018.11.01**

(71) Applicant(s)
Commonwealth Scientific and Industrial Research Organisation

(72) Inventor(s)
Walker, Andrew Allan;Weisman, Sarah;Sutherland, Tara

(74) Agent / Attorney
FB Rice Pty Ltd, L 14 90 Collins St, Melbourne, VIC, 3000, AU

(56) Related Art
WO 2013/120143 A1

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 October 2013 (03.10.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/142901 A1

(51) International Patent Classification:
C12N 15/12 (2006.01) *A61K 38/17* (2006.01)
C07K 14/435 (2006.01)

(21) International Application Number:
PCT/AU2013/000308

(22) International Filing Date:
26 March 2013 (26.03.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/615,745 26 March 2012 (26.03.2012) US

(71) Applicant: **COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION** [AU/AU]; Limestone Avenue, Campbell, Australian Capital Territory 2612 (AU).

(72) Inventors: **WALKER, Andrew Allan**; 13c Basman Street, Braddon, Australian Capital Territory 2612 (AU). **WEISMAN, Sarah**; 14 Carstensz Street, Griffith, Australian Capital Territory 2603 (AU). **SUTHERLAND, Tara**; 26 Burnside Street, Watson, Australian Capital Territory 2602 (AU).

(74) Agent: **FB RICE**; Level 14, 90 Collins Street, Melbourne, Victoria 3000 (AU).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: SILK POLYPEPTIDES

(57) Abstract: The present disclosure provides polypeptides which encode a silk polypeptide, wherein at least a portion of the polypeptide has a coiled coil structure comprising at least 10 copies of the heptad sequence *abcdefg*, and wherein at least 25% of the amino acids at position a are alanine residues and at least 20% of the amino acids at position d are aromatic residues. The present disclosure also provides nucleic acids encoding such polypeptides, and recombinant cells and/or organisms which synthesize the polypeptides. The polypeptides can be used for a variety of purposes such as in the production of personal care products, plastics, textiles, and biomedical products.

WO 2013/142901 A1

SILK POLYPEPTIDES

FIELD OF THE INVENTION

The present disclosure generally relates to polypeptides which are capable of forming a coiled coil structure and to nucleic acids encoding such polypeptides. Owing to their structure and origin, these polypeptides may be referred to as silk polypeptides. The present disclosure also generally relates to recombinant cells and/or organisms which synthesize the polypeptides described herein. The polypeptides 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

'Silk' has become a single, all encompassing description for an extremely wide range of biological materials (Sutherland et al., 2010) that are an ancient product in evolution. Silks are produced by a wide range of insects, for example by the larvae of insects where a cocoon is formed for protection during metamorphosis, by adult insects such as webspinners that spin silk to make a web-like pouch or gallery in which they live, by insects of the order Hymenoptera (which includes bees, wasps and ants) where silk is part of nest construction, and by large numbers of other arthropods, most notably the various arachnids such as spiders, where orb-webs used for catching prey are common (Sutherland et al., 2010). Despite the diversity of structures and distributions, characterised silks have certain features in common, notably, being semicrystalline materials, that is materials with regions of ordered molecular structure (crystallites) within an amorphous matrix and also, all show typically similar protein compositions, often rich in alanine, serine, and/or glycine (Sutherland et al., 2010). However, overall, very few silks have been characterised, with most research concentrating on the dragline silks of the orb-weaving spider *Nephila clavipes*, the European garden spider, *Araneus diadematus*, and the nursery web spider, *Euprosthenops australis*, and on the cocoon silk of the domesticated silkworm, *Bombyx mori*.

Female praying mantises are thought to produce a form of silk during the construction of oothecae, which are protective structures in which the female lays her eggs. Oothecae persist in the environment for many months, anchoring eggs to vegetation and protecting them from predators and extremes of weather (Britton et al., 1970). They are large structures, containing up to several hundred eggs surrounded by an extensive network of fibrous protein sheets and air pockets. Ootheca production represents a large investment of material and energy for a female mantis, and

beforehand her abdomen is swollen with large collaterial glands full of the necessary materials (Kenchington and Flower, 1969). The ootheca is produced as a foamy mass surrounding the eggs, which dries and hardens over a period of several hours.

Ootheca is a complex composite material. The structural components include proteins that are thought to fold into alpha-helical coiled coils (Rudall, 1956; Rattew, 1974; Bullough and Tulloch, 1990). These coiled coil proteins are thought to form birefringent crystallites 1-2 μm wide and 20-30 nm thick, which are embedded in layers of a non-birefringent protein material (Rattew, 1974). Also present are organic crystals of calcium citrate or calcium oxalate (Parker and Rudall, 1955; Rudall, 1956; Rudall, 1962), small molecules such as N-acyldopamines and beta-glucosides and enzymes (Yago et al., 1984; Yago et al., 1990). The ootheca of cockroaches, termites and mantises probably share an evolutionary history (Nalepa and Lenz, 2000), but mantis ootheca is different from either of these, since it hardens externally rather than internally, does not contain chitin, and its ordered protein structure is thought to comprise alpha-helical coiled coils rather than beta-sheets.

The ordered protein structure in mantis ootheca has been partially characterised using x-ray scattering, electron diffraction and electron microscopy. Rudall (Rudall, 1956; Rudall, 1962) proposed a model for *Mantis religiosa* ootheca in which two-stranded coiled coils are arranged in regular lattices. The supramolecular structure was found to form by self-assembly as the material dries in the few hours after ootheca production. Rattew (Rattew, 1974) further refined this model using electron microscopy and x-ray scattering crystallites from *Sphodromantis* ootheca. Rattew's model features two-stranded coiled coils 74 nm long, staggered by approximately one-third of their length from neighbours on either side. Bullough and Tulloch (Bullough and Tulloch, 1990) studied structural proteins from *Tenodera australasiae* ootheca using electron diffraction, producing an electron density projection to 0.43 nm consistent with most of the features of Rattew's model. Like Rattew, these authors favour a model in which proteins are almost entirely coiled coil along their length. A single coiled coil unit makes up the entirety of the material, and adjacent coiled coils are arranged antiparallel to each other. The arrangement is thought to produce layers with p2 plane group symmetry, which stack exactly in register with those above and below.

Kramer and colleagues (Kramer et al., 1973) reported three proteins in the ootheca of the Chinese mantis *Tenodera sinensis*, dubbing them oothecins I-III. Oothecins I and II had molecular weights of 43 kDa and 60 kDa respectively, with the majority of the ootheca reported to consist of oothecin I. Both oothecin I and oothecin

II were rich in glycine, serine, glutamate/glutamine and aspartate/asparagine, and neither protein responded to tests for glycoproteins or lipoproteins. Rudall (Rudall and Kenchington, 1971) reported the most abundant residues in *Mantis religiosa* ootheca as glutamate/glutamine (21.4%), alanine (14.7%), lysine (10.1%) and aspartate/asparagines (9.5%). Bullough and Tulloch (Bullough and Tulloch, 1990) and Rattew (Rattew, 1974) both reported that the mantis fibroins they studied had a molecular weight near 55 kDa, but did not publish any results or methods to support their conclusions.

Silks have long found applications as biomedical materials, as they are typically biocompatible, biodegradable and have low immunogenicity. For biomedical applications, recombinant silk can be fabricated into various formats. Included in these is the fabrication of a fibre, a sponge, films, particles or hydrogels. Variations in properties can be achieved by using different fabrication and post fabrication processing to match specific clinical needs, so that silks can be produced for applications that need high stress prior to failure, or where extensibility is required, such as in blood vessels, and where an appropriate modulus is required to modify or control cell response, for example for tissue engineering (Vepari and Kaplan, 2007). For example, silk has been used to form non-woven mats (Dal Pra et al., 2005), and has been electrospun into fibres and fibre mats with fibre sizes from nanometres to microns (Jin et al., 2002). Silk fibroin films can be cast from aqueous and non-aqueous solvents (Minoura et al., 1990). Porous sponges can be made from silk solutions, for example, by using salt or sugar as porogens with fibroin in HFIP (Nazarov et al., 2004) or in fully aqueous system (Kim et al., 2005).

As a result of their potential utility, and as silk fibres represent some of the strongest natural fibres known, they have been subject to considerable research in attempts to reproduce their synthesis. However, a recurrent problem with expression of Lepidopteran and spider silk "fibroin" genes has been low expression rates in various recombinant expression systems due to the combination of repeating nucleotide motifs that lead to deleterious recombination events, large gene size and the small number of codons for each amino acid which leads to depletion of tRNA pools. Recombinant expression leads to difficulties during translation such as translational pauses as a result of codon preferences and codon demands and extensive recombination rates leading to truncation of the genes.

Considering the unique properties of silks produced by insects, and that they are available naturally in only small amounts, there is a need for the identification of further novel nucleic acids encoding silk proteins.

SUMMARY OF THE INVENTION

The present inventors have characterised a new class of silk proteins. The proteins, nucleic acids encoding them, and examples of their use are described herein.

In a first aspect, the present invention provides an isolated and/or exogenous polynucleotide which encodes a silk polypeptide, wherein at least a portion of the polypeptide has a coiled coil structure comprising at least 10 copies of the heptad sequence *abcdefg*, and wherein at least 25% of the amino acids at position *a* are alanine residues and at least 20% of the amino acids at position *d* are aromatic residues. Preferably, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55% or at least 60% of the amino acids at position *a* are alanine residues and/or at least 25%, at least 30%, at least 35%, at least 40%, at least 45% or at least 50% of the amino acids at position *d* are aromatic residues.

The aromatic amino acids may be any naturally occurring or non-naturally occurring aromatic amino acids. For example, the aromatic residues may be selected from the group consisting of tyrosine, phenylalanine, tryptophan, histidine and proline.

In an embodiment, the polynucleotide encodes a polypeptide with a coiled coil structure comprising at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, or more copies of the heptad sequence *abcdefg*.

The heptad repeats may be contiguous in the polypeptide sequence or may not be contiguous in the polypeptide sequence. Thus, the polypeptide may comprise any number of contiguous heptad sequences, provided that the total number of heptad sequences in the polypeptide is at least 10. Preferably, the polypeptide comprises at least 6 contiguous heptad sequences, such as at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25 or more contiguous heptad sequences.

The amino acids at position *a* and/or *d* may be more hydrophobic on average than amino acids at other positions in the heptad sequence. In addition or alternatively, the average charge of amino acids at position *g* and/or *b* may be positive at neutral pH, and/or the average charge of amino acids at position *c* may be negative at neutral pH.

In an embodiment, the polynucleotide encodes a polypeptide comprising an amino acid sequence selected from:

i) an amino acid sequence as provided in any one of SEQ ID NOs 13-24, more preferably an amino acid sequence as provided in any one of SEQ ID NOs 13-18;

ii) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 13-24, more preferably an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 13-18; or

iii) a biologically active fragment of i) or ii).

The polynucleotide may encode a silk polypeptide fused to at least one other polypeptide. The at least one other polypeptide may be, for example, a polypeptide that enhances the stability of a polypeptide of the invention, a polypeptide that assists in the purification of the fusion protein, a polypeptide which promotes the formation of a coiled coil, a polypeptide which assists in the polypeptide of the invention being secreted from a cell (for example secreted from a bacterial cell), or any other polypeptide.

Particular sequences of the polynucleotides are exemplified herein. In one example, the polynucleotide comprises a sequence selected from:

i) a nucleotide sequence as provided in any one of SEQ ID NOs 1-12, more preferably a nucleotide sequence as provided in any one of SEQ ID NOs 1-6;

ii) a nucleotide sequence which is at least 30% identical to any one or more of SEQ ID NOs 1-12, more preferably a nucleotide sequence which is at least 30% identical to any one or more of SEQ ID NOs 1-6;

iii) a nucleotide sequence encoding a polypeptide as provided in any one of SEQ ID NOs 13-24, more preferably an amino acid sequence as provided in any one of SEQ ID NOs 13-18;

iv) a nucleotide sequence encoding a polypeptide which is at least 30% identical to the polypeptide as provided in any one of SEQ ID NOs 13-24, more preferably an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 13-18;

v) a nucleotide sequence encoding a biologically active fragment of a polypeptide as provided in any one of SEQ ID NOs 13-24 or a biologically active fragment of a polypeptide which is at least 30% identical to the polypeptide as provided in any one of SEQ ID NOs 13-24, more preferably a nucleotide sequence encoding a biologically active fragment of a polypeptide as provided in any one of SEQ ID NOs 13-18 or a biologically active fragment of a polypeptide which is at least 30%

identical to the polypeptide as provided in any one of SEQ ID NOs 13-18; or

vi) a sequence which hybridizes to the nucleotide sequence of any one of i) to v) under stringent conditions.

Also disclosed herein is a vector comprising at least one polynucleotide of the invention. The vector may be an expression vector, allowing recombinant expression of the polypeptides of the invention.

Also disclosed herein is a host cell comprising at least one polynucleotide of the invention, and/or at least one vector of the invention. The host cell may be a bacterial, yeast or plant cell. In one example, the host cell is an *Escherichia coli* cell. The host *Escherichia coli* cell may comprise one or more inclusion bodies comprising a polypeptide expressed by the polynucleotide of the invention and/or by the vector of the invention.

In another aspect, the present disclosure provides an isolated and/or recombinant silk polypeptide, wherein at least a portion of the polypeptide has a coiled coil structure comprising at least 10 copies of the heptad sequence *abcdefg*, and wherein at least 25% of the amino acids at position *a* are alanine residues and at least 20% of the amino acids at position *d* are aromatic residues.

The polypeptide may have any of the exemplified features described herein in relation to the encoding polynucleotide. Thus, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55% or at least 60% of the amino acids at position *a* may be alanine residues and/or at least 25%, at least 30%, at least 35%, at least 40%, at least 45% or at least 50% of the amino acids at position *d* may be aromatic residues.

The aromatic amino acids may be any naturally occurring or non-naturally occurring aromatic amino acids. For example, the aromatic residues may be selected from the group consisting of tyrosine, phenylalanine, tryptophan, histidine and proline.

In an embodiment, the polypeptide comprises at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24 or at least 25 copies of the heptad sequence *abcdefg*.

The heptad repeats may be contiguous in the polypeptide sequence or may not be contiguous in the polypeptide sequence. Thus, the polypeptide may comprise any number of contiguous heptad sequences, provided that the total number of heptad sequences in the polypeptide is at least 10. Preferably, the polypeptide comprises at least 6 contiguous heptad sequences, such as at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least

18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25 or more contiguous heptad sequences.

The amino acids at position *a* and/or *d* may be more hydrophobic on average than amino acids at other positions in the heptad sequence. In addition or alternatively, the average charge of amino acids at position *g* and/or *b* may be positive at neutral pH, and/or the average charge of amino acids at position *c* may be negative at neutral pH.

In an embodiment, the polypeptide comprises an amino acid sequence selected from:

i) an amino acid sequence as provided in any one of SEQ ID NOs 13-24, more preferably an amino acid sequence as provided in any one of SEQ ID NOs 13-18;

ii) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 13-24, more preferably an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 13-18; or

iii) a biologically active fragment of i) or ii).

The polypeptide may be fused to at least one other polypeptide. In one example, the at least one other polypeptide is selected from the group consisting of: a polypeptide that enhances the stability of a polypeptide of the invention, a polypeptide that assists in the purification of the fusion protein, a polypeptide which promotes the formation of a coiled coil, and a polypeptide which assists in the polypeptide of the invention being secreted from a cell (for example secreted from a bacterial cell).

Also disclosed herein is a composition comprising one or more silk polypeptides of the invention. In one embodiment, the composition comprises one or more dimers, each dimer comprising a first polypeptide of the invention and a second polypeptide of the invention. The composition may comprise a single type of silk polypeptide of the invention, namely each silk polypeptide in the composition may have the same sequence. Thus, in one embodiment, the composition comprises one or more homodimers. In another embodiment, the composition comprises heterodimers of two different silk polypeptides of the invention. Furthermore, the composition may comprise two or more different homodimers, or two or more different heterodimers of the invention.

In one embodiment, the composition comprises one or more heterodimers, wherein:

i) the first polypeptide comprises an amino acid sequence as provided in SEQ ID NO:13, or an amino acid sequence which is at least 30% identical to SEQ ID NO:13, or a biologically active fragment thereof, and the

second polypeptide comprises an amino acid sequence as provided in SEQ ID NO:14, or an amino acid sequence which is at least 30% identical to SEQ ID NO:14, or a biologically active fragment thereof; or

ii) the first polypeptide comprises an amino acid sequence as provided in SEQ ID NO:15, or an amino acid sequence which is at least 30% identical to SEQ ID NO:15, or a biologically active fragment thereof, and the second polypeptide comprises an amino acid sequence as provided in SEQ ID NO:16, or an amino acid sequence which is at least 30% identical to SEQ ID NO:16, or a biologically active fragment thereof; or

iii) the first polypeptide comprises an amino acid sequence as provided in SEQ ID NO:17, or an amino acid sequence which is at least 30% identical to SEQ ID NO:17, or a biologically active fragment thereof, and the second polypeptide comprises an amino acid sequence as provided in SEQ ID NO:18, or an amino acid sequence which is at least 30% identical to SEQ ID NO:18, or a biologically active fragment thereof.

The composition disclosed herein may further comprise one or more acceptable carriers, and/or may further comprise a drug. Accordingly, the composition may be provided for use as a medicine, a medical device or a cosmetic.

In another aspect, the present disclosure provides a process for preparing a polypeptide of the invention or a composition of the invention, the process comprising cultivating a host cell comprising at least one polynucleotide of the invention, and/or at least one vector of the invention, under conditions which allow expression of the polynucleotide encoding the polypeptide. The process may further comprise a step of recovering the expressed polypeptide.

In another aspect, the present disclosure provides a transgenic non-human organism comprising an exogenous polynucleotide of the invention, or a vector of the invention. The transgenic non-human organism may be a plant or a non-human animal.

In another aspect, the present disclosure provides an antibody which specifically binds a polypeptide of the invention.

In another aspect, the present disclosure provides a product comprising at least one polypeptide or composition of the invention. The product may be selected from, but not necesseraily limited to, the group consisting of: a personal care product, textiles, plastics, and biomedical products.

In another aspect, the present disclosure provides a composition comprising at least one polynucleotide of the invention, and/or at least one vector of the invention, and/or at least one host cell of the invention, and/or at least one antibody of the

invention, and one or more acceptable carriers. The one or more carriers may be pharmaceutically and/or agriculturally acceptable carriers.

In another aspect, the present disclosure provides a method of treating or preventing a disease, the method comprising administering a composition comprising a drug for treating or preventing the disease and a pharmaceutically acceptable carrier, wherein the pharmaceutically acceptable carrier comprises at least one polypeptide of the invention and/or product of the invention.

In another aspect, the present disclosure provides the use of at least one polypeptide of the invention and/or product of the invention, and a drug, in the manufacture of a medicament for treating or preventing a disease.

In another aspect, the present disclosure provides a kit comprising one or more of at least one polynucleotide of the invention, at least one vector of the invention, at least one host cell of the invention, at least one polypeptide of the invention, at least one composition of the invention, at least one antibody of the invention, and at least one product of the invention.

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

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. Ootheca material is a mixture of 5-9 proteins. The main mantis silk proteins are between 40 and 60 kDa. Left lane, *P. albofimbriata*; right lane, *T. australasiae*. Molecular weight markers corresponding to the BenchMark Protein Ladder (Invitrogen, Carlsbad, USA) are shown on the right.

Figure 2. ClustalW alignment of mantis fibroins 1. Heptad predictions shown are MARCOIL predictions for the *P. albofimbriata* sequence over 50% likelihood. Similar predictions were obtained for the other sequences. Amino acid identity or similarity is indicated by an asterisk or a colon respectively. Predicted signal sequences are shown underlined, and cysteine residues are shaded. *Pal*MF1 is *P. albofimbriata* Mantis

Fibroin 1, *Tau*MF1 is *T. australasiae* Mantis Fibroin 1, and *Amo*MF1a is *A. monstrosa* Mantis Fibroin 1a.

Figure 3. ClustalW alignments of mantis fibroins 2. Heptad predictions shown are MARCOIL predictions for the *P. albofimbriata* sequence over 50% likelihood. Similar predictions were obtained for the other sequences. Amino acid identity or similarity is indicated by an asterisk or a colon respectively. Predicted signal sequences are shown underlined, and cysteine residues are shaded. *Pal*MF2a is *P. albofimbriata* Mantis Fibroin 2a, *Tau*MF2a is *T. australasiae* Mantis Fibroin 2a, and *Amo*MF2 is *A. monstrosa* Mantis Fibroin 2.

Figure 4. Predicted alpha-helical and coiled coil domains in mantis fibroins. **A:** *P. albofimbriata* Mantis Fibroin 1; **B:** *P. albofimbriata* Mantis Fibroin 2. The solid line shows the prediction of a coiled coil domain by MARCOIL. The dotted line shows the prediction of alpha-helical structure by GOR4.

Figure 5. Trends in amino acid character between heptad positions for Mantis Fibroin 1 (dark bars) and Mantis Fibroin 2 (light bars). Values shown are average values for each species and error bars indicate the standard deviations between them. **A:** residue hydrophobicity (determined according to the Eisenberg scale). **B:** side-chain length (the length from Cα to the furthest non-hydrogen atom). **C:** charge (at pH 7).

Figure 6. Proportion of amino acid types in each heptad position. Trends in amino acid type were similar in Mantis Fibroin 1 (**A**) and Mantis Fibroin 2 (**B**). Data is pooled from all species. In both proteins, the *a* position is dominated by alanine; alanine is rare in the *d* position, which is hydrophobic due to a high proportion of aromatic residues; the *c* position is dominated by negatively charged residues; and the *b* and *g* positions by positively charged residues.

Figure 7. Mantis fibroins expressed in inclusion bodies and solubilised in 0.5% SDS. Lane M, Benchmark Protein Ladder molecular weight markers (Invitrogen); Lane 1, *T. australasiae* Mantis Fibroin 1; Lane 2, *T. australasiae* Mantis Fibroin 2a; Lane 3, *T. australasiae* Mantis Fibroin 2b; Lane 4, *A. monstrosa* Mantis Fibroin 1; Lane 5, *A. monstrosa* Mantis Fibroin 2; Lane 6, *P. albofimbriata* Mantis Fibroin 1; Lane 7, *P. albofimbriata* Mantis Fibroin 2a; Lane 8, *P. albofimbriata* Mantis Fibroin 2b. In each case between 2-5 µl of protein solution was run.

Figure 8. Recombinantly expressed Mantis fibroins are capable of forming coiled coils in solution. The spectrum shown is for *T. australasiae* Mantis Fibroin 2b.

KEY TO THE SEQUENCE LISTING

SEQ ID NO: 1 - Nucleotide sequence encoding the mature form of *Tenodera australasiae* protein Mantis Fibroin 1.

SEQ ID NO: 2 - Nucleotide sequence encoding the mature form of *Tenodera australasiae* protein Mantis Fibroin 2b.

SEQ ID NO: 3 - Nucleotide sequence encoding the mature form of *Archimantis monstrosa* protein Mantis Fibroin 1a.

SEQ ID NO: 4 - Nucleotide sequence encoding the mature form of *Archimantis monstrosa* protein Mantis Fibroin 2.

SEQ ID NO: 5 - Nucleotide sequence encoding the mature form of *Pseudomantis albofimbriata* protein Mantis Fibroin 1.

SEQ ID NO: 6 - Nucleotide sequence encoding the mature form of *Pseudomantis albofimbriata* protein Mantis Fibroin 2b.

SEQ ID NO: 7 - Nucleotide sequence encoding full length *Tenodera australasiae* protein Mantis Fibroin 1.

SEQ ID NO: 8 - Nucleotide sequence encoding full length *Tenodera australasiae* protein Mantis Fibroin 2b.

SEQ ID NO: 9 - Nucleotide sequence encoding full length *Archimantis monstrosa* protein Mantis Fibroin 1a.

SEQ ID NO: 10 - Nucleotide sequence encoding full length *Archimantis monstrosa* protein Mantis Fibroin 2.

SEQ ID NO: 11 - Nucleotide sequence encoding full length *Pseudomantis albofimbriata* protein Mantis Fibroin 1.

SEQ ID NO: 12 - Nucleotide sequence encoding full length *Pseudomantis albofimbriata* protein Mantis Fibroin 2b.

SEQ ID NO: 13 – Amino acid sequence of the mature form of *Tenodera australasiae* protein Mantis Fibroin 1.

SEQ ID NO: 14 – Amino acid sequence of the mature form of *Tenodera australasiae* protein Mantis Fibroin 2b.

SEQ ID NO: 15 – Amino acid sequence of the mature form of *Archimantis monstrosa* protein Mantis Fibroin 1a.

SEQ ID NO: 16 – Amino acid sequence of the mature form of *Archimantis monstrosa* protein Mantis Fibroin 2.

SEQ ID NO: 17 – Amino acid sequence of the mature form of *Pseudomantis albofimbriata* protein Mantis Fibroin 1.

SEQ ID NO: 18 – Amino acid sequence of the mature form of *Pseudomantis albofimbriata* protein Mantis Fibroin 2b.

SEQ ID NO: 19 - Amino acid sequence of full length *Tenodera australasiae* protein Mantis Fibroin 1.

SEQ ID NO: 20 - Amino acid sequence of full length *Tenodera australasiae* protein Mantis Fibroin 2b.

SEQ ID NO: 21 - Amino acid sequence of full length *Archimantis monstrosa* protein Mantis Fibroin 1a.

SEQ ID NO: 22 - Amino acid sequence of full length *Archimantis monstrosa* protein Mantis Fibroin 2.

SEQ ID NO: 23 - Amino acid sequence of full length *Pseudomantis albofimbriata* protein Mantis Fibroin 1.

SEQ ID NO: 24 - Amino acid sequence of full length *Pseudomantis albofimbriata* protein Mantis Fibroin 2b.

SEQ ID NO: 25 - *T. australiasiae* Mantis Fibroin 1 forward primer.

SEQ ID NO: 26 - *T. australiasiae* Mantis Fibroin 1 reverse primer.

SEQ ID NO: 27 - *T. australiasiae* Mantis Fibroin 2 forward primer.

SEQ ID NO: 28 - *T. australiasiae* Mantis Fibroin 2 reverse primer.

SEQ ID NO: 29 - *A. monstrosa* Mantis Fibroin 1 forward primer.

SEQ ID NO: 30 - *A. monstrosa* Mantis Fibroin 1 reverse primer.

SEQ ID NO: 31 - *A. monstrosa* Mantis Fibroin 2 forward primer.

SEQ ID NO: 32 - *A. monstrosa* Mantis Fibroin 2 reverse primer.

SEQ ID NO: 33 - *P. albofimbriata* Mantis Fibroin 1 forward primer.

SEQ ID NO: 34 - *P. albofimbriata* Mantis Fibroin 1 reverse primer.

SEQ ID NO: 35 - *P. albofimbriata* Mantis Fibroin 2 forward primer.

SEQ ID NO: 36 - *P. albofimbriata* Mantis Fibroin 2 reverse primer.

SEQ ID NO: 37 – JQ421298 nucleotide sequence.

SEQ ID NO: 38 – JQ421299 nucleotide sequence.

SEQ ID NO: 39 – JQ421300 nucleotide sequence.

SEQ ID NO: 40 – JQ421301 nucleotide sequence.

SEQ ID NO: 41 – JQ421302 nucleotide sequence.

SEQ ID NO: 42 – JQ421303 nucleotide sequence.

SEQ ID NO: 43 – JQ421304 nucleotide sequence.
SEQ ID NO: 44 – JQ421305 nucleotide sequence.
SEQ ID NO: 45 – JQ421306 nucleotide sequence.
SEQ ID NO: 46 – JQ421298 amino acid sequence.
SEQ ID NO: 47 – JQ421299 amino acid sequence.
SEQ ID NO: 48 – JQ421300 amino acid sequence.
SEQ ID NO: 49 – JQ421301 amino acid sequence.
SEQ ID NO: 50 – JQ421302 amino acid sequence.
SEQ ID NO: 51 – JQ421303 amino acid sequence.
SEQ ID NO: 52 – JQ421304 amino acid sequence.
SEQ ID NO: 53 – JQ421305 amino acid sequence.
SEQ ID NO: 54 – JQ421306 amino acid sequence.

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), and are incorporated herein by reference.

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 term "about", unless stated to the contrary, refers to +/- 20%, more preferably +/- 10%, even more preferably +/- 5%, of the designated value.

As used herein, the terms "silk protein" and "silk polypeptide" refer to a fibrous protein/polypeptide that can be used to produce materials such as silk fibre, silk film, silk sponges, silk particles and/or a fibrous protein complex. Typically, the silk proteins will be produced by recombinant expression. However, the silk proteins can be produced artificially such, for example, using a protein synthesizer. Silk proteins may have a sequence corresponding to a naturally occurring silk protein (for example mantis silk proteins described herein) or be a man made variant thereof. Such variants not only include small substitutions, deletions and additions, but also encompass significant rearrangement of the native sequences where, for example, heptads are re-ordered so they bear no resemblance to the primary amino acid sequence of the native protein but because of the heptad structure are still functional silk proteins.

The term "plant" includes whole plants, vegetative structures (for example, leaves, stems), roots, floral organs/structures, seed (including embryo, endosperm, and seed coat), plant tissue (for example, vascular tissue, ground tissue, and the like), cells and progeny of the same.

A "transgenic plant" refers to a plant that contains a gene construct ("transgene") not found in a wild-type plant of the same species, variety or cultivar. A "transgene" as referred to herein has the normal meaning in the art of biotechnology and includes a genetic sequence which has been produced or altered by recombinant DNA or RNA technology and which has been introduced into the plant cell. The transgene may include genetic sequences derived from a plant cell. Typically, the transgene has been introduced into the plant by human manipulation such as, for example, by transformation but any method can be used as one of skill in the art recognizes.

"Polynucleotide" refers to an oligonucleotide, nucleic acid molecule or any fragment thereof. It may be DNA or RNA of genomic or synthetic origin, double-stranded or single-stranded, and combined with carbohydrate, lipids, protein, or other materials to perform a particular activity defined herein.

"Operably linked" as used herein refers to a functional relationship between two or more nucleic acid (e.g., DNA) segments. Typically, it refers to the functional relationship of transcriptional regulatory element to a transcribed sequence. For example, a promoter is operably linked to a coding sequence, such as a polynucleotide defined herein, if it stimulates or modulates the transcription of the coding sequence in an appropriate host cell. Generally, promoter transcriptional regulatory elements that are operably linked to a transcribed sequence are physically contiguous to the

transcribed sequence, i.e., they are *cis*-acting. However, some transcriptional regulatory elements, such as enhancers, need not be physically contiguous or located in close proximity to the coding sequences whose transcription they enhance, i.e., they may be *trans*-acting.

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. Signal peptides can also be referred to as signal sequence, and are well known in the art.

As used herein, the term "fused to", or variations thereof (such as "conjugated to") are used broadly to refer to any form to covalent or non-covalent association between a polypeptide and/or polynucleotide of the invention and another polypeptide and/or polynucleotide. In one example, a polypeptide may be fused to another polypeptide by a peptide bond, which may be achieved by expression of a polynucleotide encoding a polypeptide of the invention fused in frame to a polynucleotide encoding another polypeptide.

As used herein, "transformation" is the acquisition of new genes in a cell by the incorporation of a polynucleotide.

As used herein, the term "drug" refers to any compound that can be used to treat or prevent a particular disease, examples of drugs which can be formulated with a silk protein of the invention include, but are not limited to, proteins, nucleic acids, anti-tumor agents, analgesics, antibiotics, anti-inflammatory compounds (both steroidal and non-steroidal), hormones, vaccines, labeled substances, and the like.

Polypeptides

By "isolated polypeptide" we mean a polypeptide that has generally been separated from the lipids, nucleic acids, other polypeptides, and other contaminating molecules with which it is associated in its native state. With the exception of other proteins of the invention, it is preferred that the isolated polypeptide is at least 60% free, more preferably at least 75% free, and more preferably at least 90% free from other components with which it is naturally associated. An "isolated" polypeptide can be made artificially, for example using a protein synthesizer.

The term "recombinant" in the context of a polypeptide refers to the polypeptide when produced by a cell, or in a cell-free expression system, in an altered amount or at an altered rate compared to its native state. In one embodiment the cell is a cell that

does not naturally produce the polypeptide. However, the cell may be a cell which comprises a non-endogenous gene that causes an altered, preferably increased, amount of the polypeptide to be produced. A recombinant polypeptide as disclosed herein includes polypeptides which have not been separated from other components of the transgenic (recombinant) cell, or cell-free expression system, in which it is produced, and polypeptides produced in such cells or cell-free systems which are subsequently purified away from at least some other components.

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 fragments (such as biologically active fragments), variants, mutants, modifications, analogous and/or derivatives of the polypeptides disclosed herein.

As used herein a "biologically active" fragment is a portion of a polypeptide of the invention which maintains a defined activity of the full-length polypeptide, for example, the ability to reproduce any one or more characteristics of a silk protein. Thus, a biologically active fragment of a polypeptide of the invention may be capable of forming or may form a semicrystalline material (that is, a material with regions of ordered molecular structure (crystallites) within an amorphous matrix). Alternatively or in addition, the biologically active fragment may be capable of forming or may form filamentous molecules. Thus, the biologically active fragment may be capable of being used to produce a silk fibre. Biologically active fragments can be any size as long as they maintain the defined activity.

As disclosed herein, at least a portion of each of the polypeptides has a coiled coil structure comprising at least 10 copies of the heptad sequence *abcdefg*. Thus, the portion of the polypeptide comprising copies of the heptad sequence *abcdefg* may form an alpha helical structure. The coiled coil portion allows the polypeptide to form a coiled coil structure in association with another polypeptide having a similar coiled coil structure. Thus, coiled coil structures are characterized by heptad repeats represented by the consensus sequence $(abcdefg)_n$.

The polypeptides of the invention may comprise any number of heptad repeats greater than 10. For example, the polypeptides may comprise a portion having a coiled coil structure comprising at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39 at least 40, at least 41, at least 42, at least 43, at least 44, at

least 45, at least 46, at least 47, at least 48, at least 49, at least 50, or more copies of the heptad sequence *abcdefg*. In one particular example, the polypeptides comprise a portion having a coiled coil structure comprising at least 19, or at least 23 copies of the heptad sequence *abcdefg*.

In one embodiment, the polypeptides of the invention comprise about 10 to about 30, such as about 15 to about 25, or about 19 to about 23 heptad repeats. Preferably, the polypeptides comprise about 19 or about 23 heptad repeats.

The heptad repeats may be contiguous in the polypeptide sequence or may not be contiguous in the polypeptide sequence. Thus, the polypeptide may comprise any number of contiguous heptad sequences, provided that the total number of heptad sequences in the polypeptide is at least 10. Preferably, the polypeptide comprises at least 6 contiguous heptad sequences, such as at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25 or more contiguous heptad sequences. In one example, the polypeptide comprises at least 19, such as at least 20, at least 21, at least 22, or at least 23 contiguous heptad sequences. In another example, the polypeptide comprises about 10 to about 30, such as about 15 to about 25, or about 19 to about 23 contiguous heptad repeats. Preferably, the polypeptides comprise about 19 or about 23 contiguous heptad repeats.

The heptad repeats may comprise any portion of the polypeptide of the invention. For example, the heptad repeats may comprise at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, of the length of the polypeptide, or the entire polypeptide. Thus, any portion of the polypeptide may comprise a coiled coil structure. For example, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or the entire polypeptide may comprise a coiled coil structure. In one example, at least 70% of the polypeptide comprises a coiled coil structure.

In another example, the heptad repeats may comprise about 60% to about 90%, such as about 70% to about 80% of the polypeptide. Thus, about 60% to about 90%, such as about 70% to about 80% of the polypeptide may comprise a coiled coil structure.

The remainder of the polypeptide that does not form a coiled coil structure can form any secondary protein structure. For example, the remainder of the polypeptide may form random coils, alpha helices and/or beta-sheets.

The composition of amino acids making up each heptad repeat may be the same or may differ from one heptad sequence to another. As will be understood by a person skilled in the art, a large number of combinations and permutations of different amino acids in the heptad sequence *abcdefg* can achieve the same effect of producing a coiled coil structure, which allows the formation of a coiled coil in association with another polypeptide. Guidance regarding amino acid substitutions which can be made to the polypeptides disclosed herein is provided, by way of example only, in Figures 2 and 3 and in Tables 1 and 3. Where a predicted useful amino acid substitution based on the experimental data provided herein is in any way in conflict with the exemplary substitutions provided in Table 1 it is preferred that a substitution based on the experimental data is used.

In addition, the polypeptides may comprise certain disruptions within and/or between each heptad repeat which nevertheless allow the formation of a coiled coil structure. For example, a heptad sequence may be truncated by one or more amino acids or extended by one or more amino acids, whilst still forming a coiled coil structure. Thus, as stated above, two copies of the heptad sequence *abcdefg* may be separated by one or more amino acids, which nevertheless still allows the formation of an coiled coil structure.

At least 25% of the amino acids at position *a* in the heptad repeats in the polypeptides of the invention are alanine residues. For example, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% or more of the amino acids at position *a* may be alanine residues. Preferably at least 45% of the amino acids at position *a* in the heptad repeats in the polypeptides of the invention are alanine residues.

In another example, from about 30% to about 70%, such as from about 40% to about 60%, or from about 45% to about 55% of the amino acids at position *a* in the heptad repeats in the polypeptides of the invention are alanine residues. In one example, about 50% of the amino acids at position *a* in the heptad repeats in the polypeptides of the invention are alanine residues.

In addition, at least 20% of the amino acids at position *d* in the heptad repeats in the polypeptides of the invention are aromatic residues. For example, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least

60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% or more of the amino acids at position *d* may be aromatic residues.

In another example, from about 20% to about 50%, such as from about 25% to about 45%, such as from about 30% to about 40% of the amino acids at position *d* in the heptad repeats in the polypeptides of the invention are aromatic residues. In one example, about 33% of the amino acids at position *d* in the heptad repeats in the polypeptides of the invention are aromatic residues.

The aromatic residues may be any naturally occurring or any non-naturally occurring aromatic residues. For example, the aromatic residues may be selected from the group consisting of tyrosine, phenylalanine, tryptophan, histidine and proline. A high proportion of the aromatic residues in the polypeptides of the invention may be tyrosine. For example, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, or more of the amino acids at position *d* may be tyrosine residues. Alternatively, the aromatic residues may be any non-naturally occurring aromatic residue as will be known to a person skilled in the art.

The particular aromatic residues in each heptad (if present) may be the same or may be different. Thus, any combination or permutation of aromatic residues may be present in the heptad sequences in the polypeptides of the invention.

The relative proportions of alanine at position *a* in the heptad repeats and aromatic residues at position *d* in the heptad repeats can vary, provided that at least 25% of the amino acids at position *a* are alanine residues and at least 20% of the amino acids at position *d* are aromatic residues. Thus, the polypeptides of the invention can comprise heptad repeats wherein at least 25% of the amino acids at position *a* are alanine residues and wherein the aromatic residues are present at position *d* at proportions of any one of at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% or more, and *vice versa*.

The amino acids at positions *a* and/or *d* in the heptad sequence in the polypeptides of the invention may be more hydrophobic on average than amino acids at other positions in the heptad sequence. Thus, the average hydrophobicity of amino acids at positions *a* and/or *d* may be greater than the average hydrophobicity of amino acids at each of the remaining positions in the heptad sequence. Hydrophobicity of an amino acid residue can be determined by any method known in the art. For example, hydrophobicity can be predicted based on the physiochemical properties of the amino acid side chains, or may be determined by partitioning of an amino acid between two

immiscible liquid phases. The use of these methods to determine the relative hydrophobicity of each of the naturally occurring amino acids has resulted in the production of several known hydrophobicity scales (see, by way of example only, Kallol et al. (2003); Kyte and Doolittle (1982); Eisenberg (1984); Rose and Wolfenden (1993)). Any of these, or other known hydrophobicity scales, can be used to determine the hydrophobicity (and hence, the average hydrophobicity) of the amino acids present at each position in the heptad sequence *abcdefg*. In one example, amino acid hydrophobicity is determined according to the Eisenberg scale. In a particular example, the average hydrophobicity of amino acids at positions *a* and/or *d* in the heptad sequence is positive according to the Eisenberg scale, and the average hydrophobicity of amino acids at each of the remaining positions in the heptad sequence is negative according to the Eisenberg scale. Generally, the following amino acids are considered to be more hydrophobic than others: cysteine, glycine, isoleucine, leucine, methionine and valine. Aromatic amino acids are also generally considered to be more hydrophobic than non-aromatic amino acids.

Another feature of the polypeptides of the invention is that the average length of amino acid side chain may be lower for amino acids at position *a* in the heptad sequence than amino acids at other positions in the heptad sequence. The average side chain length may be calculated, for example, as the length from the alpha-carbon to the furthest non-hydrogen atom.

Another feature of the polypeptides of the invention is that the average charge of amino acids at positions *g* and/or *b* in the heptad sequence in the polypeptides of the invention may be positive at neutral pH (pH 7). Additionally or alternatively, the average charge of amino acids at position *c* in the heptad sequence in the polypeptides of the invention may be negative at neutral pH (pH 7). The charge of naturally occurring amino acids at neutral pH is known in the art, and can be determined by methods known in the art. Generally, arginine and lysine are positively charged at neutral pH and glutamate and aspartate are negatively charged at neutral pH.

The polypeptides (and polynucleotides) of the invention can be isolated from at least Dictyopterans. Examples of suitable Dictyopterans include, but are not limited to, *Pseudomantis albofimbriata, Tenodera australasiae,* and *Archimantis monstrosa.*

The polypeptides of the invention may have a size ranging from between about 40kDa to about 60kDa. For example, the polypeptides may have a size of about 40kDa, about 41kDa, about 44kDa, about 48 kDa or about 50kDa.

The polypeptides of the invention are exemplified by a number of particular proteins whose sequences are provided in SEQ ID NOs: 13-24. As will be appreciated

from the present disclosure, the exact amino acid sequence of the polypeptides (and hence, the exact nucleic acid sequence of the polynucleotides) can vary whilst still providing a polypeptide having a structure that is capable of forming a coiled coil in association with another polypeptide. The exemplified sequences should therefore be considered as examples only, and it will be appreciated that significant variation from these particular sequences may be tolerable.

In a particular example, the present disclosure provides a polypeptide comprising an amino acid sequence selected from:

i) an amino acid sequence as provided in any one of SEQ ID NOs 13-24;

ii) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 13-24; or

iii) a biologically active fragment of i) or ii).

Preferably, the polypeptide comprises an amino acid sequence selected from:

i) an amino acid sequence as provided in any one of SEQ ID NOs 13-18;

ii) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 13-18; or

iii) a biologically active fragment of i) or ii).

The % identity of a polypeptide 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 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.

With regard to a polypeptide, it will be appreciated that % identity figures higher than those provided herein will encompass preferred embodiments. Thus, where applicable, in light of the minimum % identity figures, it is preferred that the polypeptide or polynucleotide comprises an amino acid sequence which is at least 35%, more prefereably 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 polypeptides of the invention can be prepared by introducing appropriate nucleotide changes into a nucleic acid of the invention, or by *in vitro* synthesis of the desired polypeptide. 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 polypeptide product possesses the desired characteristics.

Mutant (altered) polypeptides can be prepared using any technique known in the art. For example, a polynucleotide disclosed herein 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 and possibly also genes related to those of the invention, such as silk genes from insect or arthropod 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 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 polypeptide 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
His (H)	asn; gln
Ile (I)	leu; val; ala; met
Leu (L)	ile; val; met; ala; phe
Lys (K)	arg
Met (M)	leu; phe; ile
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; cys

Furthermore, if desired, unnatural amino acids or chemical amino acid analogues can be introduced as a substitution or addition into the polypeptides of the invention. Such amino acids include, but are not limited to, the D-isomers of the common amino acids, 2,4-diaminobutyric acid, α-amino isobutyric acid, 4-aminobutyric acid, 2-aminobutyric acid, 6-amino hexanoic acid, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β-alanine, fluoro-amino acids, designer amino acids such as β-methyl amino acids, Cα-methyl amino acids, Nα-methyl amino acids, and amino acid analogues in general.

The polypeptides of the invention can also be 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 polypeptide.

The polypeptides of the invention may or may not comprise a signal peptide. Thus, the polynucleotides of the invention may or may not encode a signal peptide. Many examples of particular signal peptides which direct the polypeptides to particular cellular locations during expression in a host cell (for example, which facilitate translocation of the polypeptides across a host cell membrane) are known in the art. Particular examples of signal peptides are provided in the specific sequences disclosed herein (see, for example, the amino acids underlined in Figures 2 and 3). The polynucleotides and polypeptides of the invention may include these specific signal peptides or may not include these specific signal peptides. Thus, the polypeptides of the invention may comprise a sequence comprising any one or more of: SEQ ID NOs 13-18; a polypeptide comprising a sequence which is at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or at least 99.9% identical to any one of SEQ ID NOs 13-18; and a polypeptide comprising a biologically active fragment thereof.

In one example, the polynucleotides and polypeptides of the invention may comprise alternative signal peptides in place of the specific signal peptides shown in Figures 2 and 3.

The polypeptides of the invention can be produced in a variety of ways, including production and recovery of natural polypeptides, production and recovery of

recombinant polypeptides, and chemical synthesis of the polypeptides. In one embodiment, an isolated polypeptide is produced by culturing a cell capable of expressing the polypeptide under conditions effective to produce the polypeptide, and recovering the polypeptide. A preferred cell to culture is a recombinant cell as disclosed herein. Effective culture conditions include, but are not limited to, effective media, bioreactor, temperature, pH and oxygen conditions that permit polypeptide production. An effective medium refers to any medium in which a cell is cultured to produce a polypeptide. Such medium typically comprises an aqueous medium having assimilable carbon, nitrogen and phosphate sources, and appropriate salts, minerals, metals and other nutrients, such as vitamins. Cells can be cultured in conventional fermentation bioreactors, shake flasks, test tubes, microtiter dishes, and petri plates. Culturing can be carried out at a temperature, pH and oxygen content appropriate for a recombinant cell. Such culturing conditions are within the expertise of one of ordinary skill in the art.

Polynucleotides

The term "isolated polynucleotide", which includes DNA, RNA, or a combination of these, single or double stranded, in the sense or antisense orientation or a combination of both, dsRNA or otherwise, refers to a polynucleotide which is at least partially separated from the polynucleotide sequences with which it is associated or linked in its native state. Preferably, the isolated polynucleotide is at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated. Furthermore, 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 polypeptide. 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 polynucleotides of the invention may comprise any nucleotide sequence which results in the expression of a polypeptide having the particular structural features described herein (i.e., wherein at least a portion of the polypeptide has an coiled coil structure comprising at least 10 copies of the heptad amino acid sequence *abcdefg*, which portion allows the polypeptide to form a coiled coil structure in association with another polypeptide having a similar coiled coil structure). Since different amino acids can be used to achieve this structure, it will be appreciated that a wide variety of nucleic acid sequences can encode the polypeptides of the invention. The redundancy of the genetic code further increases the variety of nucleotide sequences that can encode a polynucleotide of the invention. Accordingly, the specific nucleic acid sequences of the invention should be considered as exemplary only, on the understanding that alternative nucleic acid sequences, which may share as little as at least 5%, at least 10%, at least 15%, at least 20% or higher identity with the specific sequences of the invention, can still encode a polypeptide of the invention.

In one example, the polynucleotides of the invention may comprise a sequence selected from:

i) a nucleotide sequence as provided in any one of SEQ ID NOs 1-12;

ii) a nucleotide sequence which is at least 30% identical to any one or more of SEQ ID NOs 1-12;

iii) a nucleotide sequence encoding a polypeptide as provided in any one of SEQ ID NOs 13-24;

iv) a nucleotide sequence encoding a polypeptide which is at least 30% identical to the polypeptide as provided in any one of SEQ ID NOs 13-24;

v) a nucleotide sequence encoding a biologically active fragment of a polypeptide as provided in any one of SEQ ID NOs 13-24 or a biologically active fragment of a polypeptide which is at least 30% identical to the polypeptide as provided in any one of SEQ ID NOs 13-24; or

vi) a sequence which hybridizes to the nucleotide sequence of any one of i) to v) under stringent conditions.

Preferably, the polynucleotides of the invention comprise a sequence selected from:

i) a nucleotide sequence as provided in any one of SEQ ID NOs 1-6;

ii) a nucleotide sequence which is at least 30% identical to any one or more of SEQ ID NOs 1-6;

iii) a nucleotide sequence encoding a polypeptide as provided in any one of SEQ ID NOs 13-18;

iv) a nucleotide sequence encoding a polypeptide which is at least 30% identical to the polypeptide as provided in any one of SEQ ID NOs 13-18;

v) a nucleotide sequence encoding a biologically active fragment of a polypeptide as provided in any one of SEQ ID NOs 13-18 or a biologically active fragment of a polypeptide which is at least 30% identical to the polypeptide as provided in any one of SEQ ID NOs 13-18; or

vi) a sequence which hybridizes to the nucleotide sequence of any one of i) to v) under stringent conditions.

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 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.

The polynucleotides of the 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).

Also disclosed herein are polynucleotides which hybridize to a polynucleotide encoding a silk protein, or to a region flanking said polynucleotide, 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. Alternatively, the nucleic acid and/or oligonucleotides (which may also be referred to as "primers" or "probes") hybridize to the region of an insect genome of interest, such as the genome of a mantis, under conditions used in nucleic acid amplification techniques such as PCR.

The polynucleotides of the invention can be RNA, DNA, or derivatives of either. The polynucleotides can also be referred to as oligonucleotides. Although the terms polynucleotide and oligonucleotide have overlapping meaning, oligonucleotides are typically relatively short single stranded molecules. The minimum size of such oligonucleotides is the size required for the formation of a stable hybrid between an oligonucleotide and a complementary sequence on a target nucleic acid molecule. Preferably, the oligonucleotides are at least 15 nucleotides, more preferably at least 18 nucleotides, more preferably at least 19 nucleotides, more preferably at least 20 nucleotides, even more preferably at least 25 nucleotides in length.

Usually, monomers of a polynucleotide or oligonucleotide are linked by phosphodiester bonds or analogs thereof to form oligonucleotides ranging in size from relatively short monomeric units, e.g., 12-18, to several hundreds of monomeric units. Analogs of phosphodiester linkages include: phosphorothioate, phosphorodithioate, phosphoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phosphoranilidate, phosphoramidate.

The polynucleotides and oligonucleotides of the invention can be used as, for example, probes to identify nucleic acid molecules, or as primers to produce nucleic acid molecules. Polynucleotides or oligonucleotides used as a probe are typically

conjugated with a detectable label such as a radioisotope, an enzyme, biotin, a fluorescent molecule or a chemiluminescent molecule.

Vectors

The present disclosure also provides a vector, which comprises at least one isolated polynucleotide of the invention. The vector may be described as a recombinant vector, and may be any vector capable of delivering the polynucleotide into a host cell. Such a vector contains heterologous polynucleotide sequences, that is polynucleotide sequences that are not naturally found adjacent to the polynucleotides of the invention and that preferably are derived from a species other than the species from which the polynucleotide(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 vector comprises a polynucleotide of the invention 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 invention include any vectors that function (i.e., direct gene expression) in the recombinant cells of the invention, including in bacterial, fungal, endoparasite, arthropod, animal, and plant cells. Particularly preferred expression vectors can direct gene expression in plants cells. The vectors of the invention can also be used to produce the polypeptide in a cell-free expression system; such systems are well known in the art.

In particular, expression vectors disclosed herein may contain 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 polynucleotides disclosed herein. In particular, recombinant molecules (such as polynucleotides, vectors) as disclosed herein may include transcription control sequences. 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 disclosed herein. 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.

Particularly preferred transcription control sequences are promoters active in directing transcription in bacteria, such as *Escherichia coli.*

Other preferred transcription control sequences are promoters active in directing transcription in plants, either constitutively or stage and/or tissue specific, depending on the use of the plant or parts thereof. These plant promoters include, but are not limited to, promoters showing constitutive expression, such as the 35S promoter of Cauliflower Mosaic Virus (CaMV), those for leaf-specific expression, such as the promoter of the ribulose bisphosphate carboxylase small subunit gene, those for root-specific expression, such as the promoter from the glutamine synthase gene, those for seed-specific expression, such as the cruciferin A promoter from *Brassica napus*, those for tuber-specific expression, such as the class-I patatin promoter from potato or those for fruit-specific expression, such as the polygalacturonase (PG) promoter from tomato.

Recombinant molecules as disclosed herein may also (a) contain secretory signals (i.e., signal segment nucleic acid sequences) to enable an expressed polypeptide as disclosed herein to be secreted from the cell that produces the polypeptide and/or (b) contain fusion sequences which lead to the expression of nucleic acid molecules as disclosed herein as fusion proteins. Examples of suitable signal segments include any signal segment capable of directing the secretion of a polypeptide as disclosed herein. Preferred 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 a polypeptide as disclosed herein. In addition, a nucleic acid molecule as disclosed herein can be joined to a fusion segment that directs the encoded polypeptide to the proteosome, such as a ubiquitin fusion segment. Recombinant molecules may also include intervening and/or untranslated sequences surrounding and/or within the nucleic acid sequences disclosed herein.

The polypeptides and/or compositions of the invention can be produced by expressing the polynucleotides and/or vectors of the invention in a suitable expression system. The expression system may comprise a cell free *in vitro* expression system or a host cell.

Host Cells

The present disclosure also provides a host cell comprising one or more polynucleotides and/or one or more vectors of the invention. The host cell may be referred to as a recombinant cell, and may be transformed with one or more polynucleotides and/or one or more vectors of the invention. Also disclosed herein are progeny cells of such host cells. Transformation of a polynucleotide into a cell can be accomplished by any method by which a polynucleotide 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 polynucleotides 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 of the invention. Host cells either can be endogenously (i.e., naturally) capable of producing polypeptides of the invention or can be capable of producing such polypeptides after being transformed with at least one polynucleotide of the invention. Host cells can be any cell capable of producing at least one polypeptide of the invention, and include bacterial, fungal (including yeast), parasite, arthropod, animal and plant cells. The host cell may be a non-human host cell. Examples of suitable 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 plant cells such as those available from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (German Collection of Microorganisms and Cell Cultures).

In another preferred embodiment, the host cell is a bacterial cell such as an *E. coli* cell. One advantage of using host bacterial cells is that the polypeptides of the invention can be expressed so as to accumulate in inclusion bodies in the host cell. This allows the polypeptides to be conveniently recovered from the host cell. Expressed polypeptides can be recovered from inclusion bodies by any suitable method known in the art.

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 of the invention include, but are not limited to, operatively linking polynucleotides to high-copy number plasmids, integration of the polynucleotide 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 polynucleotides of the invention to correspond to the codon usage of the host cell, and the deletion of sequences that destabilize transcripts.

Transgenic Plants

The term "plant" refers to whole plants, plant organs (e.g. leaves, stems roots, etc), seeds, plant cells and the like. Plants contemplated for use in the practice of the present disclosure include both monocotyledons and dicotyledons. Target plants include, but are not limited to, the following: cereals (wheat, barley, rye, oats, rice, sorghum and related crops); beet (sugar beet and fodder beet); pomes, stone fruit and soft fruit (apples, pears, plums, peaches, almonds, cherries, strawberries, raspberries and black-berries); leguminous plants (beans, lentils, peas, soybeans); oil plants (rape,

mustard, poppy, olives, sunflowers, coconut, castor oil plants, cocoa beans, groundnuts); cucumber plants (marrows, cucumbers, melons); fibre plants (cotton, flax, hemp, jute); citrus fruit (oranges, lemons, grapefruit, mandarins); vegetables (spinach, lettuce, asparagus, cabbages, carrots, onions, tomatoes, potatoes, paprika); lauraceae (avocados, cinnamon, camphor); or plants such as maize, tobacco, nuts, coffee, sugar cane, tea, vines, hops, turf, bananas and natural rubber plants, as well as ornamentals (flowers, shrubs, broad-leaved trees and evergreens, such as conifers).

Transgenic plants, as defined in the context of the present disclosure include plants (as well as parts and cells of said plants) and their progeny which have been genetically modified using recombinant techniques to cause production of at least one polypeptide of the invention in the desired plant or plant organ. Transgenic plants can be produced using techniques known in the art, such as those generally described in A. Slater et al., Plant Biotechnology - The Genetic Manipulation of Plants, Oxford University Press (2003), and P. Christou and H. Klee, Handbook of Plant Biotechnology, John Wiley and Sons (2004).

A polynucleotide of the invention may be expressed constitutively in the transgenic plants during all stages of development. Depending on the use of the plant or plant organs, the polypeptides may be expressed in a stage-specific manner. Furthermore, the polynucleotides may be expressed tissue-specifically.

Regulatory sequences which are known or are found to cause expression of a gene encoding a polypeptide of interest in plants may be used. The choice of the regulatory sequences used depends on the target plant and/or target organ of interest. Such regulatory sequences may be obtained from plants or plant viruses, or may be chemically synthesized. Such regulatory sequences are well known to those skilled in the art.

Constitutive plant promoters are well known. Further to previously mentioned promoters, some other suitable promoters include but are not limited to the nopaline synthase promoter, the octopine synthase promoter, CaMV 35S promoter, the ribulose-1,5-bisphosphate carboxylase promoter, Adh1-based pEmu, Act1, the SAM synthase promoter and Ubi promoters and the promoter of the chlorophyll a/b binding protein. Alternatively it may be desired to have the transgene(s) expressed in a regulated fashion. Regulated expression of the polypeptides is possible by placing the coding sequence of the silk protein under the control of promoters that are tissue-specific, developmental-specific, or inducible. Several tissue-specific regulated genes and/or promoters have been reported in plants. These include genes encoding the seed storage proteins (such as napin, cruciferin, β-conglycinin, glycinin and phaseolin), zein or oil

body proteins (such as oleosin), or genes involved in fatty acid biosynthesis (including acyl carrier protein, stearoyl-ACP desaturase, and fatty acid desaturases (fad 2- 1)), and other genes expressed during embryo development (such as Bce4). Particularly useful for seed-specific expression is the pea vicilin promoter. Other useful promoters for expression in mature leaves are those that are switched on at the onset of senescence, such as the SAG promoter from *Arabidopsis*). A class of fruit-specific promoters expressed at or during anthesis through fruit development, at least until the beginning of ripening, is discussed in US 4,943,674. Other examples of tissue-specific promoters include those that direct expression in tubers (for example, patatin gene promoter), and in fiber cells (an example of a developmentally-regulated fiber cell protein is E6 fiber).

Other regulatory sequences such as terminator sequences and polyadenylation signals include any such sequence functioning as such in plants, the choice of which would be obvious to the skilled addressee. The termination region used in the expression cassette will be chosen primarily for convenience, since the termination regions appear to be relatively interchangeable. The termination region which is used may be native with the transcriptional initiation region, may be native with the polynucleotide sequence of interest, or may be derived from another source. The termination region may be naturally occurring, or wholly or partially synthetic. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions or from the genes for β-phaseolin, the chemically inducible lant gene, pIN.

Several techniques are available for the introduction of an expression construct containing a nucleic acid sequence encoding a polypeptide of interest into the target plants. Such techniques include but are not limited to transformation of protoplasts using the calcium/polyethylene glycol method, electroporation and microinjection or (coated) particle bombardment. In addition to these so-called direct DNA transformation methods, transformation systems involving vectors are widely available, such as viral and bacterial vectors (e.g. from the genus *Agrobacterium*). After selection and/or screening, the protoplasts, cells or plant parts that have been transformed can be regenerated into whole plants, using methods known in the art. The choice of the transformation and/or regeneration techniques is not critical for the purposes of the present disclosure.

To confirm the presence of the transgenes in transgenic cells and plants, a polymerase chain reaction (PCR) amplification or Southern blot analysis can be performed using methods known to those skilled in the art. Expression products of the transgenes can be detected in any of a variety of ways, depending upon the nature of

the product, and include Western blot and enzyme assay. One particularly useful way to quantitate protein expression and to detect replication in different plant tissues is to use a reporter gene, such as GUS. Once transgenic plants have been obtained, they may be grown to produce plant tissues or parts having the desired phenotype. The plant tissue or plant parts, may be harvested, and/or the seed collected. The seed may serve as a source for growing additional plants with tissues or parts having the desired characteristics.

Transgenic Organisms

The present disclosure provides transgenic organisms comprising one or more polypeptides and/or polynucleotides and/or vectors of the invention. Thus, the transgenic organism may comprise one or more recombinant polypeptides and/or exogenous polynucleotides of the invention. Preferably, the transgenic organism is non-human. For example, the transgenic organism may be a non-human animal or a plant.

Techniques for producing transgenic animals are well known in the art. A useful general textbook on this subject is Houdebine, Transgenic animals – Generation and Use (Harwood Academic, 1997).

Heterologous DNA can be introduced, for example, into fertilized mammalian ova. For instance, totipotent or pluripotent stem cells can be transformed by microinjection, calcium phosphate mediated precipitation, liposome fusion, retroviral infection or other means, the transformed cells are then introduced into the embryo, and the embryo then develops into a transgenic animal. In a highly preferred method, developing embryos are infected with a retrovirus containing the desired DNA, and transgenic animals produced from the infected embryo. In a most preferred method, however, the appropriate DNAs are coinjected into the pronucleus or cytoplasm of embryos, preferably at the single cell stage, and the embryos allowed to develop into mature transgenic animals.

Another method used to produce a transgenic animal involves microinjecting a nucleic acid into pro-nuclear stage eggs by standard methods. Injected eggs are then cultured before transfer into the oviducts of pseudopregnant recipients.

Transgenic animals may also be produced by nuclear transfer technology. Using this method, fibroblasts from donor animals are stably transfected with a plasmid incorporating the coding sequences for a binding domain or binding partner of interest under the control of regulatory sequences. Stable transfectants are then fused to enucleated oocytes, cultured and transferred into female recipients.

Recovery Methods and Production of Silk

The polypeptides of the invention may be extracted and purified from recombinant cells, such as plant, bacteria or yeast cells producing said polypeptides by a variety of methods. In one embodiment, the method involves removal of native cell proteins from homogenized cells/tissues/plants etc. by lowering pH and heating, followed by ammonium sulfate fractionation. Briefly, total soluble proteins are extracted by homogenizing cells/tissues/plants. Native proteins are removed by precipitation at pH 4.7 and then at 60°C. The resulting supernatant is then fractionated with ammonium sulfate at 40% saturation. The resulting protein will be of the order of 95% pure. Additional purification may be achieved with conventional gel or affinity chromatography.

In another example, cell lysates are treated with high concentrations of acid e.g. HCl or propionic acid to reduce pH to ~1-2 for 1 hour or more which will solubilise the silk proteins but precipitate other proteins.

Polypeptide aggregates (such as filaments and/or crystallites) can form from solutions by spontaneous self-assembly of silk proteins of the invention when the protein concentration exceeds a critical value. The aggregates may be gathered and mechanically spun into macroscopic fibres according to the method of O'Brien et al. (1994). As used herein, a "silk fibre" refers to a plurality of filaments comprising proteins of the invention which can be woven into various items such as textiles.

By the nature of the inherent coiled coil secondary structure, polypeptides of the invention may spontaneously form the coiled coil secondary structure upon dehydration. As the skilled person would be aware, the strength of the coiled coil can be enhanced through enzymatic or chemical cross-linking of lysine residues in close proximity.

In an embodiment, silk polypeptides of the invention are processed to produce a material (or product) such as, but not limited to, a sponge, particle, filament, fiber or film.

Silk filaments and/or fibres of the invention have a low processing requirement. The silk proteins require minimal processing e.g. spinning to form a strong fibre as they spontaneously forms strong coiled coils which can be reinforced with crosslinks such as lysine crosslinks. This contrasts with *B. mori* and spider recombinant silk polypeptides which require sophisticated spinning techniques in order to obtain the secondary structure (β-sheet) and strength of the fibre.

However, fibres may be spun from solutions having properties characteristic of a liquid crystal phase. The fibre concentration at which phase transition can occur is dependent on the composition of a protein or combination of proteins present in the solution. Phase transition, however, can be detected by monitoring the clarity and birefringence of the solution. Onset of a liquid crystal phase can be detected when the solution acquires a translucent appearance and registers birefringence when viewed through crossed polarizing filters.

In one fibre-forming technique, fibres can first be extruded from the protein solution through an orifice into methanol, until a length sufficient to be picked up by a mechanical means is produced. Then a fibre can be pulled by such mechanical means through a methanol solution, collected, and dried. Methods for drawing fibres are considered well-known in the art.

Further examples of methods which may be used for producing silk filaments and/or fibres of the invention are described in US 2004/0170827 and US 2005/0054830. In another example, the silk polypeptides are used to produce silk dope as decribed in WO 2011/022771.

In a preferred embodiment, one or more polypeptides of the invention are crosslinked. In one embodiment, the one or more polypeptides are crosslinked to a surface/article/product etc. of interest using techniques known in the art. Alternatively or additionally, at least some polypeptides in the silk filaments and/or fibres are crosslinked to each other. The crosslinks can comprise any suitable covalent bond or bonds. Thus, the polypeptides may be covalently crosslinked. For example, the polypeptides can be crosslinked via lysine residues, tyrosine residues, carboxyl groups, or any other suitable amino acid or functional group thereof. Preferably, the polypeptides are crosslinked via lysine residues in the proteins. Such crosslinking can be performed using chemical and/or enzymatic techniques known in the art. For example, enzymatic cross links can be catalysed by lysyl oxidase, whereas nonenzymatic cross links can be generated from glycated lysine residues (Reiser et al., 1992).

Antibodies

The present disclosure also provides antibodies to polypeptides and/or compositions of the invention or fragments thereof. Thus, the present disclosure further provides a process for the production of monoclonal or polyclonal antibodies to the polypeptides and/or compositions of the invention.

The term "binds specifically" refers to the ability of the antibody or fragment thereof to bind to at least one polypeptide and/or composition of the invention but not other known silk proteins. For example, the term "binds specifically" can be taken to mean that the antibody or fragment thereof reacts or associates more frequently, more rapidly, with greater duration and/or with greater affinity with a polypeptide and/or composition of the invention than it does with another known silk protein. In this regard, the degree of greater affinity, avidity, more readily, and/or with greater duration will depend on the application of the antibody or fragment. For example, for detection purposes the degree of specificity should be sufficiently high to permit quantification (where required). It is also to be understood by reading this definition that the term "binds specifically" does not necessarily require exclusive binding or non-detectable binding of another molecule, this is encompassed by the term "selective binding". Generally, but not necessarily, reference to binding means specific binding.

As used herein, the term "epitope" refers to a region of a polypeptide which is bound by the antibody. An epitope can be administered to an animal to generate antibodies against the epitope, however, antibodies of the invention preferably specifically bind the epitope region in the context of the entire polypeptide.

If polyclonal antibodies are desired, a selected mammal (e.g., mouse, rabbit, goat, horse, etc.) is immunised with an immunogenic polypeptide of the invention. Serum from the immunised animal is collected and treated according to known procedures. If serum containing polyclonal antibodies contains antibodies to other antigens, the polyclonal antibodies can be purified by immunoaffinity chromatography. Techniques for producing and processing polyclonal antisera are known in the art. In order that such antibodies may be made, the present disclosure also provides polypeptides, compositions, or fragments thereof haptenised to another polypeptide for use as immunogens in animals.

Monoclonal antibodies directed against the polypeptides and/or compositions of the invention can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies by hybridomas is well known. Immortal antibody-producing cell lines can be created by cell fusion, and also by other techniques such as direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. Panels of monoclonal antibodies produced can be screened for various properties; i.e., for isotype and epitope affinity.

An alternative technique involves screening phage display libraries where, for example the phage express scFv fragments on the surface of their coat with a large

variety of complementarity determining regions (CDRs). This technique is well known in the art.

For the purposes of the present disclosure, the term "antibody", unless specified to the contrary, includes fragments of whole antibodies which retain their binding activity for a target antigen. Such fragments include Fv, F(ab') and $F(ab')_2$ fragments, as well as single chain antibodies (scFv). Furthermore, the antibodies and fragments thereof may be humanised antibodies, for example as described in EP-A-239400.

Antibodies of the invention may be bound to a solid support and/or packaged into kits in a suitable container along with suitable reagents, controls, instructions and the like.

Preferably, antibodies of the invention are detectably labeled. Exemplary detectable labels that allow for direct measurement of antibody binding include radiolabels, fluorophores, dyes, magnetic beads, chemiluminescers, colloidal particles, and the like. Examples of labels which permit indirect measurement of binding include enzymes where the substrate may provide for a coloured or fluorescent product. Additional exemplary detectable labels include covalently bound enzymes capable of providing a detectable product signal after addition of suitable substrate. Examples of suitable enzymes for use in conjugates include horseradish peroxidase, alkaline phosphatase, malate dehydrogenase and the like. Where not commercially available, such antibody-enzyme conjugates are readily produced by techniques known to those skilled in the art. Further exemplary detectable labels include biotin, which binds with high affinity to avidin or streptavidin; fluorochromes (e.g., phycobiliproteins, phycoerythrin and allophycocyanins; fluorescein and Texas red), which can be used with a fluorescence activated cell sorter; haptens; and the like. Preferably, the detectable label allows for direct measurement in a plate luminometer, e.g., biotin. Such labeled antibodies can be used in techniques known in the art to detect the polypeptides of the invention.

Compositions

Also provided is a composition comprising one or more silk polypeptides of the invention. In one embodiment, the composition comprises a single type of silk polypeptide of the invention, namely each silk polypeptide in the composition has the same sequence. In another embodiment, the composition comprises two or more different silk polypeptides of the invention. For example, the composition may comprise two or more different homodimers or heterodimers as described herein.

When the composition comprises two or more polypeptides of the invention, the two or more polypeptides may be linked together by non-covalent or covalent bonds. Thus, the coiled coils may comprise two or more polypeptides held together by Van der Waals forces. Alternatively, or in addition, the two or more polypeptides may be crosslinked together by covalent bonds. For example, one or more cysteine residues in each polypeptide may be linked together by a disulfide bond. Accordingly, the composition of the invention may comprise one or more disulfide bonds between two or more polypeptides which associate to form a coiled coil along at least a portion of their length. Alternatively, the two or more polypeptides may be crosslinked by any other suitable covalent bonds as will be understood by a person skilled in the art.

The composition may comprise additional polypeptides other than those polypeptides of the invention. For example, the additional polypeptides may form an amorphous matrix which can support the coiled coil structures formed by the polypeptides of the invention.

Preferably, the composition does not include one or more polypeptides, other than those of the present invention, which are present in naturally occurring mantis oothecae. For example, the composition may not include any one or more of the proteins present in mantis oothecae other than the mantis fibroin polypeptides of the invention. Thus, the composition may not include any one or more of the proteins identified by the GenBank accession numbers JQ421298 (SEQ ID NO: 46), JQ421299 (SEQ ID NO: 47), JQ421300 (SEQ ID NO: 48), JQ421301 (SEQ ID NO: 49), JQ421302 (SEQ ID NO: 50), JQ421303 (SEQ ID NO: 51), JQ421304 (SEQ ID NO: 52), JQ421305 (SEQ ID NO: 53) and JQ421306 (SEQ ID NO: 54) in Table 2. Accordingly, the composition may not include any one or more of the proteins encoded by the nucleotide sequences identified by the GenBank accession numbers JQ421298 (SEQ ID NO: 37), JQ421299 (SEQ ID NO: 38), JQ421300 (SEQ ID NO: 39), JQ421301 (SEQ ID NO: 40), JQ421302 (SEQ ID NO: 41), JQ421303 (SEQ ID NO: 42), JQ421304 (SEQ ID NO: 43), JQ421305 (SEQ ID NO: 44) and JQ421306 (SEQ ID NO: 45). In addition, or alternatively, the composition may not include any other proteins identified in naturally occurring mantis ootheca.

The compositions of the invention may further comprise an "acceptable carrier". Examples of such acceptable carriers include water, saline, Ringer's solution, dextrose solution, Hank's solution, and other aqueous physiologically balanced salt solutions. Nonaqueous vehicles, such as fixed oils, sesame oil, ethyl oleate, or triglycerides may also be used.

The "acceptable carrier" may be a "pharmaceutically acceptable carrier" and/or an "agriculturally acceptable carrier". The term pharmaceutically acceptable carrier refers to molecular entities and compositions that do not produce an allergic, toxic or otherwise adverse reaction when administered to an animal, particularly a mammal, and more particularly a human. Useful examples of pharmaceutically acceptable carriers or diluents include, but are not limited to, solvents, dispersion media, coatings, stabilizers, protective colloids, adhesives, thickeners, thixotropic agents, penetration agents, sequestering agents and isotonic and absorption delaying agents that do not affect the activity of the polypeptides of the invention. The proper fluidity can be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. More generally, the polypeptides of the invention can be combined with any non-toxic solid or liquid additive corresponding to the usual formulating techniques. The term agriculturally acceptable carrier refers to molecular entities and compositions that do not damage plants or act as pollutants or harmful environmental contaminants when applied in an agricultural setting. Suitable agriculturally acceptable carriers will be known to a person skilled in the art.

As outlined herein, in some embodiments a polypeptide and/or product of the invention is itself used as a pharmaceutically acceptable carrier. Thus, the composition may further comprise an active agent such as a drug or a cosmetic agent. The composition may therefore be provided for use as a medicine, in a medical device, and/or as a cosmetic.

Other suitable compositions are described herein with reference to specific uses of the polypeptides in particular products.

Uses

Silk proteins are useful for the creation of new biomaterials because of their exceptional toughness and strength. However, to date the fibrous proteins of spiders and insects are large proteins (over 100kDa) and consist of highly repetitive amino acid sequences. These proteins are encoded by large genes containing highly biased codons making them particularly difficult to produce in recombinant systems. By comparison, the silk proteins of the invention are short and non-repetitive. These properties make the genes encoding these proteins particularly attractive for recombinant production of new biomaterials.

The present disclosure therefore provides a product comprising one or more polypeptides, materials and/or compositions of the invention. The polypeptides and/or

materials may be present in any structural form described herein, such as in the form of one or more filaments, fibres, crystallites, etc. The particular form may be selected according to the desired properties of the product.

The silk proteins, materials, and/or products of the invention can be used for a broad and diverse array of medical, military, industrial and commercial applications. Silk fibres 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 proteins, materials, and/or products of the invention also have applications in compositions for personal care products such as cosmetics, skin care, hair care and hair colouring; and in coating of particles, such as pigments.

The silk proteins may be used in their native form or they may be modified to form derivatives, which provide a more beneficial effect. For example, the silk protein 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 silk proteins 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".

Silk fibres of the invention 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 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 protein 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 protein 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 at least one silk protein, material, and/or product of the invention include, 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 protein in a cosmetically acceptable medium. An effective amount of a silk protein 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 protein in a cosmetically acceptable medium. An effective amount of a silk protein, material, and/or product of the invention 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 protein, material, and/or product of the invention in a cosmetically acceptable medium. An effective amount of a silk protein 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 colouring 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 silk proteins, silk proteins, materials, and/or products of the invention 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 colour pigments, organic pigments, and inert particles. The inorganic colour 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 silk proteins, materials, and/or products of the invention 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 silk proteins, materials, and/or products of the invention may also be used in soap (see, for example, US 2005/0130857).

Pigment and Cosmetic Particle Coating

The effective amount of a silk protein, material, and/or product of the invention 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 silk protein, material, and/or product of the invention to be used depends on the type of pigment or cosmetic particle being coated. For example, the amount of silk protein, material, and/or product of the invention used with inorganic colour pigments is preferably between about 0.01% and 20% by weight. In the case of organic pigments, the preferred amount 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 protein, material, and/or product of the invention to the particles while tumbling or mixing, forming a slurry of the silk protein, material, and/or product and the particles and drying, spray drying a solution of the silk protein, material, and/or product of the invention onto the particles or lyophilizing a slurry of the silk protein, material, and/or product of the invention and the particles. These coated pigments and cosmetic particles may be used in cosmetic formulations, paints, inks and the like.

Biomedical

The silk proteins, materials, and/or products of the invention 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 proteins, materials, and/or products. For the purpose of a wound-healing bandage, an effective amount of silk protein, material, and/or product the invention 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 protein, material, and/or product. Alternatively, the solution containing silk protein, material, and/or product may be sprayed onto the surface of the material to be coated using a spray gun. Additionally, the solution containing the silk proteins, materials, and/or products 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 proteins, materials, and/or products of the invention may also be used in the form of a film as a wound dressing material. The use of silk proteins, 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 protein, material, and/or product. For a film for wound care, an effective amount of silk protein, material, and/or product 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 proteins, materials, and/or products of the invention may also be used in 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 proteins, materials, and/or products of the invention may also be used in 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 disclosure, 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 proteins, materials, and/or products of the invention may also be used in 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 disclosure include cartilage (both articular and meniscal), bone, muscle, skin and blood vessels.

The silk proteins, materials, and/or products of the invention may also be used in 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 proteins, materials, and/or products may also be used in dermatological compositions (see, for example, US 2005/0019297). Furthermore, the silk proteins, materials, and/or products and derivatives thereof may also be used in sustained release compositions (see, for example, US 2004/0005363).

Textiles

The silk proteins, materials, and/or products of the invention may also be applied to 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 protein, material, and/or product. For the purpose of fibre coating for use in textiles, an effective amount of silk protein, material, and/or product 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 protein 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 protein-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 protein composition onto woven fibres. An ideal embodiment of this application is the use of silk proteins to coat stretchable weaves such as used for stockings.

Composite Materials

The silk proteins, materials, and/or products of the invention 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 of the invention 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 which can include silk proteins of the invention are generally described in JP 2000139755.

Advanced Materials

Silks of the present disclosure have considerable toughness and stand 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 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 - Mantis ootheca consists mostly of two proteins, Mantis Fibroin 1 and Mantis Fibroin 2

Characterisation of mantis ootheca structural components has been hindered by a number of significant experimental difficulties. For example, the rigid structural nature of the solid ootheca impedes the performance of mass spectroscopy. Even the addition of proteases is often not sufficient to separate the structural proteins so as to allow their complete characterisation by mass spectroscopy, due to the extent of aggregation of the structural components in the insoluble ootheca. In addition, given that only a small sample of proteins may be derivable from the insoluble ootheca, the relative abundance of those proteins in the entire ootheca (and therefore, the importance of those proteins in determining the oothecal structure) cannot clearly be determined.

The present inventors attempted to overcome the problems associated with obtaining samples from insoluble oothecae by obtaining samples of ootheca that can be solubilised. However, obtaining ootheca that can be solubilised presents another challenge, as it must be obtained during ootheca production. This requires collecting both female and male mantises of each species from the wild, housing them individually to avoid cannibalism, and feeding them every 1-2 days (which itself requires a prey species to be either maintained in culture or purchased regularly).

Females that have not mated must then be mated with males in a large enclosure after feeding the female to excess, in order to avoid cannibalism before fertilisation. A fertilised female fed regularly normally produces an ootheca every 2-5 weeks, and often at night, so it takes some time and diligence to catch a female during ootheca production.

These experimental difficulties have no doubt contributed to the lack of further characterisation of ootheca components to date. The present examples demonstrate the inventors' attempts to overcome these difficulties.

Females of three species of mantises (*Pseudomantis albofimbriata*, *Tenodera australasiae*, and *Archimantis monstrosa*) were collected from Canberra (ACT, Australia) and fed field crickets (*Acheta domestica*, Pisces Live Food, Canberra, Australia). Solubilised ootheca was prepared by collecting wet ootheca material as it was produced during oviposition and immediately solubilising it in 6M guanidinium hydrochloride with protease inhibitors (Complete mini EDTA-free, Roche) and 5 mM EDTA. Guanidinium hydrochloride was removed from solubilised samples by repeated concentration with a Centricon-10 centrifugal filter device (Millipore, Belerica, USA) and dilution with 0.1% sodium dodecyl sulphate (SDS). Solubilised ootheca proteins were separated by polyacrylamide gel electrophoresis (PAGE) using NuPage 4-12% Bis-Tris gels and 2-(N-morpholino)ethanesulphonic acid running buffer (Invitrogen, Carlsbad, USA) and stained with Coomassie Brilliant Blue (Sigma-Aldrich, St. Louis, USA).

Five to seven protein bands were identified using SDS-PAGE, depending on the species investigated (Figure 1). The dominant bands migrated between the 40-60 kDa marker bands, and other proteins were present at lower abundance. Solutions generated from *P. albofimbriata* oothecae contained an intense band at 48 kDa, and weak bands at 30, 58, 80, and 95 kDa, whereas those from *T. australasiae* contained intense bands at 44 and 50 kDa, a medium band at 58 kDa, and weak bands at 25, 30, 33 and 90 kDa (Figure 1).

EXAMPLE 2 - Identification of Mantis ootheca proteins by liquid chromatography/mass spectrometry

Females of three species of mantises (*Pseudomantis albofimbriata*, *Tenodera australasiae*, and *Archimantis monstrosa*) were collected from Canberra (ACT, Australia) and fed field crickets (*Acheta domestica*, Pisces Live Food, Canberra, Australia). Females were dissected in phosphate buffered saline, pH 7.0. Glands for protein preparations were stored in 2.5% SDS with protease inhibitors (Complete mini

EDTA-free, Roche, Basel, Switzerland) and 5 mM EDTA. Glands from which cDNA libraries were constructed were stored in RNAlater (Ambion, Austin, USA), at -80°C.

Total RNA was prepared from primary collaterial glands using RNAqueous-4PCR (Ambion, Austin, USA) and mRNA isolated using Micro-FastTrack 2.0 (Invitrogen, Carlsbad, USA). Libraries of cDNAs were constructed using Cloneminer II cDNA Library Construction kit (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions. Clones containing cDNA inserts larger than 300 bp were sequenced by Micromon Services (Monash University, Melbourne, Australia) using an Applied Biosystems 3730S Genetic Analyser and Applied Biosystems PRISM BigDye Terminator Mix cycling chemistry. A library of possible protein sequences was generated.

Individual protein bands from SDS-PAGE gels of solubilised oothecae (Figure 1) were excised, digested with trypsin (Sigma-Aldrich, St. Louis, USA) and used for liquid chromatography/mass spectroscopy (LC-MS) according to previously published methods (Sutherland et al., 2006). Ootheca proteins were identified by matching the size of tryptic fragments identified by LC-MS with *in silico* digests of proteins encoded by mantis cDNAs using SpectrumMill (Agilent, Santa Clara, USA). This process led to the confident identification of five *P. albofimbriata* ootheca proteins and eight *T. australasiae* ootheca proteins (Table 2).

EXAMPLE 3 - Two proteins, Mantis Fibroin 1 and Mantis Fibroin 2, are the main structural proteins in mantis ootheca

The strength of protein bands on the SDS-PAGE gels (Figure 1) and the copy numbers of sequences in the cDNA libraries were analysed to draw conclusions about which were the main structural proteins in mantis ootheca.

The most abundant sequences in the cDNA libraries from each species encoded two proteins we called Mantis Fibroin 1 (*P. albofimbriata*: SEQ ID NO: 5; *T. australasiae*: SEQ ID NO: 1; *A. monstrosa*: SEQ ID NO: 3 and allelic variant described in Table 3) and Mantis Fibroin 2 (*P. albofimbriata*: SEQ ID NO: 6 and allelic variant described in Table 3; *T. australasiae*: SEQ ID NO: 2 and allelic variant described in Table 3; *A. monstrosa*: SEQ ID NO: 4). The remainder of each library consisted mostly of single-copy sequences. For *P. albofimbriata* Mantis Fibroin 1 and Mantis Fibroin 2 account respectively for 24.6% and 24.6% of sequences, for *T. australasiae* 31.5% and 13.1%, and for *A. monstrosa* 40.7% and 38.4%. The third and fourth most abundant sequences are much less common (*P. albofimbriata*, 10.5% and 3.5%; *T. australasiae*, 5.4% and 4.3%; *A. monstrosa*, 2.3% and 2.3%). The proportion of sequences encoding

Mantis Fibroin 1 and Mantis Fibroin 2 were close to equally abundant in the *P. albofimbriata* and *A. monstrosa* libraries, while sequences encoding Mantis Fibroin 1 are more common than those encoding Mantis Fibroin 2 in the *T. australasiae* library.

Table 2: Identification of praying mantis egg case proteins.

Sequence (protein name or GenBank accession number)	Abundance in library (%)	SpectrumMill score[a] (# peptides)	Database homology	Protein band intensity[b]	MW on gel (kDa)[b]	predicted MW[c]
Pseudomantis albofimbriata:						
Fibroin 1	24.6	78.1 (5)	nd	high	48	41.2
Fibroin 2	24.6	119.2 (10)	nd	high	48	40.1
JQ421298	10.5	71.1 (6)	nd	very low	30	nc
JQ421299	3.5	79.0 (7)	yes[d]	low	58	nc
JQ421300	1.8	25.2 (2)	nd	high	48	nc
Tenodera australasiae:						
Fibroin 1	31.5	226.9 (16)	nd	high	50	41.3
Fibroin 2	13.1	96.3 (8)	nd	high	44	39.7
JQ421303	5.4	45.8(4)	nd	very low	25	20.5
JQ421301	4.3	127.4 (10)	yes[d]	very low	25	22.4
JQ421305	4.3	149.1 (12)	nd	very low	30	nc
JQ421302	3.3	94.6 (8)	nd	very low	33	28.6
JQ421304	2.2	23.4 (2)	nd	high	50	41.1
JQ421306	1.1	47.3 (3)	yes[e]	medium	58	nc
Archimantis monstrosa:						
Fibroin 1	40.7	-	nd	-	-	40.7
Fibroin 2	38.4	-	nd	-	-	40.5

[a]Spectrum Mill scores <20 are not considered a confident identification, [b]See Figure 1, [c]Without signal peptide, [d]cuticular protein, [e]catylase, nd: not detected; nc: could not be calculated because only a partial sequence was available.

P. albofimbriata Mantis Fibroin 1 and Mantis Fibroin 2 were both detected in the most intense band, at 48 kDa. The product of JQ421300 was also detected in the 48 kDa band, but JQ421300 was only present as a single sequence in the library. Also, only two peptides from the product of JQ421300 were detected from this band, compared to five from Mantis Fibroin 1 and ten from Mantis Fibroin 2 respectively, suggesting the product of JQ421300 is present only at very low levels. JQ421298 is common in the *P. albofimbriata* library (10.5%), but its product is unlikely to be a major structural protein as it was found to correspond only to the very weak 30 kDa band on the protein gel.

For *T. australasiae*, Mantis Fibroin 1 corresponded to the 50 kDa band and Mantis Fibroin 2 to the 44 kDa band, the most intense bands on the gel. The product of JQ421304 was also detected in the 50 kDa band, but only two copies of JQ421304 were detected in the library, and only two peptides were detected from its product compared to 16 from Mantis Fibroin 1.

For *A. monstrosa*, sequences encoding Mantis Fibroin 1 and Mantis Fibroin 2 make up 79.1% of the entire library. Thus, all our results suggest that Mantis Fibroin 1 and Mantis Fibroin 2 are the principle structural proteins in mantis ootheca.

EXAMPLE 4 - Conserved sequence features of Mantis Fibroin 1 and Mantis Fibroin 2 in subfamily Mantinae

The inventors analysed the sequences of Mantis Fibroin 1 and Mantis Fibroin 2 in terms of their divergence, amino acid composition, and conserved sequence features.

Despite having similar length and composition, Mantis Fibroin 1 and Mantis Fibroin 2 do not appear to be derived from a common ancestor and no convincing primary sequence alignment between them could be found. For each of *T. australasiae* Mantis Fibroin 2, *P. albofimbriata* Mantis Fibroin 2, and *A. monstrosa* Mantis Fibroin 1, there were two groups of sequences in our libraries differing by 1-2 single nucleotide polymorphisms, and present in approximately equal numbers. Since each of our libraries was constructed from a single female, these results are consistent with the two groups corresponding to transcripts from maternal and paternal alleles. Allelic differences in mantis fibroins are reported in Table 3.

Mantis Fibroin 1 has a predicted signal peptide, of 22 residues. Depending on species the mature protein comprises 361-365 residues with a predicted molecular weight of 40.7-41.3 kDa. The amino acid composition of Mantis Fibroin 1 in all species is similar, the most abundant amino acids in the mature sequence being ala (12.9-15.9), glu (11.0-12.7), ser (8.5-10.5), and lys (8.5-9.3).

Table 3: Allelic differences in mantis fibroin nucleotide and amino acid sequences.

Species	Alleles	SNP	Amino acid subsitution
Tenodera australasiae	Mantis Fibroin 1	-	-
	Mantis Fibroin 2a	228T, 772G	silent, 258Ala
	Mantis Fibroin 2b	228C, 772A	silent, 258Thr
Pseudomantis albofimbriata	Mantis Fibroin 1	-	-
	Mantis Fibroin 2a	353C	118Ser
	Mantis Fibroin 2b	353A	118Tyr
Archimantis monstrosa	Mantis Fibroin 1a	15A	5Ile
	Mantis Fibroin 1b	15G*	5Met
	Mantis Fibroin 2	-	-

A protein alignment (Figure 2) of Mantis Fibroin 1 from the three species generated by ClustalW indicated strong sequence alignment of the signal peptide, moderate conservation of the alpha helical region and low or no sequence conservation at the N- and C-termini of the mature protein. The overall pairwise identity was 45-51% and the overall pairwise similarity 71-76% (Table 4). The position of the cysteine residues was not conserved.

Mantis Fibroin 2 was predicted to have an 18 residue signal peptide, leaving a mature sequence of 342-351 residues with a predicted molecular weight of 39.7-40.5 kDa. Mantis Fibroin 2 contained 10.2-13.9% ala, 13.1-13.5% glu, and 9.7-10.2% lys. The location of cysteines was not conserved. Low or no conservation was found in the N-terminal region of the mature protein (Figure 3). The overall identity of Mantis Fibroin 2 between species was 71-77% and the overall similarity 49-57% (Table 5).

Table 4: Divergence of mantis fibroin sequences. Values given are percent amino acid identity, with percent amino acid character, as calculated using the GONNET matrix, shown in brackets. Values on the upper right are for Mantis Fibroin 1 and values on the lower left for Mantis Fibroin 2.

	Pseudomantis	*Tenodera*	*Archimantis*	
Pseudomantis		47(76)	51(75)	Mantis Fibroin 1
Tenodera	53(74)		45(71)	
Archimantis	57(77)	49(71)		
	Mantis Fibroin 2			

EXAMPLE 5 - Mantis fibroins are predicted to have large coiled coil domains and to form dimers

The secondary and tertiary structure of mantis fibroins was investigated using bioinformatic prediction programs. Secondary structure was predicted using GOR4 (Combet et al., 2000). Coiled coil domains were predicted using MARCOIL (Delorenzi and Speed, 2002). Oligomerisation states were predicted using PrOCoil (Mahrenholz et al., 2011).

The secondary structure prediction program GOR4 (Combet et al., 2000) predicts that the majority of Mantis Fibroin 1 (71.6-79.2%, depending on species) forms alpha-helices, with the remainder at each end and some short regions between them predicted to form random coils (19.1-24.5%) and beta-sheets (1.7-7.0%). MARCOIL (Delorenzi and Speed, 2002) predicted a 23-heptad coiled coil region stretching approximately over residues 40-200 of the mature protein. The region from residues 200-330 was predicted to be mostly alpha-helical but not coiled coil (Figure 4A).

Mantis Fibroin 2 was predicted to have many structural similarities to Mantis Fibroin 1 (Figure 4B), with the majority of the sequence predicted to form alpha-helices (72.9-77.5%), with lesser amounts of random coils (16.7-20.2%) and beta-sheets (4.3-8.1%). The coiled coil domain is set slightly further back in the sequence, spanning approximately from residues 80-240. Residues from 30-80 and 240-330 were mostly predicted to be alpha-helical but not to form coiled coils. Similar predictions were obtained for Mantis Fibroin 1 and Mantis Fibroin 2 from *Archimantis* and *Tenodera*, though the exact boundaries of the coiled coil domain varied, and the coiled coil domain in *Tenodera* Mantis Fibroin 2 became low around the region of residue 206.

PrOCoil (Mahrenholz et al., 2011) is a support vector machine algorithm that classifies sequences either as dimer-forming or trimer-forming depending on pairwise correlations of residues. PrOCoil classified Mantis Fibroin 1 and Mantis Fibroin 2 from each of the three species as dimers, consistent with previous reports that crystallites in mantis ootheca are made up of dimeric coiled coils.

EXAMPLE 6 - Mantis fibroins coiled coil domains have an unusual core composition

Residues for which a heptad position was predicted with over 50% confidence by MARCOIL were used to investigate patterns in amino acid properties compared to heptad position. As similar trends were observed for each species, data for the three species for each of Mantis Fibroin 1 and Mantis Fibroin 2 were combined (Figures 5 and 6).

As expected, residues in positions *a* and *d* were on average significantly more hydrophobic than residues in other positions ($p<0.05$; Figure 5A). Approximately half of the residues in the *a* position of Mantis Fibroin 1 and Mantis Fibroin 2 were alanine (Figure 6) which has resulted in a much shorter average side-chain length than the side chain length in other positions ($p<0.03$; Figure 5B). The *d* position of the core contained a low proportion of alanine residues, its hydrophobicity being due to a high proportion of aromatic residues, principally tyrosine (Figure 6). The average charge of residues at neutral pH was also calculated. Residues in position *g* ($p<0.01$) and to a lesser extent *b* ($p<0.05$) had an average positive charge, while residues in *c* positions had an average negative charge ($p<0.01$).

Details of particular amino acid residues and number of heptad repeats in the coiled coil domains of the mantis fibroin proteins compared to core components and number of heptad repeats in the coiled coil domains of the *G. gallus* alpha-tropomyosin protein, illustrating the unique features of the mantis fibroin proteins, are shown in Table 5.

EXAMPLE 7 - Mantis fibroin primary sequences are unique

The inventors compared mantis fibroin primary sequences to protein sequence databases and classes of proteins with similar properties. They were unable to find any homology between mantis fibroins and any sequence in Genbank's non-redundant protein sequence database or the translated non-redundant nucleotide sequence database.

Table 5: Mantis Fibroin coiled coil composition compared to *G. gallus* alpha-tropomyosin coiled coil composition.

Species	Protein	% Ala in (a) position	% Aromatics in (d) position (FWYHP)	Heptads in CC domain*
Mantis *P. albofimbriata*	Fibroin 1	55.5	42	19
	Fibroin 2	47.8	26.1	23
Mantis *T. australasiae*	Fibroin 1	66.6	31.6	19
	Fibroin 2	47.8	30.4	23
Mantis *A. monstrosa*	Fibroin 1	33.3	47.4	19
	Fibroin 2	47.8	43.4	23
Chicken *G. gallus*	alpha-tropomyosin	20	7.5	40

Mantis fibroins appear to be the only examples of proteins with long coiled coil domains incorporating a significant proportion of aromatic residues in the hydrophobic core. While the frequency of aromatic residues in the core of previously described coiled coil proteins overall is almost zero (Woolfson, 2005), approximately one third of residues in the *d* position of mantis fibroins are aromatic. Since aromatic residues are metabolically expensive to produce, and a large quantity of protein is used to make an ootheca, the retention of aromatic residues in significant numbers in mantis fibroins indicates a functional role for them.

The inventors conducted a search of the literature for any coiled coils that systematically included aromatic residues in the core. Among natural proteins, the only class of proteins identified were the ferritin-class Alacoils (Gernert, 1995). Three of the five original examples of this class feature aromatics at the *d* position, but the longest of these spans five heptads, whereas the coiled coil domains in mantis fibroins are predicted to span approximately 23 heptads.

Thus, mantis fibroins are distinct from any other fibrous coiled coil proteins due to the composition of their hydrophobic core, and distinct from any of the proteins described as Alacoils due (at least) to the length of their coiled coil domains.

EXAMPLE 8 - Mantis fibroins can be expressed in *E. coli*

Recombinant expression constructs (without the signal peptides) were made by PCR amplification of full-length sequences encoding mantis fibroins. The primers used were as follows, with NcoI and BamHI sites underlined, and coding sequences italicised:

T. australiasiae Mantis Fibroin 1, GGA ATT CCC ATG GGC *TCT CCC TTG GAA GAC AAA TAC* (SEQ ID NO: 25) and C GGC GGA TCC TTA TTA *CAG ACC TTC GCC GGA AC* (SEQ ID NO: 26);

T. australiasiae Mantis Fibroin 2, GGA ATT CCC ATG *GGC AAG AAA CAT GAAG TAA TGA* (SEQ ID NO: 27) and C GGC GGA TCC TTA TTA *TCC GTG GTA GTT GGA GTG* G (SEQ ID NO: 28);

A. monstrosa Mantis Fibroin 1, GGA ATT CCC ATG GGC *TCT CCC TTG GAA GAA AAA TAT G* (SEQ ID NO: 29) and C GGC GGA TCC TTA TTA *ACT CAT TCC TTC ACC TTC AGT T* (SEQ ID NO: 30);

A. monstrosa Mantis Fibroin 2, GGA ATT CCC ATG *GGC AAG AAA CAC GAA GCA* (SEQ ID NO: 31) and C GGC GGA TCC TTA TTA *TGC TCC GTG GTA GTT GGA* (SEQ ID NO: 32);

P. albofimbriata Mantis Fibroin 1, GGA ATT CCC ATG GGC *TCA CCC TTG GAA GAA AAA TAT* (SEQ ID NO: 33) and C GGC GGA TCC TTA TTA *TTC ATC GCC GTA AGA CAT TT* (SEQ ID NO: 34);

P. albofimbriata Mantis Fibroin 2, GGA ATT CCC ATG *GGC AAG AAT CAC GAA GTA ATG* (SEQ ID NO: 35) and C GGC GGA TCC TTA TTA *TGC TCC GTG GTA GTT GGA G* (SEQ ID NO: 36).

PCR amplicons were then digested with Nco1 and BamH1 (New England Biosciences) and cloned into pET14b vector (Novagen) using T4 DNA ligase (New England Biosciences). Successful cloning was verified by restriction digest and DNA sequencing.

E. coli Rosetta 2 DE3 cells (Novagen) were transformed with the ligated vectors and allowed to grow on LB plates containing 100 µg/ml ampicillin and 34 µg/ml chloramphenicol overnight. Several colonies from each plate were picked into 5 ml Overnight Express Autoinduction media (Merck) contained in a 50 ml Falcon tube. Cultures were grown overnight on a 200 rpm shaker at 37°C. Cultures were pelleted by centrifugation at 13000 rpm for 10 minutes and stored at -80°C for later use.

To prepare protein solutions, pellets were resuspended in 1 ml BugBuster Master Mix (Novagen) and incubated at 10 minutes on an axial rotator. Inclusion bodies were separated from the soluble fraction by centrifugation and resuspended in 0.2% SDS. Material not soluble in 1 ml 0.2% SDS was pelleted by centrifugation and resuspended in 1 ml 0.5% SDS. Material not soluble in 1 ml 0.5% SDS was pelleted by centrifugation and resuspended in 1 ml 1.0% SDS. Expression of mantis fibroins was strong in the inclusion bodies (Figure 7) but low or absent in the soluble fraction.

EXAMPLE 9 - Mantis Fibroins expressed in *E. coli* can form coiled coils in solution

The secondary and tertiary structures of mantis fibroins expressed in *E. coli* were examined using circular dichroism spectroscopy. Mantis fibroin solutions in 1.0% SDS were dialysed against 0.03% SDS using 10K MWCO Slide-A-Lyzer cassettes (Pierce, Illinois, USA). Far ultraviolet circular dichroism spectra of fibroin solutions between 180 and 300 nm were measured using an 0.2 mm path-length cell (20/O/Q/0.2, Starna, Atascadero, USA) in a Chirascan spectrometer (Applied Photophysics, Leatherhead, UK) at 25°C. Spectra were collected between 180-300 nm with 1 nm increments and an integration time of 0.5 seconds. Three spectra were averaged and the spectrum of a blank that contained 0.03% SDS but no protein was subtracted.

An example spectrum, of *T. australasiae* Mantis Fibroin 2b, is shown in Figure 8. The pattern consists of two negative peaks at 209 nm and 220 nm, which is typical of alpha-helices. The small broad positive peak at 265 nm is probably due to aromatic residues (Tanaka et al., 2006). The ratio of the circular dichroism at 220 nm to that at 209 nm exceeds unity (1.03), indicating that the majority of protein is present as coiled coils rather than isolated alpha-helices (Monera et al., 1993). This result indicates that mantis fibroins expressed in *E. coli* are capable of taking on the same tertiary structures that are present in natural solid ootheca (Rudall, 1956; Rudall, 1962).

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 61/615,745 filed 26 March 2013, the entire contents of whoch are incorporation herein by reference.

All publications discussed above 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

Britton et al. (1970). Insects of Australia: A text book for students and research workers. Carlton, Melbourne University Press.

Bullough and Tulloch (1990) J. Mol. Bio. 215: 161-173.

Christou and Klee (2004) Handbook of Plant Biotechnology, John Wiley and Sons.

Combet et al. (2000) Trends in Biochemical Sciences 25(3): 147-150.

Delorenzi and Speed (2002) Bioinformatics 18(4): 617-625.

Dal Pra et al. (2005) Biomaterials 26:1987-99.

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

Eisenberg (1984) Annual Review of Biochemistry 53: 595–623.

Gernert et al. (1995) Protein Sci. 4: 2252-2260.

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

Hepurn et al. (1979) Insect Biochem. 9:66.

Houdebine (1997) Transgenic animals – Generation and Use. Harwood Academic.

Jin et al. (2002) Biomacromolecules 3:1233-9.

Kallol et al. (2003) Journal of Chromatography A, 1000, 637–655.

Kenchington and Flower (1969) Journal of Microscopy 89: 263-281.

Kim et al. (2005) Biomaterials 26:2775-85.

Kramer et al. (1973) Insect Biochemistry 3: 297-302.

Kyte and Doolittle (1982) Journal of Molecular Biology 157 (1): 105–32.

Mahrenholz et al. (2011) Mol. Cell Proteomics 10: M110 004994.

Minoura et al. (1990) Biomaterials 11: 430-4.

Monera et al. (1993) Journal of Biological Chemistry 268: 19218-19227.

Nalepa and Lenz (2000) Proc. Bio. Sci. 267: 1809-1813.

Nazarov et al. (2004) Biomacromolecules 5:718-26.

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

O'Brien et al. (1994). Design, Synthesis and Fabrication of Novel Self-Assembling Fibrillar Proteins. Silk Polymers: Materials Science and Biotechnology. Kaplan, Adams, Farmer and Viney, American Chemical Society, Washington, D.C. pp. 104-117.

Parker and Rudall (1955) Biochim. Biophys. Acta 17: 287-287.

Rattew (1974). D. Phil. Thesis, University of Oxford.

Reiser et al. (1992) Nucleic Acids Research 32 (Web Server issue): W321-W326.

Rose and Wolfenden (1993) Annual Review of Biophysics and Biomolecular Structure 22, 381–415.

Rudall (1956) Lectures on the Scientific Basis of Medicine 5: 217-230.
Rudall (1962). Silk and other cocoon proteins. Comparative Biochemistry. M. Florkin and H. S. Mason. New York, Academic Press. 4: 397-433.
Rudall and Kenchington (1971) Annual Review of Entomology 16: 73-96.
Slater et al. (2003) Plant Biotechnology - The Genetic Manipulation of Plants, Oxford University Press.
Sutherland et al. (2006) Genome Res. 16:1414-1421.
Sutherland et al. (2010) Annu. Rev. Entomol. 55:171-88.
Tanaka et al. (2006) Chirality 18: 652-661.
Vepari and Kaplan (2007) Prog. Polym. Sci. 32:991-1007.
Woolfson et al. (2005) 70: 79-112.
Yago et al. (1984) Insect Biochemistry 14: 7-9.
Yago et al. (1990) Insect Biochemistry 20: 745-750.

CLAIMS

1. An isolated and/or recombinant silk polypeptide, wherein at least a portion of the polypeptide has a coiled coil structure comprising at least 10 copies of the heptad sequence *abcdefg*, and wherein at least 25% of the amino acids at position *a* are alanine residues and at least 20% of the amino acids at position *d* are aromatic residues.

2. The polypeptide of claim 1, wherein at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55% or at least 60% of the amino acids at position *a* are alanine residues and/or at least 25%, at least 30%, at least 35%, at least 40%, at least 45% or at least 50% of the amino acids at position *d* are aromatic residues.

3. The polypeptide of claim 1 or claim 2, wherein the aromatic residues are selected from the group consisting of tyrosine, phenylalanine, tryptophan, histidine and proline.

4. The polypeptide of any one of claims 1 to 3, wherein at least a portion of the polypeptide has a coiled coil structure comprising at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24 or at least 25 copies of the heptad sequence *abcdefg*.

5. The polypeptide of any one of claims 1 to 4, wherein the amino acids at position *a* and/or *d* are more hydrophobic on average than amino acids at other positions in the heptad sequence.

6. The polypeptide of any one of claims 1 to 5, wherein the average charge of amino acids at position *g* and/or *b* is positive at neutral pH, and/or wherein the average charge of amino acids at position *c* is negative at neutral pH.

7. The polypeptide of any one of claims 1 to 6, which comprises an amino acid sequence selected from:

i) an amino acid sequence as provided in any one of SEQ ID NOs 13-24;

ii) an amino acid sequence which is at least 30% identical to any one or more of SEQ ID NOs 13-24; or

iii) a biologically active fragment of i) or ii).

2013239320 02 Oct 2018

8. The polypeptide according to any one of claims 1 to 7, which is fused to at least one other polypeptide.

9. A composition comprising one or more silk polypeptides according to any one of claims 1 to 8.

10. The composition of claim 9, further comprising one or more acceptable carriers.

11. The composition of claim 9 or claim 10, further comprising a drug.

12. A vector comprising at least one polynucleotide which encodes a silk polypeptide according to any one of claims 1 to 8.

13. A host cell comprising at least one vector of claim 12.

14. A process for preparing a polypeptide according to any one of claims 1 to 8 or a composition according to any one of claims 9 to 11, the process comprising cultivating a host cell of claim 13, or a vector of claim 12, under conditions which allow expression of the polynucleotide encoding the polypeptide, and recovering the expressed polypeptide.

15. A transgenic non-human organism comprising a vector of claim 12 or a host cell of claim 13.

16. An antibody which specifically binds a polypeptide according to any one of claims 1 to 7.

17. A product comprising at least one polypeptide according to any one of claims 1 to 8 or a composition according to any one of claims 9 to 11.

18. The product of claim 17, wherein the product is selected from the group consisting of: a personal care product, textiles, plastics, and biomedical products.

2013239320 02 Oct 2018

19. A composition comprising one or more of i) at least one vector according to claim 12, ii) at least one host cell according to claim 13, and iii) at least one antibody of claim 16, and one or more acceptable carriers.

20. A method of treating or preventing a disease, the method comprising administering a composition comprising a drug for treating or preventing the disease and a pharmaceutically acceptable carrier, wherein the pharmaceutically acceptable carrier comprises at least one polypeptide according to any one of claims 1 to 8 and/or product of claim 17 or claim 18.

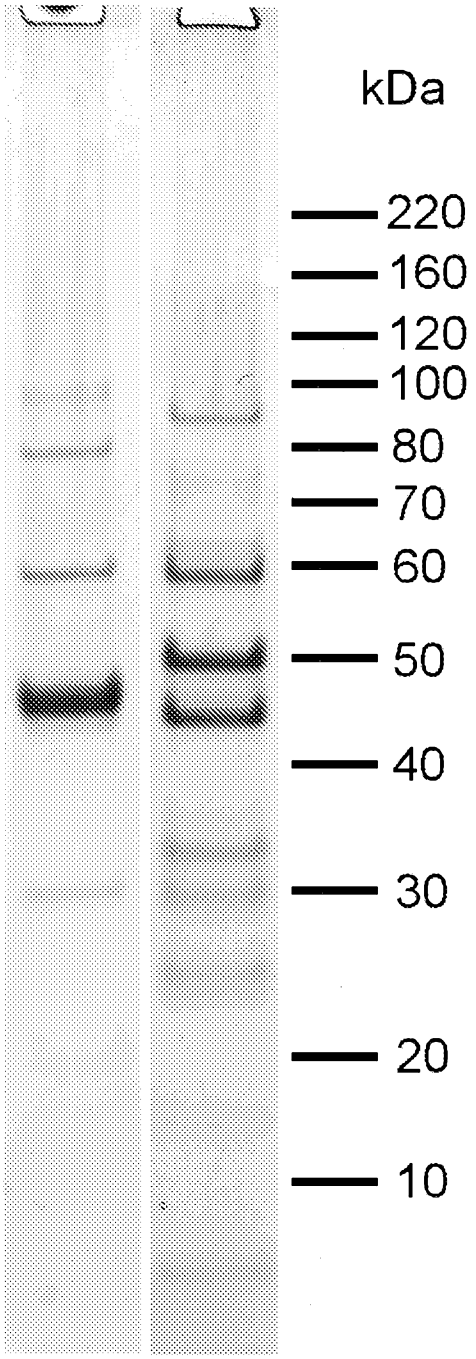

Figure 1

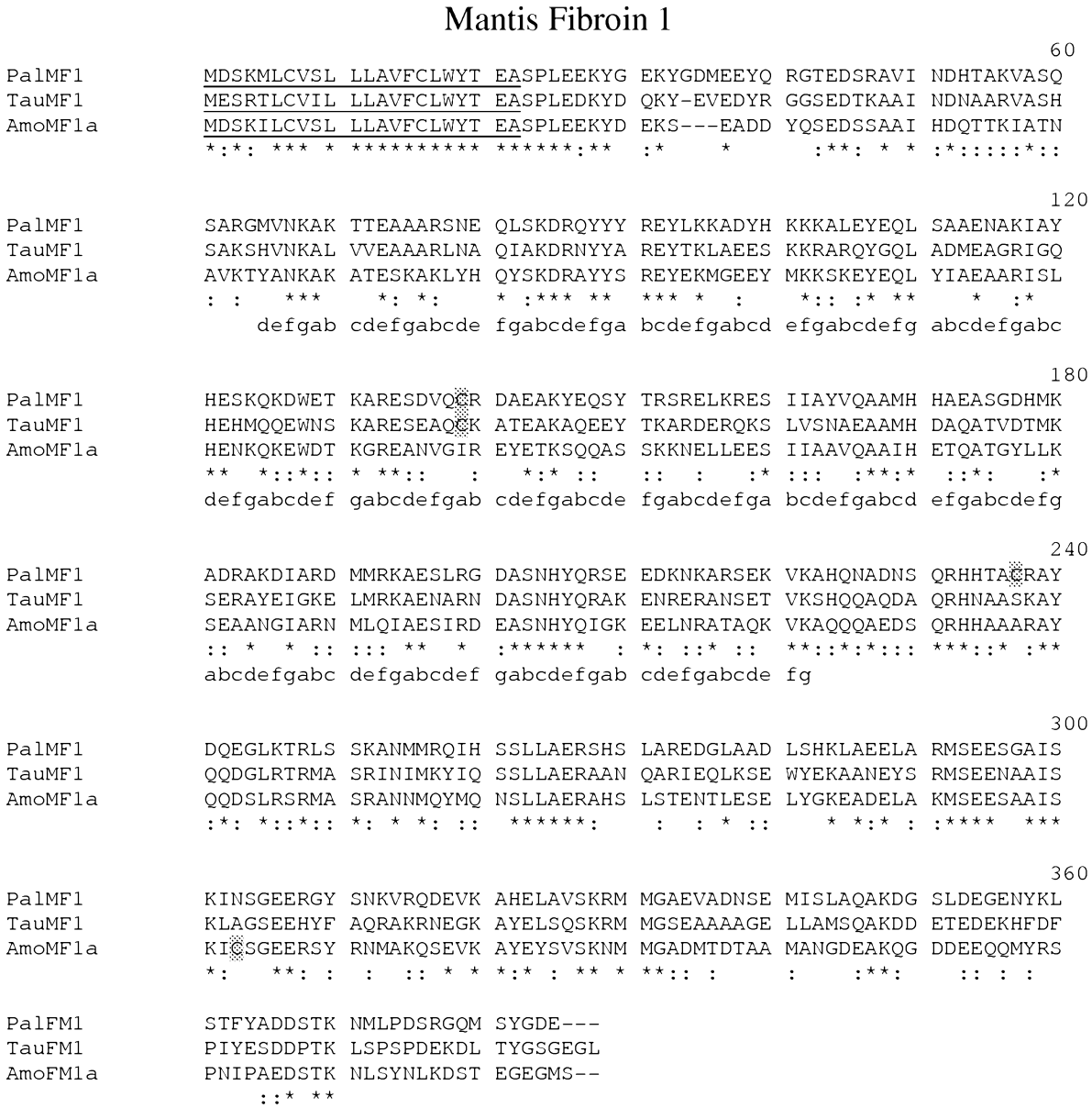
Mantis Fibroin 1
PalMF1
TauMF1
AmoMF1a
PalFM1
TauFM1
AmoFM1a

Figure 2

Mantis Fibroin 2

```
                                                                                                  60
PalMF2a         MKFHIAFVLL VVFGAAQAGK NHEVMTYGSG YKTMGDEGGS GVGNEGEDYQ DN---EGATA
TauMF2a         MKFHIAFVLL VIFGAAQAGK KHEVMTYGSG YKHMGGETYE DVGTGNR--- -----LGSTA
AmoMF2          MKFHIVFVLL VVFGAAQAGK KHEALTFGSG YKSTYGEGET FDDEDDQALR NERVPVGALS
                 ***** **** *:******** :** :*:*** **    *                    *: :

                                                                                                 120
PalMF2a         ATILDESTHH TEEARDIFGT RSEAHAYSAE MFADLVREKR QASIESHKKA EDYAVRANEE
TauMF2a         FDIMEAADEN TERASHTFGS KSAAYSSDAD LFIELLREKR ETRANHGKRA ESQAVLANES
AmoMF2          AAIINPYALH SEEGRIAYDT SSQYYANKAE GSADLSREKK QMHGEYHGKA ATYASRANEA
                  *::    : :*     : :  *  ::  *:    :* ***: :   :   :*    *  ***
                                                       cde fgabcdefga bcdefgabcd

                                                                                                 180
PalMF2a         SKKSQLLKRQ ARDKQAIAKQ YEEKAQKYDR ISKQQDIKEQ DDYRKSDAES EEYKRSIAVA
TauMF2a         YQKSQLHKRQ AKDKQAISKE YEERAQKHDR LSKEQDMKEH DDYRKSNAED TELRNSVERS
AmoMF2          YKKSQLHKRQ AKDKQAIAKE YEERAQKHES RSKALDVRDQ DDERKSVTEM EEYARALKIA
                 :**** *** *:*****:*: ***:***::   **  *:::: ** *** :*   *   ::  :
                 efgabcdefg abcdefgabc defgabcdef gabcdefgab cdefgabcde fgabcdefga

                                                                                                 240
PalMF2a         NAALALASAY EEASRMELDA TGEMEQQSKE LYTKSEEYNK VAEE▒ITRAK KEKELARIEE
TauMF2a         NYDHVMALGY HELSQLEMGE TNQ▒EQLSRE LQSRAEEYFN LAKELKEKAK KEKENARIKK
AmoMF2          NLALVFAGIY QETGRLQLEA TNVFEQFHKM LSTKGEEYKK QAEEYKEKAN KEKEEAAIQQ
                *    :*  *  *  ::::   *   **  :  * :: *** :  *:*   :*: **** * *::
                 bcdefgabcd efgabcdefg abcdefgabc defgabcdef gabcdefgab cdefgabcde

                                                                                                 300
PalMF2a         AKGKEAEAKS QEYENFATDN NKKYNAMKFY GWEFKMNAEN ERHNADY▒RI KSRYLAQLSN
TauMF2a         AKAKEEEARA EEYENAFTEN SKKVLAYKFY ELEFGMKALN EHHQAESARV RHHFLQILEQ
AmoMF2          AKSKELNAKA QEYENIFIES SKKLAANRYY ELEFKMKAEN ERHHAELARI RSRFLSRLAN
                ** ** :*:: :****   :   **  * ::*   ** *:* * *:*:*:  *: : ::*  * :
                 fgabcdefga bcdefga

                                                                                                 360
PalMF2a         YNREQAEALY HFAAAQRKDA ELFHRYAMEL YKQTRVLTAS AAQIMKQHRY TGQEIYSKQP
TauMF2a         HNSQHADMLW GYAQQEDKDG RSFTQYATEL SKQTKMLTAT AAHLMKQHRY TGMEMYSKQP
AmoMF2          YNREQAEAVL RFARSERKDG EIFRRNAIEL YKETRALAAT AARVMKQHRY TGQEIYTKQP
                :* ::*: :   :*  : **. . * : * **  *:*: *:*: **::****** ** *:*:***

PalMF2a         FPHSNYHGA
TauMF2a         FPHSNYHG-
AmoMF2          FPHSNYHGA
                ********
```

Figure 3

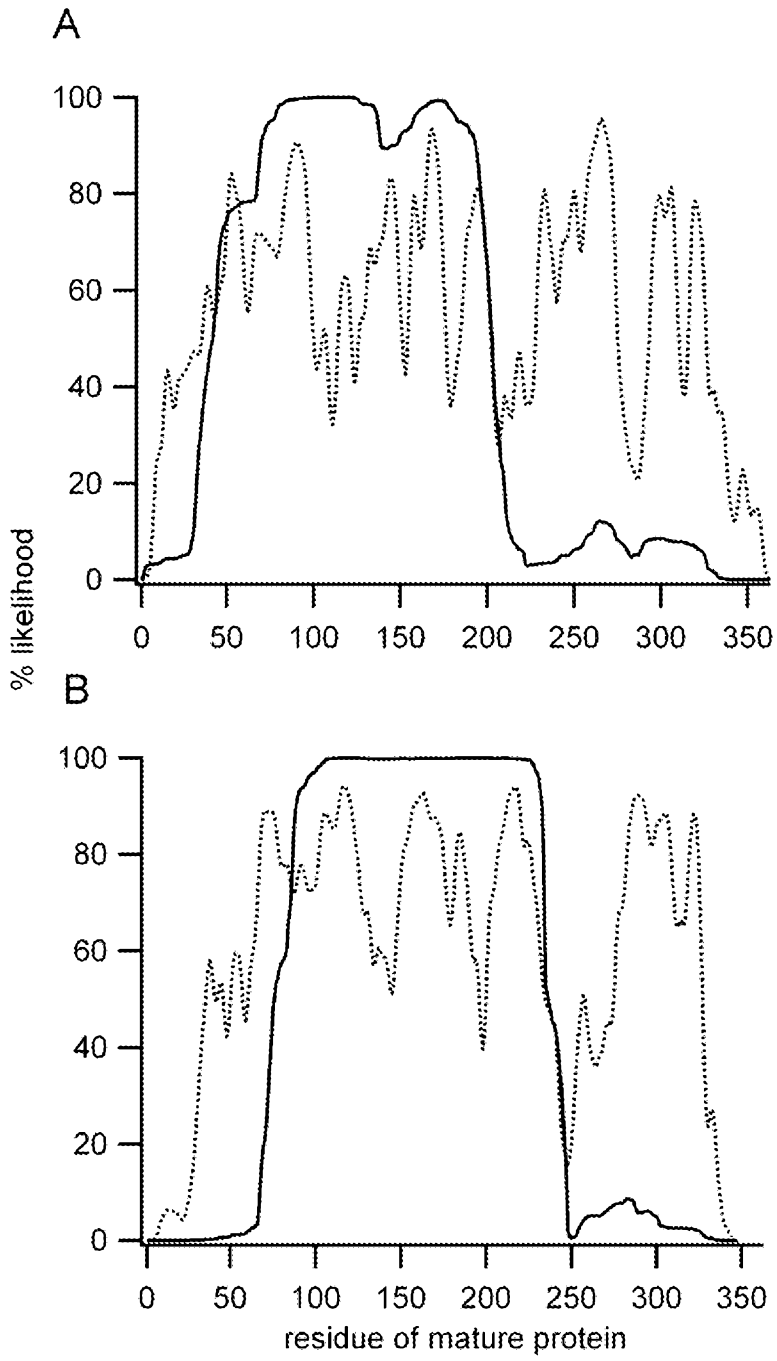
A
100
80
60
40
20
0
% likelihood
0
50
100
150
200
250
300
350
B
100
80
60
40
20
0
0
50
100
150
200
250
300
350
residue of mature protein

Figure 4

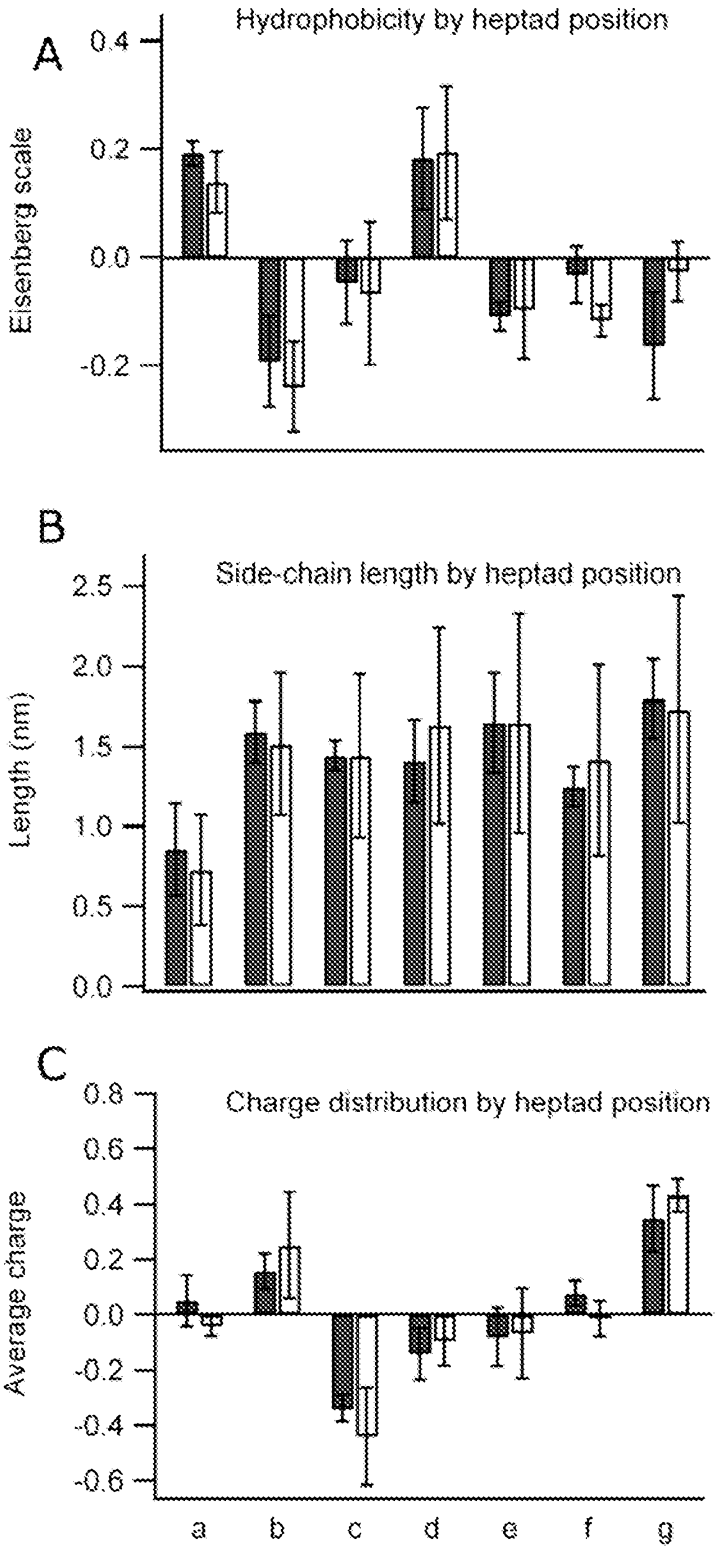
A
Hydrophobicity by heptad position
Eisenberg scale
0.4
0.2
0.0
-0.2
B
Side-chain length by heptad position
Length (nm)
2.5
2.0
1.5
1.0
0.5
0.0
C
Charge distribution by heptad position
Average charge
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
a
b
c
d
e
f
g

Figure 5

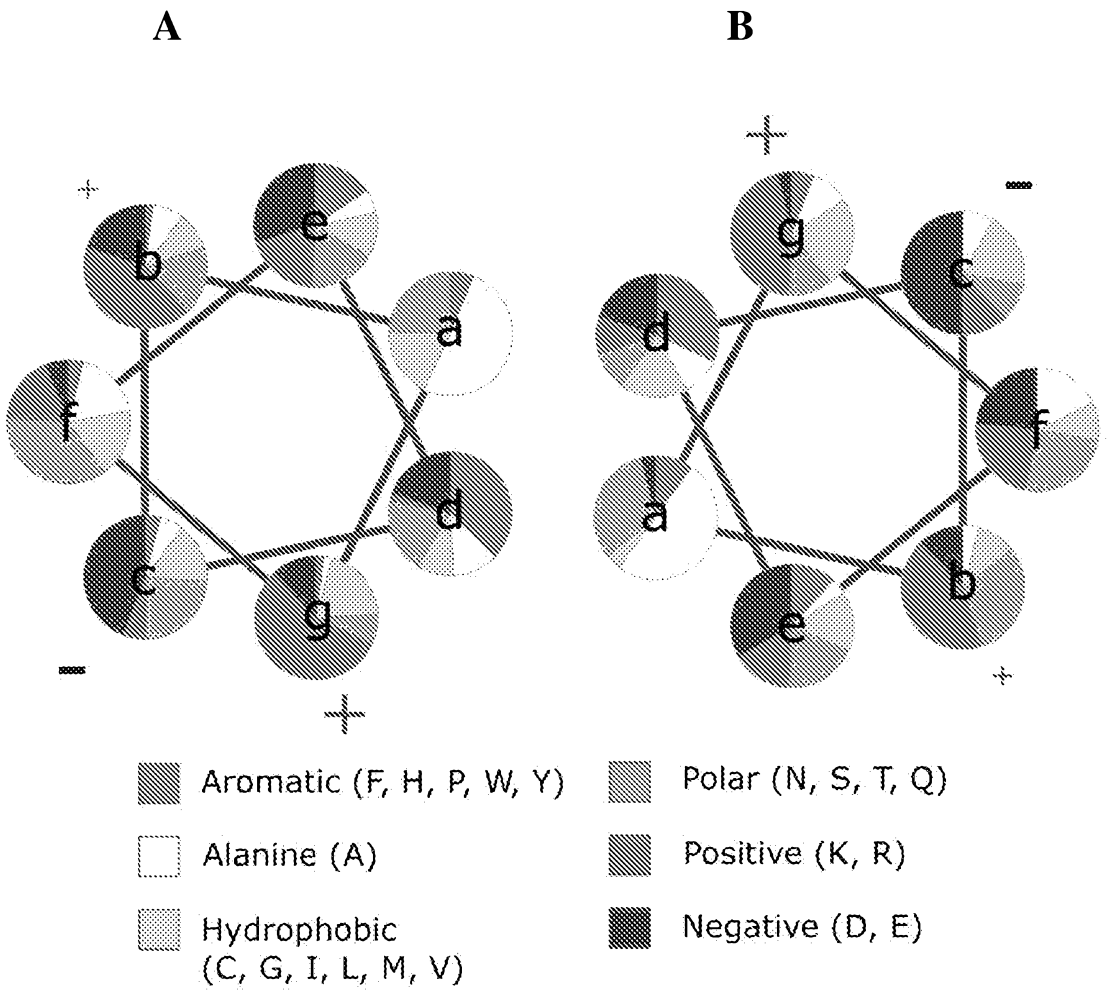
A
B
e
b
a
f
d
c
g
g
c
d
f
a
b
e
Aromatic (F, H, P, W, Y)
Alanine (A)
Hydrophobic
(C, G, I, L, M, V)
Polar (N, S, T, Q)
Positive (K, R)
Negative (D, E)

Figure 6

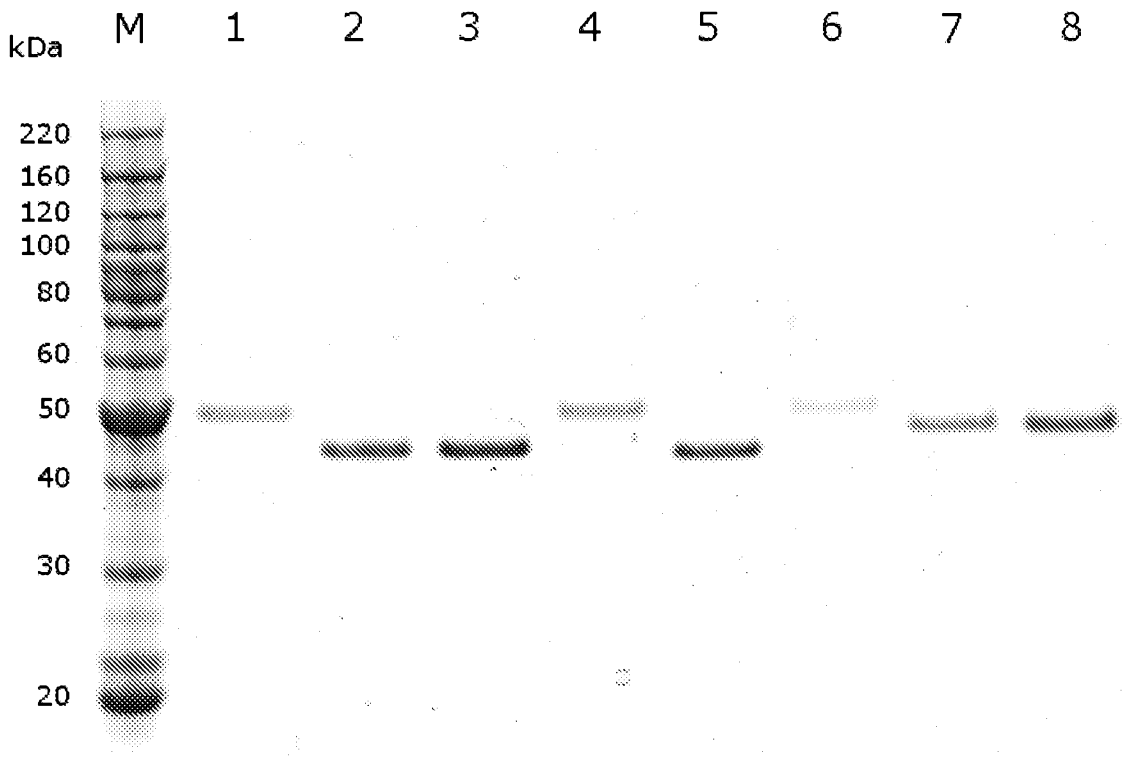

Figure 7

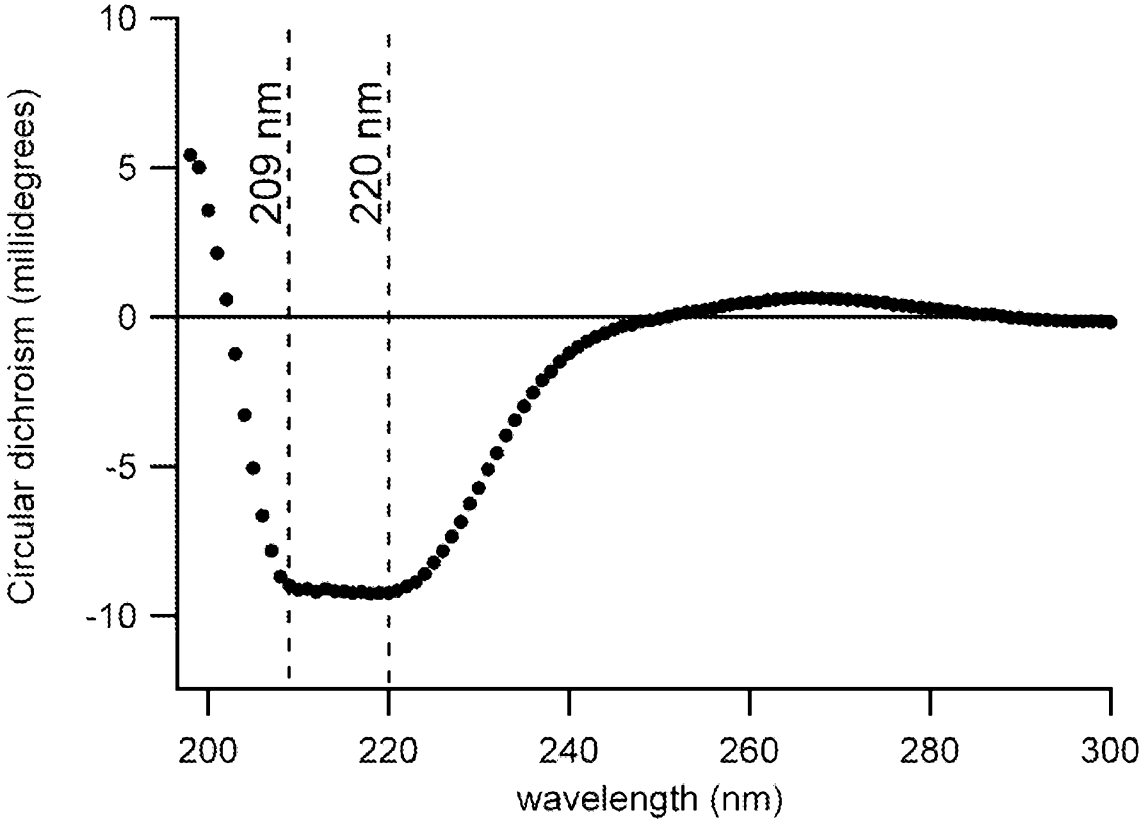
10
5
0
-5
-10
209 nm
220 nm
Circular dichroism (millidegrees)
200
220
240
260
280
300
wavelength (nm)

Figure 8

SEQUENCE LISTING

<110> Commonwealth Scientific and Indistrial Research Organisation

<120> Silk Polypeptides

<130> 513061

<150> 61/615,745

<151> 2012-03-26

<160> 54

<170> PatentIn version 3.5

<210> 1

<211> 1098

<212> DNA

<213> Tenodera australasiae

<400> 1

tctcccttgg aagacaaata cgaccaaaaa tatgaagtgg aagattaccg aggaggttcc 60

gaggacacca aggcagctat aaatgataac gcggcacgag ttgctagcca ttcggcgaag 120

tcccatgtca acaaagctct agttgtagaa gcagcagcga ggcttaatgc gcagatcgcc 180

aaggacagaa actactatgc tagggagtac accaagttgg ctgaagaatc caagaagagg 240

gctaggcaat atggacaact cgccgacatg gaagccggca gaattggcca gcatgaacac 300

atgcagcaag aatggaactc taaagctaga gaatccgagg cacagtgcaa ggctacagag 360

gcgaaagcac aagaggagta cacaaaggct agagatgaga ggcagaagag cttagtttct 420

aacgccgaag cagctatgca cgacgctcaa gctactgtgg atacgatgaa gtccgagaga 480

gcgtacgaga tcggcaagga attgatgcga aaggcagaga atgctagaaa tgatgccagc 540

aatcactacc agagagccaa ggaaaacaga gagagagcta attcagagac agtgaagtct 600

caccaacagg ctcaagacgc gcagagacat aatgctgctt ccaaagccta tcagcaggat 660

ggccttagga ctcgcatggc ttccagaatc aacattatga agtacatcca gagtagcttg 720

ttggccgaac gcgctgcaaa ccaagctagg atagagcaat taaaatccga atggtacgaa 780

aaagcagcca acgagtattc caggatgtct gaagagaacg ctgccattag caaacttgcc 840

ggaagcgaag aacattattt tgcgcaaagg gctaaacgaa atgaaggtaa ggcttatgaa 900

ttatcccaga gcaaaaggat gatgggttct gaggcggcag cagcaggcga actgcttgcc 960

atgtcacaag caaaggatga tgaaaccgag gacgaaaagc atttcgattt ccccatctat 1020

gaatccgatg accccaccaa attgtctcct tctcctgatg agaaagattt aacgtatggt 1080

tccggcgaag gtctgtaa 1098

<210> 2

<211> 1029

<212> DNA

<213> Tenodera australasiae

<400> 2

ggcaagaaac atgaagtaat gacctacggg tccggctaca aacatatggg cggcgaaact 60

tatgaggatg ttggaactgg aaataggttg ggctctaccg ccttcgacat tatggaagcg 120

gctgacgaaa atactgagcg agctagtcat acgtttggct ccaaatctgc agcctactcc 180

agcgatgctg acctgttcat cgagcttctc agagagaaga gagagacccg cgcgaaccac 240

ggcaaaagag cagagagcca ggctgtttta gccaacgagt cataccaaaa gtctcagttg 300

cacaagagac aagctaagga caagcaagcc atttccaagg aatacgagga aagggctcag 360

aaacacgacc ggttgagcaa agaacaggac atgaaggaac atgatgatta tagaaagagc 420

aatgcggaag atacagagct tcggaatagc gtagaaagaa gcaattatga tcatgttatg 480

gcactgggct atcatgaact gagtcagtta gaaatgggag agaccaatca atgcgagcaa 540

ttgtccagag aactgcagtc cagggctgaa gaatacttca atttagctaa agaactcaaa 600

gaaaaggcaa agaaggagaa ggagaatgct aggatcaaaa aagctaaggc caaagaggaa 660

gaagcgagag ctgaagagta cgaaaatgct ttcaccgaaa acagcaagaa ggtgctaaca 720

tacaaattct acgagctgga gttcggcatg aaggcactga acgagcatca tcaagccgaa 780

agtgccagag tcaggcacca cttcttacaa attctggagc aacacaatag ccagcacgct 840

gacatgctct gggggtatgc tcaacaagaa gataaggatg gtagatcctt cactcagtat 900

gcaactgagc ttagtaaaca gacaaagatg ttgaccgcaa ctgccgctca tctcatgaag 960

cagcatcgtt atactggcat ggaaatgtac tccaaacagc cattccccca ctccaactac 1020

cacggataa 1029

<210> 3

<211> 1086

<212> DNA

<213> Archimantis monstrosa

<400> 3

tctcccttgg aagaaaaata tgacgaaaaa tctgaagcgg atgattatca atcggaggac 60

agctcggcag ctattcatga ccaaacgaca aaaattgcta cgaacgccgt gaagacgtat 120

gccaacaaag ctaaagctac tgaatcaaag gcaaagcttt atcaccagta ctccaaggac 180

agagcgtact attctagaga atacgagaag atgggtgaag agtacatgaa gaagtctaag 240

gaatatgaac aactgtacat agctgaagcg gccaggattt cccttcatga aaacaaacag 300

aaggaatggg ataccaaagg tagagaagcc aatgtaggca ttagggaata cgaaacgaaa 360

agccaacagg cgtcctcaaa gaaaaatgag ctactggaag agagcataat tgccgcagtc 420

caagcagcca ttcacgagac tcaggcaaca gggtacctat tgaagtccga ggcagcgaac 480

ggaattgcca ggaatatgtt gcaaatagca gagagtatac gggatgaagc ctcgaaccac 540

tatcagatag gcaaggaaga gctaaatagg gccactgccc agaaggttaa ggcccaacaa 600

caggctgaag attctcagag acatcatgct gctgccagag cctatcagca ggattccctt 660

aggtctcgca tggcctccag agccaacaat atgcagtaca tgcagaacag cttgttggcc 720

gaacgcgctc atagcctaag tacggagaac acacttgaat ccgaactgta tggaaaagaa 780

gccgatgaat tagccaagat gtctgaagag agcgctgcca ttagcaaaat ctgtagcggt 840

gaagaacgta gttataggaa catggctaaa caaagtgaag ttaaggctta tgaatattct 900

gtgagcaaaa acatgatggg tgccgatatg acagacacag ccgcaatggc taacggggat 960

gaagcaaagc agggggatga cgaagagcaa cagatgtaca ggagccccaa cattcccgcc 1020

gaagactcca ccaaaaacct gtcttataat cttaaagact caactgaagg tgaaggaatg 1080

agttaa 1086

<210> 4

<211> 1056

<212> DNA

<213> Archimantis monstrosa

<400> 4

ggcaagaaac acgaagcatt aacctttggt tccggctaca agtctacgta cggagaaggc 60

gaaactttcg atgatgaaga tgatcaagct ctaaggaacg agagagttcc ggtgggcgct 120

ctctccgctg caattataaa cccgtatgcc ttgcattctg aggaaggtag aatagcgtac 180

gacacgtcat ctcaatatta tgccaacaaa gccgagggtt ccgctgacct ctccagagag 240

aagaaacaga tgcatgggga ataccacggt aaagcagcga cctatgcgtc cagagccaac 300

gaggcttaca agaaatctca attgcacaag agacaagcca aggacaagca agccatcgcc 360

aaggaatacg aggaaagagc acagaaacac gagagtagga gcaaagcact ggacgtaaga 420

gaccaggatg atgaaagaaa gagcgtgacc gaaatggaag agtatgcacg tgccctcaaa 480

atagccaatt tggcactggt ctttgcaggt atttatcagg aaacgggccg tttgcaatta 540

gaagcgacca atgtattcga gcagttccac aagatgctga gtaccaaggg cgaagagtac 600

aaaaagcaag ctgaggaata caaagaaaag gcaaataaag agaaggagga agctgctatc 660

caacaagcta agagcaaaga gttaaacgca aaggctcaag agtacgaaaa tatcttcatc 720

gaaagcagca agaagcttgc cgctaacaga tactacgagc tggagtttaa gatgaaagca 780

gagaacgagc gtcatcatgc cgagttagcc agaatcaggt ccaggttctt gtcacggctg 840

gccaactaca acagagaaca ggctgaggca gtcttgcggt ttgctcgttc cgaaaggaag 900

gacggtgaga ttttccgtag gaatgcaatt gagctttata aagagacaag ggcattggcc 960

gcaactgctg ctcgtgtcat gaagcagcac cgttatactg gccaggaaat ctacaccaag 1020

cagccattcc cccactccaa ctaccacgga gcataa 1056

<210> 5

<211> 1092

<212> DNA

<213> Pseudomantis albofimbriata

<400> 5

tcacccttgg aagaaaaata tggcgaaaaa tatggagata tggaggaata tcaaaggggc 60

accgaggaca gcagggcagt tattaatgac cacacggcaa aagttgctag ccaatcggca 120

cggggcatgg taaacaaagc taaaactact gaagcagcgg ccaggtctaa tgaacagctc 180

tccaaggaca gacaatacta ctatcgggag tacttaaaga aggctgatta ccacaagaag 240

aaggctttag aatatgaaca actatccgcg gctgaaaacg ccaagattgc ctatcacgaa 300

agcaaacaga aggattggga gaccaaagct agagaatccg acgtacagtg tagggatgca 360

gaagcgaaat acgaacagtc gtacaccagg agtagagagc tgaagagaga gagcataatt 420

gcctatgtcc aagcagccat gcaccatgct gaggcaagtg gggaccacat gaaggccgac 480

agagcgaaag acattgccag ggatatgatg cgaaaagcag agagtctgcg gggtgacgcc 540

tcgaaccact accagagatc cgaggaagac aaaaataagg cccgttccga gaaggttaag 600

gcccatcaaa atgctgacaa ttcccagaga caccatactg cttgcagagc ctatgaccag 660

gaaggcctta agactcgtct atcctccaaa gccaacatga tgaggcaaat ccatagcagc 720

ttgttggccg aacgctctca cagcttagct agggaagatg gacttgcagc cgatttgtct 780

cacaaactag ccgaagaact agccaggatg tctgaagaga gcggtgccat tagcaaaatt 840

aacagtggtg aagaacgtgg atattcgaac aaggttagac aagatgaagt taaggctcat 900

gaacttgctg tgagcaaaag gatgatgggt gctgaggtag cagacaactc cgaaatgatt 960

tccctggcac aagcaaagga tgggtcactc gatgagggag aaaattacaa gctctccacg 1020

ttttacgccg atgactccac aaaaaacatg cttcctgaca gtagaggcca aatgtcttac 1080

ggcgatgaat aa 1092

<210> 6

<211> 1047

<212> DNA

<213> Pseudomantis albofimbriata

<400> 6

ggcaagaatc acgaagtaat gacctatggt tccggctaca aaactatggg cgacgaaggg 60

ggaagcggtg ttggtaatga aggtgaagat taccaagata atgagggagc taccgccgcc 120

acaattttag atgagtctac ccatcatact gaggaagcta gagacatttt cggcacgaga 180

tctgaagctc acgcctacag tgccgagatg ttcgccgacc tcgtcagaga gaagagacaa 240

gcttccattg aaagccacaa gaaagcagag gactatgcgg tccgtgccaa cgaggagtac 300

aagaagtctc aattgctcaa gagacaagcc agggacaagc aagccatcgc caaacaatac 360

gaggaaaagg cacagaaata cgacaggata tccaaacaac aggacataaa agaacaggat 420

gattatagaa agagcgatgc cgaatcggaa gagtacaagc gtagcatcgc tgtagccaat 480

gcggcactgg ccttggcaag tgcctatgag gaagcgagcc gtatggagct cgatgcgacc 540

ggtgaaatgg agcagcagtc caaggaactg tacaccaagt ccgaggagta caataaggta 600

gctgaggaat gcattacaag ggcaaagaaa gagaaggaat tagctaggat cgaagaagct 660

aagggcaaag aggccgaagc aaagtctcaa gagtacgaaa atttcgccac cgacaacaac 720

aagaagtaca acgctatgaa attctacggg tgggagttta agatgaaggc agagaacgag 780

cgtcataatg ccgactattg cagaatcaaa tccaggtact tggcacagct gtccaactac 840

aacagagaac aggctgaggc actttaccat tttgctgctg cccaaagaaa ggacgctgag 900

ctcttccata ggtatgcaat ggagctttac aaacagacaa gggtattgac cgcatctgct 960

gctcaaatca tgaagcagca ccgttatact ggtcaggaga tctacagcaa gcagccattc 1020

ccccactcca actaccacgg agcataa 1047

<210> 7

<211> 1164

<212> DNA

<213> Tenodera australasiae

<400> 7

atggaatcca gaacactatg tgtaatcctg cttcttgcag tgttctgcct ttggtacacc 60

gaggcatctc ccttggaaga caaatacgac caaaaatatg aagtggaaga ttaccgagga 120

ggttccgagg acaccaaggc agctataaat gataacgcgg cacgagttgc tagccattcg 180

gcgaagtccc atgtcaacaa agctctagtt gtagaagcag cagcgaggct taatgcgcag 240

atcgccaagg acagaaacta ctatgctagg gagtacacca agttggctga agaatccaag 300

aagagggcta ggcaatatgg acaactcgcc gacatggaag ccggcagaat tggccagcat 360

gaacacatgc agcaagaatg gaactctaaa gctagagaat ccgaggcaca gtgcaaggct 420

acagaggcga aagcacaaga ggagtacaca aaggctagag atgagaggca gaagagctta 480

gtttctaacg ccgaagcagc tatgcacgac gctcaagcta ctgtggatac gatgaagtcc 540

gagagagcgt acgagatcgg caaggaattg atgcgaaagg cagagaatgc tagaaatgat 600

gccagcaatc actaccagag agccaaggaa aacagagaga gagctaattc agagacagtg 660

aagtctcacc aacaggctca agacgcgcag agacataatg ctgcttccaa agcctatcag 720

caggatggcc ttaggactcg catggcttcc agaatcaaca ttatgaagta catccagagt 780

agcttgttgg ccgaacgcgc tgcaaaccaa gctaggatag agcaattaaa atccgaatgg 840

tacgaaaaag cagccaacga gtattccagg atgtctgaag agaacgctgc cattagcaaa 900

cttgccggaa gcgaagaaca ttattttgcg caaagggcta aacgaaatga aggtaaggct 960

tatgaattat cccagagcaa aaggatgatg ggttctgagg cggcagcagc aggcgaactg 1020

cttgccatgt cacaagcaaa ggatgatgaa accgaggacg aaaagcattt cgatttcccc 1080

atctatgaat ccgatgaccc caccaaattg tctccttctc ctgatgagaa agatttaacg 1140

tatggttccg gcgaaggtct gtaa 1164

<210> 8

<211> 1083

<212> DNA

<213> Tenodera australasiae

<400> 8

atgaagttcc atattgcttt cgtgcttctg gtcatctttg gcgcggctca ggccggcaag 60

aaacatgaag taatgaccta cgggtccggc tacaaacata tgggcggcga aacttatgag 120

gatgttggaa ctggaaatag gttgggctct accgccttcg acattatgga agcggctgac 180

gaaaatactg agcgagctag tcatacgttt ggctccaaat ctgcagccta ctccagcgat 240

gctgacctgt tcatcgagct tctcagagag aagagagaga cccgcgcgaa ccacggcaaa 300

agagcagaga gccaggctgt tttagccaac gagtcatacc aaaagtctca gttgcacaag 360

agacaagcta aggacaagca agccatttcc aaggaatacg aggaaagggc tcagaaacac 420

gaccggttga gcaaagaaca ggacatgaag gaacatgatg attatagaaa gagcaatgcg 480

gaagatacag agcttcggaa tagcgtagaa agaagcaatt atgatcatgt tatggcactg 540

ggctatcatg aactgagtca gttagaaatg ggagagacca atcaatgcga gcaattgtcc 600

agagaactgc agtccagggc tgaagaatac ttcaatttag ctaaagaact caaagaaaag 660

gcaaagaagg agaaggagaa tgctaggatc aaaaaagcta aggccaaaga ggaagaagcg 720

agagctgaag agtacgaaaa tgctttcacc gaaaacagca agaaggtgct aacatacaaa 780

ttctacgagc tggagttcgg catgaaggca ctgaacgagc atcatcaagc cgaaagtgcc 840

agagtcaggc accacttctt acaaattctg gagcaacaca atagccagca cgctgacatg 900

ctctgggggt atgctcaaca agaagataag gatggtagat ccttcactca gtatgcaact 960

gagcttagta aacagacaaa gatgttgacc gcaactgccg ctcatctcat gaagcagcat 1020

cgttatactg gcatggaaat gtactccaaa cagccattcc cccactccaa ctaccacgga 1080

taa 1083

<210> 9

<211> 1152

<212> DNA

<213> Archimantis monstrosa

<400> 9

atggactcca aaatactatg tgtaagcctg ctccttgcag tgttctgcct ttggtacacc 60

gaggcatctc ccttggaaga aaaatatgac gaaaaatctg aagcggatga ttatcaatcg 120

gaggacagct cggcagctat tcatgaccaa acgacaaaaa ttgctacgaa cgccgtgaag 180

acgtatgcca acaaagctaa agctactgaa tcaaaggcaa agctttatca ccagtactcc 240

aaggacagag cgtactattc tagagaatac gagaagatgg gtgaagagta catgaagaag 300

tctaaggaat atgaacaact gtacatagct gaagcggcca ggatttccct tcatgaaaac 360

aaacagaagg aatgggatac caaaggtaga gaagccaatg taggcattag ggaatacgaa 420

acgaaaagcc aacaggcgtc ctcaaagaaa aatgagctac tggaagagag cataattgcc 480

gcagtccaag cagccattca cgagactcag gcaacagggt acctattgaa gtccgaggca 540

gcgaacggaa ttgccaggaa tatgttgcaa atagcagaga gtatacggga tgaagcctcg 600

aaccactatc agataggcaa ggaagagcta aatagggcca ctgcccagaa ggttaaggcc 660

caacaacagg ctgaagattc tcagagacat catgctgctg ccagagccta tcagcaggat 720

tcccttaggt ctcgcatggc ctccagagcc aacaatatgc agtacatgca gaacagcttg 780

ttggccgaac gcgctcatag cctaagtacg gagaacacac ttgaatccga actgtatgga 840

aaagaagccg atgaattagc caagatgtct gaagagagcg ctgccattag caaaatctgt 900

agcggtgaag aacgtagtta taggaacatg gctaaacaaa gtgaagttaa ggcttatgaa 960

tattctgtga gcaaaaacat gatgggtgcc gatatgacag acacagccgc aatggctaac 1020

ggggatgaag caaagcaggg ggatgacgaa gagcaacaga tgtacaggag ccccaacatt 1080

cccgccgaag actccaccaa aaacctgtct tataatctta aagactcaac tgaaggtgaa 1140

ggaatgagtt aa 1152

<210> 10

<211> 1110

<212> DNA

<213> Archimantis monstrosa

<400> 10

atgaagttcc acattgtctt cgttcttctg gtcgtctttg gcgcggctca ggctggcaag 60

aaacacgaag cattaacctt tggttccggc tacaagtcta cgtacggaga aggcgaaact 120

ttcgatgatg aagatgatca agctctaagg aacgagagag ttccggtggg cgctctctcc 180

gctgcaatta taaacccgta tgccttgcat tctgaggaag gtagaatagc gtacgacacg 240

tcatctcaat attatgccaa caaagccgag ggttccgctg acctctccag agagaagaaa 300

cagatgcatg gggaatacca cggtaaagca gcgacctatg cgtccagagc caacgaggct 360

tacaagaaat ctcaattgca caagagacaa gccaaggaca agcaagccat cgccaaggaa 420

tacgaggaaa gagcacagaa acacgagagt aggagcaaag cactggacgt aagagaccag 480

gatgatgaaa gaaagagcgt gaccgaaatg gaagagtatg cacgtgccct caaaatagcc 540

aatttggcac tggtctttgc aggtatttat caggaaacgg gccgtttgca attagaagcg 600

accaatgtat tcgagcagtt ccacaagatg ctgagtacca agggcgaaga gtacaaaaag 660

caagctgagg aatacaaaga aaaggcaaat aaagagaagg aggaagctgc tatccaacaa 720

gctaagagca aagagttaaa cgcaaaggct caagagtacg aaaatatctt catcgaaagc 780

agcaagaagc ttgccgctaa cagatactac gagctggagt ttaagatgaa agcagagaac 840

gagcgtcatc atgccgagtt agccagaatc aggtccaggt tcttgtcacg gctggccaac 900

tacaacagag aacaggctga ggcagtcttg cggtttgctc gttccgaaag gaaggacggt 960

gagattttcc gtaggaatgc aattgagctt tataaagaga caagggcatt ggccgcaact 1020

gctgctcgtg tcatgaagca gcaccgttat actggccagg aaatctacac caagcagcca 1080

ttcccccact ccaactacca cggagcataa 1110

<210> 11

<211> 1158

<212> DNA

<213> Pseudomantis albofimbriata

<400> 11

atggactcca aaatgctgtg tgtaagcctg ctcctagcag tgttctgcct ttggtacacc 60

gaggcatcac ccttggaaga aaaatatggc gaaaaatatg gagatatgga ggaatatcaa 120

aggggcaccg aggacagcag ggcagttatt aatgaccaca cggcaaaagt tgctagccaa 180

tcggcacggg gcatggtaaa caaagctaaa actactgaag cagcggccag gtctaatgaa 240

cagctctcca aggacagaca atactactat cgggagtact taaagaaggc tgattaccac 300

aagaagaagg ctttagaata tgaacaacta tccgcggctg aaaacgccaa gattgcctat 360

cacgaaagca aacagaagga ttgggagacc aaagctagag aatccgacgt acagtgtagg 420

gatgcagaag cgaaatacga acagtcgtac accaggagta gagagctgaa gagagagagc 480

ataattgcct atgtccaagc agccatgcac catgctgagg caagtgggga ccacatgaag 540

gccgacagag cgaaagacat tgccagggat atgatgcgaa aagcagagag tctgcggggt 600

gacgcctcga accactacca gagatccgag gaagacaaaa ataaggcccg ttccgagaag 660

gttaaggccc atcaaaatgc tgacaattcc cagagacacc atactgcttg cagagcctat 720

gaccaggaag gccttaagac tcgtctatcc tccaaagcca acatgatgag gcaaatccat 780

agcagcttgt tggccgaacg ctctcacagc ttagctaggg aagatggact tgcagccgat 840

ttgtctcaca aactagccga agaactagcc aggatgtctg aagagagcgg tgccattagc 900

aaaattaaca gtggtgaaga acgtggatat tcgaacaagg ttagacaaga tgaagttaag 960

gctcatgaac ttgctgtgag caaaaggatg atgggtgctg aggtagcaga caactccgaa 1020

atgatttccc tggcacaagc aaaggatggg tcactcgatg agggagaaaa ttacaagctc 1080

tccacgtttt acgccgatga ctccacaaaa aacatgcttc ctgacagtag aggccaaatg 1140

tcttacggcg atgaataa 1158

<210> 12

<211> 1101

<212> DNA

<213> Pseudomantis albofimbriata

<400> 12

atgaagttcc acattgcctt cgttcttctg gtcgtctttg gcgcggctca ggccggcaag 60

aatcacgaag taatgaccta tggttccggc tacaaaacta tgggcgacga agggggaagc 120

ggtgttggta atgaaggtga agattaccaa gataatgagg gagctaccgc cgccacaatt 180

ttagatgagt ctacccatca tactgaggaa gctagagaca ttttcggcac gagatctgaa 240

gctcacgcct acagtgccga gatgttcgcc gacctcgtca gagagaagag acaagcttcc 300

attgaaagcc acaagaaagc agaggactat gcggtccgtg ccaacgagga gtacaagaag 360

tctcaattgc tcaagagaca agccagggac aagcaagcca tcgccaaaca atacgaggaa 420

aaggcacaga aatacgacag gatatccaaa caacaggaca taaaagaaca ggatgattat 480

agaaagagcg atgccgaatc ggaagagtac aagcgtagca tcgctgtagc caatgcggca 540

ctggccttgg caagtgccta tgaggaagcg agccgtatgg agctcgatgc gaccggtgaa 600

atggagcagc agtccaagga actgtacacc aagtccgagg agtacaataa ggtagctgag 660

gaatgcatta caagggcaaa gaaagagaag gaattagcta ggatcgaaga agctaagggc 720

aaagaggccg aagcaaagtc tcaagagtac gaaaatttcg ccaccgacaa caacaagaag 780

tacaacgcta tgaaattcta cgggtgggag tttaagatga aggcagagaa cgagcgtcat 840

aatgccgact attgcagaat caaatccagg tacttggcac agctgtccaa ctacaacaga 900

gaacaggctg aggcacttta ccattttgct gctgcccaaa gaaaggacgc tgagctcttc 960

cataggtatg caatggagct ttacaaacag acaagggtat tgaccgcatc tgctgctcaa 1020

atcatgaagc agcaccgtta tactggtcag gagatctaca gcaagcagcc attcccccac 1080

tccaactacc acggagcata a 1101

<210> 13

<211> 365

<212> PRT

<213> Tenodera australasiae

<400> 13

Ser Pro Leu Glu Asp Lys Tyr Asp Gln Lys Tyr Glu Val Glu Asp Tyr

1 5 10 15

Arg Gly Gly Ser Glu Asp Thr Lys Ala Ala Ile Asn Asp Asn Ala Ala

20 25 30

Arg Val Ala Ser His Ser Ala Lys Ser His Val Asn Lys Ala Leu Val

35 40 45

Val Glu Ala Ala Ala Arg Leu Asn Ala Gln Ile Ala Lys Asp Arg Asn

50 55 60

Tyr Tyr Ala Arg Glu Tyr Thr Lys Leu Ala Glu Glu Ser Lys Lys Arg

65 70 75 80

Ala Arg Gln Tyr Gly Gln Leu Ala Asp Met Glu Ala Gly Arg Ile Gly

85 90 95

Gln His Glu His Met Gln Gln Glu Trp Asn Ser Lys Ala Arg Glu Ser

100 105 110

Glu Ala Gln Cys Lys Ala Thr Glu Ala Lys Ala Gln Glu Glu Tyr Thr

115 120 125

Lys Ala Arg Asp Glu Arg Gln Lys Ser Leu Val Ser Asn Ala Glu Ala

130 135 140

Ala Met His Asp Ala Gln Ala Thr Val Asp Thr Met Lys Ser Glu Arg

145 150 155 160

Ala Tyr Glu Ile Gly Lys Glu Leu Met Arg Lys Ala Glu Asn Ala Arg

165 170 175

Asn Asp Ala Ser Asn His Tyr Gln Arg Ala Lys Glu Asn Arg Glu Arg

180 185 190

Ala Asn Ser Glu Thr Val Lys Ser His Gln Gln Ala Gln Asp Ala Gln

195 200 205

Arg His Asn Ala Ala Ser Lys Ala Tyr Gln Gln Asp Gly Leu Arg Thr

210 215 220

Arg Met Ala Ser Arg Ile Asn Ile Met Lys Tyr Ile Gln Ser Ser Leu

225 230 235 240

Leu Ala Glu Arg Ala Ala Asn Gln Ala Arg Ile Glu Gln Leu Lys Ser

245 250 255

Glu Trp Tyr Glu Lys Ala Ala Asn Glu Tyr Ser Arg Met Ser Glu Glu

260 265 270

Asn Ala Ala Ile Ser Lys Leu Ala Gly Ser Glu Glu His Tyr Phe Ala

275 280 285

Gln Arg Ala Lys Arg Asn Glu Gly Lys Ala Tyr Glu Leu Ser Gln Ser

290 295 300

Lys Arg Met Met Gly Ser Glu Ala Ala Ala Ala Gly Glu Leu Leu Ala

305 310 315 320

Met Ser Gln Ala Lys Asp Asp Glu Thr Glu Asp Glu Lys His Phe Asp

325 330 335

Phe Pro Ile Tyr Glu Ser Asp Asp Pro Thr Lys Leu Ser Pro Ser Pro

340 345 350

Asp Glu Lys Asp Leu Thr Tyr Gly Ser Gly Glu Gly Leu

355 360 365

<210> 14

<211> 342

<212> PRT

<213> Tenodera australasiae

<400> 14

Gly Lys Lys His Glu Val Met Thr Tyr Gly Ser Gly Tyr Lys His Met

1 5 10 15

Gly Gly Glu Thr Tyr Glu Asp Val Gly Thr Gly Asn Arg Leu Gly Ser

20 25 30

Thr Ala Phe Asp Ile Met Glu Ala Ala Asp Glu Asn Thr Glu Arg Ala

35 40 45

Ser His Thr Phe Gly Ser Lys Ser Ala Ala Tyr Ser Ser Asp Ala Asp

50 55 60

Leu Phe Ile Glu Leu Leu Arg Glu Lys Arg Glu Thr Arg Ala Asn His

65 70 75 80

Gly Lys Arg Ala Glu Ser Gln Ala Val Leu Ala Asn Glu Ser Tyr Gln

85 90 95

Lys Ser Gln Leu His Lys Arg Gln Ala Lys Asp Lys Gln Ala Ile Ser

100 105 110

Lys Glu Tyr Glu Glu Arg Ala Gln Lys His Asp Arg Leu Ser Lys Glu

115 120 125

Gln Asp Met Lys Glu His Asp Asp Tyr Arg Lys Ser Asn Ala Glu Asp

130 135 140

Thr Glu Leu Arg Asn Ser Val Glu Arg Ser Asn Tyr Asp His Val Met

145 150 155 160

Ala Leu Gly Tyr His Glu Leu Ser Gln Leu Glu Met Gly Glu Thr Asn

165 170 175

Gln Cys Glu Gln Leu Ser Arg Glu Leu Gln Ser Arg Ala Glu Glu Tyr

180 185 190

Phe Asn Leu Ala Lys Glu Leu Lys Glu Lys Ala Lys Lys Glu Lys Glu

195 200 205

Asn Ala Arg Ile Lys Lys Ala Lys Ala Lys Glu Glu Glu Ala Arg Ala

210 215 220

Glu Glu Tyr Glu Asn Ala Phe Thr Glu Asn Ser Lys Lys Val Leu Thr

225 230 235 240

Tyr Lys Phe Tyr Glu Leu Glu Phe Gly Met Lys Ala Leu Asn Glu His

245 250 255

His Gln Ala Glu Ser Ala Arg Val Arg His His Phe Leu Gln Ile Leu

260 265 270

Glu Gln His Asn Ser Gln His Ala Asp Met Leu Trp Gly Tyr Ala Gln

275 280 285

Gln Glu Asp Lys Asp Gly Arg Ser Phe Thr Gln Tyr Ala Thr Glu Leu

290 295 300

Ser Lys Gln Thr Lys Met Leu Thr Ala Thr Ala Ala His Leu Met Lys

305 310 315 320

Gln His Arg Tyr Thr Gly Met Glu Met Tyr Ser Lys Gln Pro Phe Pro

325 330 335

His Ser Asn Tyr His Gly

340

<210> 15

<211> 361

<212> PRT

<213> Archimantis monstrosa

<400> 15

Ser Pro Leu Glu Glu Lys Tyr Asp Glu Lys Ser Glu Ala Asp Asp Tyr

1 5 10 15

Gln Ser Glu Asp Ser Ser Ala Ala Ile His Asp Gln Thr Thr Lys Ile

20 25 30

Ala Thr Asn Ala Val Lys Thr Tyr Ala Asn Lys Ala Lys Ala Thr Glu

35 40 45

Ser Lys Ala Lys Leu Tyr His Gln Tyr Ser Lys Asp Arg Ala Tyr Tyr

50 55 60

Ser Arg Glu Tyr Glu Lys Met Gly Glu Glu Tyr Met Lys Lys Ser Lys

65 70 75 80

Glu Tyr Glu Gln Leu Tyr Ile Ala Glu Ala Ala Arg Ile Ser Leu His

85 90 95

Glu Asn Lys Gln Lys Glu Trp Asp Thr Lys Gly Arg Glu Ala Asn Val

100 105 110

Gly Ile Arg Glu Tyr Glu Thr Lys Ser Gln Gln Ala Ser Ser Lys Lys

115 120 125

Asn Glu Leu Leu Glu Glu Ser Ile Ile Ala Ala Val Gln Ala Ala Ile

130 135 140

His Glu Thr Gln Ala Thr Gly Tyr Leu Leu Lys Ser Glu Ala Ala Asn

145 150 155 160

Gly Ile Ala Arg Asn Met Leu Gln Ile Ala Glu Ser Ile Arg Asp Glu

165 170 175

Ala Ser Asn His Tyr Gln Ile Gly Lys Glu Glu Leu Asn Arg Ala Thr

180 185 190

Ala Gln Lys Val Lys Ala Gln Gln Gln Ala Glu Asp Ser Gln Arg His

195 200 205

His Ala Ala Ala Arg Ala Tyr Gln Gln Asp Ser Leu Arg Ser Arg Met

210 215 220

Ala Ser Arg Ala Asn Asn Met Gln Tyr Met Gln Asn Ser Leu Leu Ala

225 230 235 240

Glu Arg Ala His Ser Leu Ser Thr Glu Asn Thr Leu Glu Ser Glu Leu

245 250 255

Tyr Gly Lys Glu Ala Asp Glu Leu Ala Lys Met Ser Glu Glu Ser Ala

260 265 270

Ala Ile Ser Lys Ile Cys Ser Gly Glu Glu Arg Ser Tyr Arg Asn Met

275 280 285

Ala Lys Gln Ser Glu Val Lys Ala Tyr Glu Tyr Ser Val Ser Lys Asn

290 295 300

Met Met Gly Ala Asp Met Thr Asp Thr Ala Ala Met Ala Asn Gly Asp

305 310 315 320

Glu Ala Lys Gln Gly Asp Asp Glu Glu Gln Gln Met Tyr Arg Ser Pro

325 330 335

Asn Ile Pro Ala Glu Asp Ser Thr Lys Asn Leu Ser Tyr Asn Leu Lys

340 345 350

Asp Ser Thr Glu Gly Glu Gly Met Ser

355 360

<210> 16

<211> 351

<212> PRT

<213> Archimantis monstrosa

<400> 16

Gly Lys Lys His Glu Ala Leu Thr Phe Gly Ser Gly Tyr Lys Ser Thr

1 5 10 15

Tyr Gly Glu Gly Glu Thr Phe Asp Asp Glu Asp Asp Gln Ala Leu Arg

20 25 30

Asn Glu Arg Val Pro Val Gly Ala Leu Ser Ala Ala Ile Ile Asn Pro

35 40 45

Tyr Ala Leu His Ser Glu Glu Gly Arg Ile Ala Tyr Asp Thr Ser Ser

50 55 60

Gln Tyr Tyr Ala Asn Lys Ala Glu Gly Ser Ala Asp Leu Ser Arg Glu

65 70 75 80

Lys Lys Gln Met His Gly Glu Tyr His Gly Lys Ala Ala Thr Tyr Ala

85 90 95

Ser Arg Ala Asn Glu Ala Tyr Lys Lys Ser Gln Leu His Lys Arg Gln

100 105 110

Ala Lys Asp Lys Gln Ala Ile Ala Lys Glu Tyr Glu Glu Arg Ala Gln

115 120 125

Lys His Glu Ser Arg Ser Lys Ala Leu Asp Val Arg Asp Gln Asp Asp

130 135 140

Glu Arg Lys Ser Val Thr Glu Met Glu Glu Tyr Ala Arg Ala Leu Lys

145 150 155 160

Ile Ala Asn Leu Ala Leu Val Phe Ala Gly Ile Tyr Gln Glu Thr Gly

165 170 175

Arg Leu Gln Leu Glu Ala Thr Asn Val Phe Glu Gln Phe His Lys Met

180 185 190

Leu Ser Thr Lys Gly Glu Glu Tyr Lys Lys Gln Ala Glu Glu Tyr Lys

195 200 205

Glu Lys Ala Asn Lys Glu Lys Glu Glu Ala Ala Ile Gln Gln Ala Lys

210 215 220

Ser Lys Glu Leu Asn Ala Lys Ala Gln Glu Tyr Glu Asn Ile Phe Ile

225 230 235 240

Glu Ser Ser Lys Lys Leu Ala Ala Asn Arg Tyr Tyr Glu Leu Glu Phe

245 250 255

Lys Met Lys Ala Glu Asn Glu Arg His His Ala Glu Leu Ala Arg Ile

260 265 270

Arg Ser Arg Phe Leu Ser Arg Leu Ala Asn Tyr Asn Arg Glu Gln Ala

275 280 285

Glu Ala Val Leu Arg Phe Ala Arg Ser Glu Arg Lys Asp Gly Glu Ile

290 295 300

Phe Arg Arg Asn Ala Ile Glu Leu Tyr Lys Glu Thr Arg Ala Leu Ala

305 310 315 320

Ala Thr Ala Ala Arg Val Met Lys Gln His Arg Tyr Thr Gly Gln Glu

325 330 335

Ile Tyr Thr Lys Gln Pro Phe Pro His Ser Asn Tyr His Gly Ala

340 345 350

<210> 17

<211> 363

<212> PRT

<213> Pseudomantis albofimbriata

<400> 17

Ser Pro Leu Glu Glu Lys Tyr Gly Glu Lys Tyr Gly Asp Met Glu Glu
1 5 10 15

Tyr Gln Arg Gly Thr Glu Asp Ser Arg Ala Val Ile Asn Asp His Thr
20 25 30

Ala Lys Val Ala Ser Gln Ser Ala Arg Gly Met Val Asn Lys Ala Lys
35 40 45

Thr Thr Glu Ala Ala Ala Arg Ser Asn Glu Gln Leu Ser Lys Asp Arg
50 55 60

Gln Tyr Tyr Tyr Arg Glu Tyr Leu Lys Lys Ala Asp Tyr His Lys Lys
65 70 75 80

Lys Ala Leu Glu Tyr Glu Gln Leu Ser Ala Ala Glu Asn Ala Lys Ile

85 90 95

Ala Tyr His Glu Ser Lys Gln Lys Asp Trp Glu Thr Lys Ala Arg Glu

100 105 110

Ser Asp Val Gln Cys Arg Asp Ala Glu Ala Lys Tyr Glu Gln Ser Tyr

115 120 125

Thr Arg Ser Arg Glu Leu Lys Arg Glu Ser Ile Ile Ala Tyr Val Gln

130 135 140

Ala Ala Met His His Ala Glu Ala Ser Gly Asp His Met Lys Ala Asp

145 150 155 160

Arg Ala Lys Asp Ile Ala Arg Asp Met Met Arg Lys Ala Glu Ser Leu

165 170 175

Arg Gly Asp Ala Ser Asn His Tyr Gln Arg Ser Glu Glu Asp Lys Asn

180 185 190

Lys Ala Arg Ser Glu Lys Val Lys Ala His Gln Asn Ala Asp Asn Ser

195 200 205

Gln Arg His His Thr Ala Cys Arg Ala Tyr Asp Gln Glu Gly Leu Lys

210 215 220

Thr Arg Leu Ser Ser Lys Ala Asn Met Met Arg Gln Ile His Ser Ser

225 230 235 240

Leu Leu Ala Glu Arg Ser His Ser Leu Ala Arg Glu Asp Gly Leu Ala

245 250 255

Ala Asp Leu Ser His Lys Leu Ala Glu Glu Leu Ala Arg Met Ser Glu

260 265 270

Glu Ser Gly Ala Ile Ser Lys Ile Asn Ser Gly Glu Glu Arg Gly Tyr

275 280 285

Ser Asn Lys Val Arg Gln Asp Glu Val Lys Ala His Glu Leu Ala Val

290 295 300

Ser Lys Arg Met Met Gly Ala Glu Val Ala Asp Asn Ser Glu Met Ile

305 310 315 320

Ser Leu Ala Gln Ala Lys Asp Gly Ser Leu Asp Glu Gly Glu Asn Tyr

325 330 335

Lys Leu Ser Thr Phe Tyr Ala Asp Asp Ser Thr Lys Asn Met Leu Pro

340 345 350

Asp Ser Arg Gly Gln Met Ser Tyr Gly Asp Glu

355 360

<210> 18

<211> 348

<212> PRT

<213> Pseudomantis albofimbriata

<400> 18

Gly Lys Asn His Glu Val Met Thr Tyr Gly Ser Gly Tyr Lys Thr Met

1 5 10 15

Gly Asp Glu Gly Gly Ser Gly Val Gly Asn Glu Gly Glu Asp Tyr Gln

20 25 30

Asp Asn Glu Gly Ala Thr Ala Ala Thr Ile Leu Asp Glu Ser Thr His

35 40 45

His Thr Glu Glu Ala Arg Asp Ile Phe Gly Thr Arg Ser Glu Ala His

50 55 60

Ala Tyr Ser Ala Glu Met Phe Ala Asp Leu Val Arg Glu Lys Arg Gln

65 70 75 80

Ala Ser Ile Glu Ser His Lys Lys Ala Glu Asp Tyr Ala Val Arg Ala

85 90 95

Asn Glu Glu Tyr Lys Lys Ser Gln Leu Leu Lys Arg Gln Ala Arg Asp

100 105 110

Lys Gln Ala Ile Ala Lys Gln Tyr Glu Glu Lys Ala Gln Lys Tyr Asp

115 120 125

Arg Ile Ser Lys Gln Gln Asp Ile Lys Glu Gln Asp Asp Tyr Arg Lys

130 135 140

Ser Asp Ala Glu Ser Glu Glu Tyr Lys Arg Ser Ile Ala Val Ala Asn

145 150 155 160

Ala Ala Leu Ala Leu Ala Ser Ala Tyr Glu Glu Ala Ser Arg Met Glu

165 170 175

Leu Asp Ala Thr Gly Glu Met Glu Gln Gln Ser Lys Glu Leu Tyr Thr

180 185 190

Lys Ser Glu Glu Tyr Asn Lys Val Ala Glu Glu Cys Ile Thr Arg Ala

195 200 205

Lys Lys Glu Lys Glu Leu Ala Arg Ile Glu Glu Ala Lys Gly Lys Glu

210 215 220

Ala Glu Ala Lys Ser Gln Glu Tyr Glu Asn Phe Ala Thr Asp Asn Asn

225 230 235 240

Lys Lys Tyr Asn Ala Met Lys Phe Tyr Gly Trp Glu Phe Lys Met Lys

245 250 255

Ala Glu Asn Glu Arg His Asn Ala Asp Tyr Cys Arg Ile Lys Ser Arg

260 265 270

Tyr Leu Ala Gln Leu Ser Asn Tyr Asn Arg Glu Gln Ala Glu Ala Leu

275 280 285

Tyr His Phe Ala Ala Ala Gln Arg Lys Asp Ala Glu Leu Phe His Arg

290 295 300

Tyr Ala Met Glu Leu Tyr Lys Gln Thr Arg Val Leu Thr Ala Ser Ala

305 310 315 320

Ala Gln Ile Met Lys Gln His Arg Tyr Thr Gly Gln Glu Ile Tyr Ser

325 330 335

Lys Gln Pro Phe Pro His Ser Asn Tyr His Gly Ala

340 345

<210> 19

<211> 387

<212> PRT

<213> Tenodera australasiae

<400> 19

Met Glu Ser Arg Thr Leu Cys Val Ile Leu Leu Leu Ala Val Phe Cys

1 5 10 15

Leu Trp Tyr Thr Glu Ala Ser Pro Leu Glu Asp Lys Tyr Asp Gln Lys

20 25 30

Tyr Glu Val Glu Asp Tyr Arg Gly Gly Ser Glu Asp Thr Lys Ala Ala

35 40 45

Ile Asn Asp Asn Ala Ala Arg Val Ala Ser His Ser Ala Lys Ser His

50 55 60

Val Asn Lys Ala Leu Val Val Glu Ala Ala Ala Arg Leu Asn Ala Gln

65 70 75 80

Ile Ala Lys Asp Arg Asn Tyr Tyr Ala Arg Glu Tyr Thr Lys Leu Ala

85 90 95

Glu Glu Ser Lys Lys Arg Ala Arg Gln Tyr Gly Gln Leu Ala Asp Met

100 105 110

Glu Ala Gly Arg Ile Gly Gln His Glu His Met Gln Gln Glu Trp Asn

115 120 125

Ser Lys Ala Arg Glu Ser Glu Ala Gln Cys Lys Ala Thr Glu Ala Lys

130 135 140

Ala Gln Glu Glu Tyr Thr Lys Ala Arg Asp Glu Arg Gln Lys Ser Leu

145 150 155 160

Val Ser Asn Ala Glu Ala Ala Met His Asp Ala Gln Ala Thr Val Asp

165 170 175

Thr Met Lys Ser Glu Arg Ala Tyr Glu Ile Gly Lys Glu Leu Met Arg

180 185 190

Lys Ala Glu Asn Ala Arg Asn Asp Ala Ser Asn His Tyr Gln Arg Ala

195 200 205

Lys Glu Asn Arg Glu Arg Ala Asn Ser Glu Thr Val Lys Ser His Gln

210 215 220

Gln Ala Gln Asp Ala Gln Arg His Asn Ala Ala Ser Lys Ala Tyr Gln

225 230 235 240

Gln Asp Gly Leu Arg Thr Arg Met Ala Ser Arg Ile Asn Ile Met Lys

245 250 255

Tyr Ile Gln Ser Ser Leu Leu Ala Glu Arg Ala Ala Asn Gln Ala Arg

260 265 270

Ile Glu Gln Leu Lys Ser Glu Trp Tyr Glu Lys Ala Ala Asn Glu Tyr

275 280 285

Ser Arg Met Ser Glu Glu Asn Ala Ala Ile Ser Lys Leu Ala Gly Ser

290 295 300

Glu Glu His Tyr Phe Ala Gln Arg Ala Lys Arg Asn Glu Gly Lys Ala

305 310 315 320

Tyr Glu Leu Ser Gln Ser Lys Arg Met Met Gly Ser Glu Ala Ala Ala

325 330 335

Ala Gly Glu Leu Leu Ala Met Ser Gln Ala Lys Asp Asp Glu Thr Glu

340 345 350

Asp Glu Lys His Phe Asp Phe Pro Ile Tyr Glu Ser Asp Asp Pro Thr

355 360 365

Lys Leu Ser Pro Ser Pro Asp Glu Lys Asp Leu Thr Tyr Gly Ser Gly

370 375 380

Glu Gly Leu

385

<210> 20

<211> 360

<212> PRT

<213> Tenodera australasiae

<400> 20

Met Lys Phe His Ile Ala Phe Val Leu Leu Val Ile Phe Gly Ala Ala

1 5 10 15

Gln Ala Gly Lys Lys His Glu Val Met Thr Tyr Gly Ser Gly Tyr Lys

20 25 30

His Met Gly Gly Glu Thr Tyr Glu Asp Val Gly Thr Gly Asn Arg Leu

35 40 45

Gly Ser Thr Ala Phe Asp Ile Met Glu Ala Ala Asp Glu Asn Thr Glu

50 55 60

Arg Ala Ser His Thr Phe Gly Ser Lys Ser Ala Ala Tyr Ser Ser Asp

65 70 75 80

Ala Asp Leu Phe Ile Glu Leu Leu Arg Glu Lys Arg Glu Thr Arg Ala

85 90 95

Asn His Gly Lys Arg Ala Glu Ser Gln Ala Val Leu Ala Asn Glu Ser

100 105 110

Tyr Gln Lys Ser Gln Leu His Lys Arg Gln Ala Lys Asp Lys Gln Ala

115 120 125

Ile Ser Lys Glu Tyr Glu Glu Arg Ala Gln Lys His Asp Arg Leu Ser

130 135 140

Lys Glu Gln Asp Met Lys Glu His Asp Asp Tyr Arg Lys Ser Asn Ala

145 150 155 160

Glu Asp Thr Glu Leu Arg Asn Ser Val Glu Arg Ser Asn Tyr Asp His

165 170 175

Val Met Ala Leu Gly Tyr His Glu Leu Ser Gln Leu Glu Met Gly Glu

180 185 190

Thr Asn Gln Cys Glu Gln Leu Ser Arg Glu Leu Gln Ser Arg Ala Glu

195 200 205

Glu Tyr Phe Asn Leu Ala Lys Glu Leu Lys Glu Lys Ala Lys Lys Glu

210 215 220

Lys Glu Asn Ala Arg Ile Lys Lys Ala Lys Ala Lys Glu Glu Glu Ala

225 230 235 240

Arg Ala Glu Glu Tyr Glu Asn Ala Phe Thr Glu Asn Ser Lys Lys Val

245 250 255

Leu Thr Tyr Lys Phe Tyr Glu Leu Glu Phe Gly Met Lys Ala Leu Asn

260 265 270

Glu His His Gln Ala Glu Ser Ala Arg Val Arg His His Phe Leu Gln

275 280 285

Ile Leu Glu Gln His Asn Ser Gln His Ala Asp Met Leu Trp Gly Tyr

290 295 300

Ala Gln Gln Glu Asp Lys Asp Gly Arg Ser Phe Thr Gln Tyr Ala Thr

305 310 315 320

Glu Leu Ser Lys Gln Thr Lys Met Leu Thr Ala Thr Ala Ala His Leu

325 330 335

Met Lys Gln His Arg Tyr Thr Gly Met Glu Met Tyr Ser Lys Gln Pro

340 345 350

Phe Pro His Ser Asn Tyr His Gly

355 360

<210> 21

<211> 383

<212> PRT

<213> Archimantis monstrosa

<400> 21

Met Asp Ser Lys Ile Leu Cys Val Ser Leu Leu Leu Ala Val Phe Cys

1 5 10 15

Leu Trp Tyr Thr Glu Ala Ser Pro Leu Glu Glu Lys Tyr Asp Glu Lys

20 25 30

Ser Glu Ala Asp Asp Tyr Gln Ser Glu Asp Ser Ser Ala Ala Ile His

35 40 45

Asp Gln Thr Thr Lys Ile Ala Thr Asn Ala Val Lys Thr Tyr Ala Asn

50 55 60

Lys Ala Lys Ala Thr Glu Ser Lys Ala Lys Leu Tyr His Gln Tyr Ser

65 70 75 80

Lys Asp Arg Ala Tyr Tyr Ser Arg Glu Tyr Glu Lys Met Gly Glu Glu

85 90 95

Tyr Met Lys Lys Ser Lys Glu Tyr Glu Gln Leu Tyr Ile Ala Glu Ala

100 105 110

Ala Arg Ile Ser Leu His Glu Asn Lys Gln Lys Glu Trp Asp Thr Lys

115 120 125

Gly Arg Glu Ala Asn Val Gly Ile Arg Glu Tyr Glu Thr Lys Ser Gln

130 135 140

Gln Ala Ser Ser Lys Lys Asn Glu Leu Leu Glu Glu Ser Ile Ile Ala

145 150 155 160

Ala Val Gln Ala Ala Ile His Glu Thr Gln Ala Thr Gly Tyr Leu Leu

165 170 175

Lys Ser Glu Ala Ala Asn Gly Ile Ala Arg Asn Met Leu Gln Ile Ala

180 185 190

Glu Ser Ile Arg Asp Glu Ala Ser Asn His Tyr Gln Ile Gly Lys Glu

195 200 205

Glu Leu Asn Arg Ala Thr Ala Gln Lys Val Lys Ala Gln Gln Gln Ala

210 215 220

Glu Asp Ser Gln Arg His His Ala Ala Ala Arg Ala Tyr Gln Gln Asp

225 230 235 240

Ser Leu Arg Ser Arg Met Ala Ser Arg Ala Asn Asn Met Gln Tyr Met

245 250 255

Gln Asn Ser Leu Leu Ala Glu Arg Ala His Ser Leu Ser Thr Glu Asn

260 265 270

Thr Leu Glu Ser Glu Leu Tyr Gly Lys Glu Ala Asp Glu Leu Ala Lys

275 280 285

Met Ser Glu Glu Ser Ala Ala Ile Ser Lys Ile Cys Ser Gly Glu Glu

290 295 300

Arg Ser Tyr Arg Asn Met Ala Lys Gln Ser Glu Val Lys Ala Tyr Glu

305 310 315 320

Tyr Ser Val Ser Lys Asn Met Met Gly Ala Asp Met Thr Asp Thr Ala

325 330 335

Ala Met Ala Asn Gly Asp Glu Ala Lys Gln Gly Asp Asp Glu Glu Gln

340 345 350

Gln Met Tyr Arg Ser Pro Asn Ile Pro Ala Glu Asp Ser Thr Lys Asn

355 360 365

Leu Ser Tyr Asn Leu Lys Asp Ser Thr Glu Gly Glu Gly Met Ser

370 375 380

<210> 22

<211> 369

<212> PRT

<213> Archimantis monstrosa

<400> 22

Met Lys Phe His Ile Val Phe Val Leu Leu Val Val Phe Gly Ala Ala

1 5 10 15

Gln Ala Gly Lys Lys His Glu Ala Leu Thr Phe Gly Ser Gly Tyr Lys

20 25 30

Ser Thr Tyr Gly Glu Gly Glu Thr Phe Asp Asp Glu Asp Asp Gln Ala

35 40 45

Leu Arg Asn Glu Arg Val Pro Val Gly Ala Leu Ser Ala Ala Ile Ile

50 55 60

Asn Pro Tyr Ala Leu His Ser Glu Glu Gly Arg Ile Ala Tyr Asp Thr

65 70 75 80

Ser Ser Gln Tyr Tyr Ala Asn Lys Ala Glu Gly Ser Ala Asp Leu Ser

85 90 95

Arg Glu Lys Lys Gln Met His Gly Glu Tyr His Gly Lys Ala Ala Thr

100 105 110

Tyr Ala Ser Arg Ala Asn Glu Ala Tyr Lys Lys Ser Gln Leu His Lys

115 120 125

Arg Gln Ala Lys Asp Lys Gln Ala Ile Ala Lys Glu Tyr Glu Glu Arg

130 135 140

Ala Gln Lys His Glu Ser Arg Ser Lys Ala Leu Asp Val Arg Asp Gln

145 150 155 160

Asp Asp Glu Arg Lys Ser Val Thr Glu Met Glu Glu Tyr Ala Arg Ala

165 170 175

Leu Lys Ile Ala Asn Leu Ala Leu Val Phe Ala Gly Ile Tyr Gln Glu

180 185 190

Thr Gly Arg Leu Gln Leu Glu Ala Thr Asn Val Phe Glu Gln Phe His

195 200 205

Lys Met Leu Ser Thr Lys Gly Glu Glu Tyr Lys Lys Gln Ala Glu Glu

210 215 220

Tyr Lys Glu Lys Ala Asn Lys Glu Lys Glu Glu Ala Ala Ile Gln Gln

225 230 235 240

Ala Lys Ser Lys Glu Leu Asn Ala Lys Ala Gln Glu Tyr Glu Asn Ile

245 250 255

Phe Ile Glu Ser Ser Lys Lys Leu Ala Ala Asn Arg Tyr Tyr Glu Leu

260 265 270

Glu Phe Lys Met Lys Ala Glu Asn Glu Arg His His Ala Glu Leu Ala

275 280 285

Arg Ile Arg Ser Arg Phe Leu Ser Arg Leu Ala Asn Tyr Asn Arg Glu

290 295 300

Gln Ala Glu Ala Val Leu Arg Phe Ala Arg Ser Glu Arg Lys Asp Gly

305 310 315 320

Glu Ile Phe Arg Arg Asn Ala Ile Glu Leu Tyr Lys Glu Thr Arg Ala

325 330 335

Leu Ala Ala Thr Ala Ala Arg Val Met Lys Gln His Arg Tyr Thr Gly

340 345 350

Gln Glu Ile Tyr Thr Lys Gln Pro Phe Pro His Ser Asn Tyr His Gly

355 360 365

Ala

<210> 23

<211> 385

<212> PRT

<213> Pseudomantis albofimbriata

<400> 23

Met Asp Ser Lys Met Leu Cys Val Ser Leu Leu Leu Ala Val Phe Cys

1 5 10 15

Leu Trp Tyr Thr Glu Ala Ser Pro Leu Glu Glu Lys Tyr Gly Glu Lys

20 25 30

Tyr Gly Asp Met Glu Glu Tyr Gln Arg Gly Thr Glu Asp Ser Arg Ala

35 40 45

Val Ile Asn Asp His Thr Ala Lys Val Ala Ser Gln Ser Ala Arg Gly

50 55 60

Met Val Asn Lys Ala Lys Thr Thr Glu Ala Ala Ala Arg Ser Asn Glu

65 70 75 80

Gln Leu Ser Lys Asp Arg Gln Tyr Tyr Tyr Arg Glu Tyr Leu Lys Lys

85 90 95

Ala Asp Tyr His Lys Lys Lys Ala Leu Glu Tyr Glu Gln Leu Ser Ala

100 105 110

Ala Glu Asn Ala Lys Ile Ala Tyr His Glu Ser Lys Gln Lys Asp Trp

115 120 125

Glu Thr Lys Ala Arg Glu Ser Asp Val Gln Cys Arg Asp Ala Glu Ala

130 135 140

Lys Tyr Glu Gln Ser Tyr Thr Arg Ser Arg Glu Leu Lys Arg Glu Ser

145 150 155 160

Ile Ile Ala Tyr Val Gln Ala Ala Met His His Ala Glu Ala Ser Gly

165 170 175

Asp His Met Lys Ala Asp Arg Ala Lys Asp Ile Ala Arg Asp Met Met

180 185 190

Arg Lys Ala Glu Ser Leu Arg Gly Asp Ala Ser Asn His Tyr Gln Arg

195 200 205

Ser Glu Glu Asp Lys Asn Lys Ala Arg Ser Glu Lys Val Lys Ala His

210 215 220

Gln Asn Ala Asp Asn Ser Gln Arg His His Thr Ala Cys Arg Ala Tyr

225 230 235 240

Asp Gln Glu Gly Leu Lys Thr Arg Leu Ser Ser Lys Ala Asn Met Met

245 250 255

Arg Gln Ile His Ser Ser Leu Leu Ala Glu Arg Ser His Ser Leu Ala

260 265 270

Arg Glu Asp Gly Leu Ala Ala Asp Leu Ser His Lys Leu Ala Glu Glu

275 280 285

Leu Ala Arg Met Ser Glu Glu Ser Gly Ala Ile Ser Lys Ile Asn Ser

290 295 300

Gly Glu Glu Arg Gly Tyr Ser Asn Lys Val Arg Gln Asp Glu Val Lys

305 310 315 320

Ala His Glu Leu Ala Val Ser Lys Arg Met Met Gly Ala Glu Val Ala

325 330 335

Asp Asn Ser Glu Met Ile Ser Leu Ala Gln Ala Lys Asp Gly Ser Leu

340 345 350

Asp Glu Gly Glu Asn Tyr Lys Leu Ser Thr Phe Tyr Ala Asp Asp Ser

355 360 365

Thr Lys Asn Met Leu Pro Asp Ser Arg Gly Gln Met Ser Tyr Gly Asp

370 375 380

Glu

385

<210> 24

<211> 366

<212> PRT

<213> Pseudomantis albofimbriata

<400> 24

Met Lys Phe His Ile Ala Phe Val Leu Leu Val Val Phe Gly Ala Ala

1 5 10 15

Gln Ala Gly Lys Asn His Glu Val Met Thr Tyr Gly Ser Gly Tyr Lys

20 25 30

Thr Met Gly Asp Glu Gly Gly Ser Gly Val Gly Asn Glu Gly Glu Asp

35 40 45

Tyr Gln Asp Asn Glu Gly Ala Thr Ala Ala Thr Ile Leu Asp Glu Ser

50 55 60

Thr His His Thr Glu Glu Ala Arg Asp Ile Phe Gly Thr Arg Ser Glu

65 70 75 80

Ala His Ala Tyr Ser Ala Glu Met Phe Ala Asp Leu Val Arg Glu Lys

85 90 95

Arg Gln Ala Ser Ile Glu Ser His Lys Lys Ala Glu Asp Tyr Ala Val

100 105 110

Arg Ala Asn Glu Glu Tyr Lys Lys Ser Gln Leu Leu Lys Arg Gln Ala

115 120 125

Arg Asp Lys Gln Ala Ile Ala Lys Gln Tyr Glu Glu Lys Ala Gln Lys

130 135 140

Tyr Asp Arg Ile Ser Lys Gln Gln Asp Ile Lys Glu Gln Asp Asp Tyr

145 150 155 160

Arg Lys Ser Asp Ala Glu Ser Glu Glu Tyr Lys Arg Ser Ile Ala Val

165 170 175

Ala Asn Ala Ala Leu Ala Leu Ala Ser Ala Tyr Glu Glu Ala Ser Arg

180 185 190

Met Glu Leu Asp Ala Thr Gly Glu Met Glu Gln Gln Ser Lys Glu Leu

195 200 205

Tyr Thr Lys Ser Glu Glu Tyr Asn Lys Val Ala Glu Glu Cys Ile Thr

210 215 220

Arg Ala Lys Lys Glu Lys Glu Leu Ala Arg Ile Glu Glu Ala Lys Gly

225 230 235 240

Lys Glu Ala Glu Ala Lys Ser Gln Glu Tyr Glu Asn Phe Ala Thr Asp

245 250 255

Asn Asn Lys Lys Tyr Asn Ala Met Lys Phe Tyr Gly Trp Glu Phe Lys

260 265 270

Met Lys Ala Glu Asn Glu Arg His Asn Ala Asp Tyr Cys Arg Ile Lys

275 280 285

Ser Arg Tyr Leu Ala Gln Leu Ser Asn Tyr Asn Arg Glu Gln Ala Glu

290 295 300

Ala Leu Tyr His Phe Ala Ala Ala Gln Arg Lys Asp Ala Glu Leu Phe

305 310 315 320

His Arg Tyr Ala Met Glu Leu Tyr Lys Gln Thr Arg Val Leu Thr Ala

325 330 335

Ser Ala Ala Gln Ile Met Lys Gln His Arg Tyr Thr Gly Gln Glu Ile

340 345 350

Tyr Ser Lys Gln Pro Phe Pro His Ser Asn Tyr His Gly Ala

355 360 365

<210> 25

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 25

ggaattccca tgggctctcc cttggaagac aaatac 36

<210> 26

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 26

cggcggatcc ttattacaga ccttcgccgg aac 33

<210> 27

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 27

ggaattccca tgggcaagaa acatgaagta atga 34

<210> 28

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 28

cggcggatcc ttattatccg tggtagttgg agtgg 35

<210> 29

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 29

ggaattccca tgggctctcc cttggaagaa aaatatg 37

<210> 30

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 30

cggcggatcc ttattaactc attccttcac cttcagtt 38

<210> 31

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 31

ggaattccca tgggcaagaa acacgaagca 30

<210> 32

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 32

cggcggatcc ttattatgct ccgtggtagt tgga 34

<210> 33

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 33

ggaattccca tgggctcacc cttggaagaa aaatat 36

<210> 34

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 34

cggcggatcc ttattattca tcgccgtaag acattt 36

<210> 35

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 35

ggaattccca tgggcaagaa tcacgaagta atg 33

<210> 36

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 36

cggcggatcc ttattatgct ccgtggtagt tggag 35

<210> 37

<211> 837

<212> DNA

<213> Pseudomantis albofimbriata

<400> 37

atggctggcc tggtgaaatt cggtcttctt ctcattcttg gagttttaca ggtctctgca 60

ggtgaataca aagatataca gcttggaaat atagcatgtg aagttaaatg ttttgaaaaa 120

accagtgctc tcacggaggg taatcaaaga tccgagagga ttatgaagag cgactcggaa 180

atcaaagtaa agctgattca ggctgtagaa ggtgttatat cagcagctga gcagtgctgt 240

ggaaaggaat ctgaagaaaa gaaatgccta gaagaagttg tttatcccgc actgctaaag 300

atcatagatc ctctgagtat attcttccac gttaacgttg gactcttcac gcagacttta 360

gcaaatttat cattggttgc aacctacgct cagcgtgaac tagattgttg ttgcaagctt 420

gaggattgtg tgcagaaaat tgaatttcct actgccagct acaagggagg tattaaagaa 480

tgtaaaggtg aatactgcgt tcgtaacctc aatcagttgt tggcgaaact ccgcgattat 540

ttgccagcga tggttaccag caaacttcta gaggatcgac tgattagttc ttacttcaat 600

tttcaaaagt acaaggaaca agtccctaaa gaaggaaaat gcaaggagga aaagggattg 660

gctgaactca acaaggaaat aaatagtttt aatttacttg tgaaatatta tagttcgtac 720

caagtagcat taaattcgtt cgatgaggag atattgagtc ttacatctgc tgctgcaata 780

gaaattctgg gatgtgccag cgaaaaagat ataaagcaat acatcccatg taattaa 837

<210> 38

<211> 729

<212> DNA

<213> Pseudomantis albofimbriata

<400> 38

tttggtggtg ctgcgttctc tgagggtatt agcggtggtg ctcgaaccag taggatcgaa 60

accgacgagg aactggatag ctcctctatg aaaattgaga aggtgagaac ttaccctgta 120

atgcgtgatg tggtggtcaa gaaagaaatt cccaagactg tgcatgaacc ttatcccgtg 180

atggtggaca aaccataccc cgtcaaagtc ccaacaccct ataaagttac gtactctaaa 240

gaaatcctta agccagtcga gaggaaggtt ccaatgccct atccagttaa gcagccctat 300

cctgtataca tacaccataa agtacctttt ttcatcgaga agaaggtgcc atatcccgtc 360

aggttaccct ataagcactt ctacgcagtt cccaaacctt accccgtcta ctacgagaag 420

aaggttccct atccagtcga gaagaaggtg ccatacgtga taaaagttcc ctacagggtc 480

ccctattacg ttgacaagcc ctacgccgtt cccgtcgtga agaaggtgcc ttacgaggta 540

gagaagaagg taccttatcc agtcaaggtc cccttcgcag ctccctatcc agttactaag 600

atcgtgacag tcaaagaaga agtccccgtg tacaaacccg tcaagtacag caaattggtt 660

accatcaaga agaaggttcc cgtcataaag aaggagaaag agtttttcga agtaaagaag 720

aagtactag 729

<210> 39

<211> 900

<212> DNA

<213> Pseudomantis albofimbriata

<400> 39

aatgacagaa agaattactt cattgaatac agcaaactgg cgaagtatat caaggaaacc 60

cacaagaaga ctcaagagaa ggcagctata gaggaagctt cattggaggg tatcgaaaac 120

aggcgtaaag agtgggaggt caagtcgagg gaacctttga ccaacgctaa gagctatgaa 180

atacgtgcgc agaaattata tgctgatgca cacgaattgt ggagggggat agagacatct 240

gttatagctg cagctgctgc agaggcagaa gcctcaggct atgacaatat tgccaacaag 300

gaacgaaact ggagcatgga tatgaagagc aaagccgaag aaatgctcag cgaatccgct 360

agataccata atacgaaagc aaaggaagct gagcagaggg caatgggtaa caggattaag 420

gcccaacagc acgccgagga agcacagaga cattttcagg ccgcaaaagc cctagaggaa 480

caggccctga aggccgaagc agcatctctg gataacgaag ccaagtacat tgccttgaca 540

attcaggcag aggaagccgc aaacatgtgc aaagaggatg agatgtgtgc cacaatggaa 600

gaaatgatgc acagcgagct aaacaagcga gcacaaatag ctgatggaat gtacaaagta 660

atccagaaag aagtggagtt ttacgaacat aaggcaaaga aacaggaggg tgaggcaacc 720

aggtactctc aaagtgccca tcaaattggt agcgatttcg tttcccatag cgcttcaatg 780

ttaaagttgg caacggccgg ggaggattcc gaaatagaac atgacgagga aatgccttat 840

aaccagaaat taaaaactat gcctttcgtg gctcaaaaaa atgatcgatt aattttgtga 900

<210> 40

<211> 627

<212> DNA

<213> Tenodera australasiae

<400> 40

atgaagacct tgatctgcct cttgctgctt tccggcaccg cattggctgg gaaatatgaa 60

gaaattgaaa aagtgaagtt catcgataga gtcaaagagg tcaaagtacc agtttctaaa 120

cctgttccct acgctgtgaa tcaagccgtt ctgtatccag ttgacgtacc ttatcccata 180

acagttgaaa aacaggtacc ttataaggtg atacatgaaa aagagcatat agtgccgata 240

catcatgcct attccgctgt gatgaaggag aaacagccat tctatgtgga aaagaaagtc 300

ccctatccag tggaagtgcc agtaaagatt cccgagcctt acaaggtaga ggtgccagtg 360

ccggtacctc gcccatatgt ggtaaaggag attgaggagc gtgaagttaa ggtgcctgta 420

aaggtggaaa aagccgtccc ctatccaaaa cctgaaccgt atcgcgtata ccttcccaaa 480

ggctactcag acaagctaga cactagattg ctagatgaac atacctcata ccctttcaag 540

ggacatgaag gcattgaagc ttcccacgtc gatttgcatg agaagatagg tgaactggat 600

atacacaaaa agaaagagaa gcgataa 627

<210> 41

<211> 801

<212> DNA

<213> Tenodera australasiae

<400> 41

atggctatcc tgttgaacct agccgttttc tgcatcttcg gagccataca ggtttccagt 60

gaacagcctc ctaaaaccac agagctcaca gttgtgacaa gagaattaaa atgttgcgat 120

ggaatacaag aactatttga atatattaaa ggaaggcaca cattgaatga cagcatagaa 180

gcaaaggttg gctccattcc cgctcttgat actgaaatca ggcggttaca gagactaggt 240

tgcggaaaag aactcgaaga aaaaaagaac cacgaggaag ccagtgaagc tgtaaagaaa 300

attgaagaga ttttaataaa attgttggag gtctccgacc gactactaac agatacgctg 360

aaattgctca gtcttctggg ttttgcttca aatgatgaga gatgcatttg caagaattat 420

gcatgtgttc agaaacttga atatcccgag aaaaaagtaa aggaagtaaa agaatgtaaa 480

gaaggaaaat gtaataagag actcagagag ctactcaatt ccttgtaccg aatgcaagct 540

gatttagaag actgtgttca actgtatgaa cttgctgaac aaaacctgca gcaagtgtca 600

aaactgaggg aagaagtcca gaaagatacg gagtgtaagg atgaaaagtt ggcgaatgac 660

tataaggaaa aactgagcct ctcttcgctg tacttgtcga aagccaacta cttagattat 720

aaagttgttt atgaacgcat tccttatctg ttgaaagaat gccaagaaac tattaaggaa 780

attaattctt gtggaaaata a 801

<210> 42

<211> 606

<212> DNA

<213> Tenodera australasiae

<400> 42

atgaagaaat attcagtgca gttggcagca ttggtgatac tttgcgcgat aagcttttgt 60

tcagctggag aggagaagaa agaggaaaag aaagaggaga agaaagagaa gaaatgtcca 120

ccagagggag ttgaactcaa acctgttgag ctctctgaag aggatatcaa gcatcttgtt 180

caccactggc atttggtgaa acataacaat aatttgtacg gcaaagattt gtcctgctct 240

gaatacacgt atggcaagtc tgaaaaaggc gtcttcttcc atggttcagc ctataacgaa 300

agctccaaca aggtgtgggc cacaaccgga gtttttaccc agaaggaacg aggcgtgtac 360

gaaggagtct ttgaagacaa acgggtcaat ggagagttct tgttcttgta tgtgagcaag 420

aaaatataca ttttggccgg ctgcataaat gacaaagaat ttgcactctg ggtaaagacc 480

gtagatggga agcttgataa agagcaagaa gagattgttg aggcgacact caagaagaac 540

ggcttgccaa acgacttaaa acaggctcct aatgataact gtgccaagaa gaagtccagt 600

gtctaa 606

<210> 43

<211> 1155

<212> DNA

<213> Tenodera australasiae

<400> 43

atggatatcc ggttcctgct tcttttcatt ctccctgctg tggcaacggc tggtacctat 60

tcagatgaga ccgtgtcttt gggaatggat aatgaagata aagaatttga caactcaggc 120

agttataaga cggaaggtga atatagcgtc gctggttccg tgagcgcaga catttttaac 180

gagggcagcc acttggctgg cgccgcggca gaagaatggg gcatgaaggc taagacggtt 240

gagcaaaatg ccaaatggtt cagaaacttg ctcaatgaga agcagatgat tttgggtcag 300

gtgaagaaga ggcttggctt actgcatgcc ttggccacgg aaaacctgca gaatgcagaa 360

caatatcgtc tgtctgctgc cgagaagtgc caatcctcca aagctttaga gaccaaggcg 420

cgagagtccc acaatttgaa caaggattat gctgttcaag tacatgaact gagggctaag 480

ctgaacaatg aattgaaaca gcacctggac gctctcatca gtagtcgtgt ggctgaggct 540

aacgccaata ccgatatagc tgaggcggag atacagagaa gctgggcaga gttcttggaa 600

aagcaagcac aggatatgct atctgaagct gccaaatggg agcgaagctc taccgaaaat 660

ctgaactcgg ccaagaaatt caagttggaa gctgctgcag aaaaggccaa ggccaaggag 720

ctcgagaaga aacgtaagga ggcattggat gctgctatgg aggctgagag gaaagcacta 780

gatttctatc agttccacct ggagttccac atgaaggcac aggcagaaga acatcaaggt 840

gggcagatca aagtcaaagc caagtacctc caaatgatgg ccgcttcgtt gttggatagt 900

gcttcgcagc ttcgtatgtt tgccaacctg caaaataaag acaaatcact ctgggagaag 960

aagacagcaa gtgcaaaggc agaggctaac ctccttgagg cctcggcgga taagatgctc 1020

cagaatatcg aaaaactcaa aaagtcattg caggcgtcat ctaaggtggc acagaaaagt 1080

aacgaattag ttcctcagga agaaaagagc gttaagagtg agccaagagc aataactcca 1140

gataaaaagg tgtag 1155

<210> 44

<211> 810

<212> DNA

<213> Tenodera australasiae

<400> 44

atggcgatcc cgattaaact tgcactacta ttctccatct tgggagtcgc acaggtgtgc 60

tcttctgaat acaaaactga tgacctagat atacgtactt acgaagtgaa atgccatcaa 120

aaagttgaaa tgctggtgca gaatgtgcaa aacagctcat cgatttatgc aagggcatct 180

gaatcgaaag aacaccttat caagattggc acacatttgg cagaattact cgtaaagtat 240

gcttgtgaaa aagaaatcac tgacaagaag ggtatagaag aaataattgc tactctgcaa 300

gagttcagga gctctttgga aaattatata caagaagttg ccgtcaaata ttttggagct 360

ttgtacagca ctaccttgta ttcaacgtgg gctttggaag agattcattg ttgttgcagc 420

aaacttaata tctgcctaaa gaaatatgag taccccataa gcaagaacga ggtcaaagaa 480

tgtaagccta acaaatgtaa caaggcggtt gttggcggtg gatataaaga actaataaca 540

tggttatcaa aggtagaaga atcggaatgg ctactgaaag aattaattag gctgattgag 600

gaactgctga aatacaaaga tgaatacatc aaagatccaa aatgtgcgaa cgaaaaggct 660

gctgcactta taggagagct cgaccagaga attgtgctga ttgtatgtga cttgaactac 720

cgccagcatg aattggtaaa atatgattta cttatgttgc acagggccga cgaaattttg 780

acgggcttca aagaatgtaa aaaggagtga 810

<210> 45

<211> 195

<212> DNA

<213> Tenodera australasiae

<400> 45

gacaacttca gccaggccca agtgttctgg aataacgtgc tgaagccaga cgagaaggaa 60

cgttttgtgg agaacgttta tgaaactctt aagcttgtag aaccagaatt gcagctggac 120

gttacgaagt attttggtgc agttcaccca gatatacggg aaatgctggc tgcgaaattg 180

gagcttaaga cataa 195

<210> 46

<211> 278

<212> PRT

<213> Pseudomantis albofimbriata

<400> 46

Met Ala Gly Leu Val Lys Phe Gly Leu Leu Leu Ile Leu Gly Val Leu
1 5 10 15

Gln Val Ser Ala Gly Glu Tyr Lys Asp Ile Gln Leu Gly Asn Ile Ala
20 25 30

Cys Glu Val Lys Cys Phe Glu Lys Thr Ser Ala Leu Thr Glu Gly Asn
35 40 45

Gln Arg Ser Glu Arg Ile Met Lys Ser Asp Ser Glu Ile Lys Val Lys

50 55 60

Leu Ile Gln Ala Val Glu Gly Val Ile Ser Ala Ala Glu Gln Cys Cys

65 70 75 80

Gly Lys Glu Ser Glu Glu Lys Lys Cys Leu Glu Glu Val Val Tyr Pro

85 90 95

Ala Leu Leu Lys Ile Ile Asp Pro Leu Ser Ile Phe Phe His Val Asn

100 105 110

Val Gly Leu Phe Thr Gln Thr Leu Ala Asn Leu Ser Leu Val Ala Thr

115 120 125

Tyr Ala Gln Arg Glu Leu Asp Cys Cys Cys Lys Leu Glu Asp Cys Val

130 135 140

Gln Lys Ile Glu Phe Pro Thr Ala Ser Tyr Lys Gly Gly Ile Lys Glu

145 150 155 160

Cys Lys Gly Glu Tyr Cys Val Arg Asn Leu Asn Gln Leu Leu Ala Lys

165 170 175

Leu Arg Asp Tyr Leu Pro Ala Met Val Thr Ser Lys Leu Leu Glu Asp

180 185 190

Arg Leu Ile Ser Ser Tyr Phe Asn Phe Gln Lys Tyr Lys Glu Gln Val

195 200 205

Pro Lys Glu Gly Lys Cys Lys Glu Glu Lys Gly Leu Ala Glu Leu Asn

210 215 220

Lys Glu Ile Asn Ser Phe Asn Leu Leu Val Lys Tyr Tyr Ser Ser Tyr

225 230 235 240

Gln Val Ala Leu Asn Ser Phe Asp Glu Glu Ile Leu Ser Leu Thr Ser

245 250 255

Ala Ala Ala Ile Glu Ile Leu Gly Cys Ala Ser Glu Lys Asp Ile Lys

260 265 270

Gln Tyr Ile Pro Cys Asn

275

<210> 47

<211> 242

<212> PRT

<213> Pseudomantis albofimbriata

<400> 47

Phe Gly Gly Ala Ala Phe Ser Glu Gly Ile Ser Gly Gly Ala Arg Thr

1 5 10 15

Ser Arg Ile Glu Thr Asp Glu Glu Leu Asp Ser Ser Ser Met Lys Ile

20 25 30

Glu Lys Val Arg Thr Tyr Pro Val Met Arg Asp Val Val Val Lys Lys

35 40 45

Glu Ile Pro Lys Thr Val His Glu Pro Tyr Pro Val Met Val Asp Lys

50 55 60

Pro Tyr Pro Val Lys Val Pro Thr Pro Tyr Lys Val Thr Tyr Ser Lys

65 70 75 80

Glu Ile Leu Lys Pro Val Glu Arg Lys Val Pro Met Pro Tyr Pro Val

85 90 95

Lys Gln Pro Tyr Pro Val Tyr Ile His His Lys Val Pro Phe Phe Ile

100 105 110

Glu Lys Lys Val Pro Tyr Pro Val Arg Leu Pro Tyr Lys His Phe Tyr

115 120 125

Ala Val Pro Lys Pro Tyr Pro Val Tyr Tyr Glu Lys Lys Val Pro Tyr

130 135 140

Pro Val Glu Lys Lys Val Pro Tyr Val Ile Lys Val Pro Tyr Arg Val

145 150 155 160

Pro Tyr Tyr Val Asp Lys Pro Tyr Ala Val Pro Val Val Lys Lys Val

165 170 175

Pro Tyr Glu Val Glu Lys Lys Val Pro Tyr Pro Val Lys Val Pro Phe

180 185 190

Ala Ala Pro Tyr Pro Val Thr Lys Ile Val Thr Val Lys Glu Glu Val

195 200 205

Pro Val Tyr Lys Pro Val Lys Tyr Ser Lys Leu Val Thr Ile Lys Lys

210 215 220

Lys Val Pro Val Ile Lys Lys Glu Lys Glu Phe Phe Glu Val Lys Lys

225 230 235 240

Lys Tyr

<210> 48

<211> 299

<212> PRT

<213> Pseudomantis albofimbriata

<400> 48

Asn Asp Arg Lys Asn Tyr Phe Ile Glu Tyr Ser Lys Leu Ala Lys Tyr
1 5 10 15

Ile Lys Glu Thr His Lys Lys Thr Gln Glu Lys Ala Ala Ile Glu Glu
20 25 30

Ala Ser Leu Glu Gly Ile Glu Asn Arg Arg Lys Glu Trp Glu Val Lys
35 40 45

Ser Arg Glu Pro Leu Thr Asn Ala Lys Ser Tyr Glu Ile Arg Ala Gln
50 55 60

Lys Leu Tyr Ala Asp Ala His Glu Leu Trp Arg Gly Ile Glu Thr Ser
65 70 75 80

Val Ile Ala Ala Ala Ala Ala Glu Ala Glu Ala Ser Gly Tyr Asp Asn
85 90 95

Ile Ala Asn Lys Glu Arg Asn Trp Ser Met Asp Met Lys Ser Lys Ala

100 105 110

Glu Glu Met Leu Ser Glu Ser Ala Arg Tyr His Asn Thr Lys Ala Lys

115 120 125

Glu Ala Glu Gln Arg Ala Met Gly Asn Arg Ile Lys Ala Gln Gln His

130 135 140

Ala Glu Glu Ala Gln Arg His Phe Gln Ala Ala Lys Ala Leu Glu Glu

145 150 155 160

Gln Ala Leu Lys Ala Glu Ala Ala Ser Leu Asp Asn Glu Ala Lys Tyr

165 170 175

Ile Ala Leu Thr Ile Gln Ala Glu Glu Ala Ala Asn Met Cys Lys Glu

180 185 190

Asp Glu Met Cys Ala Thr Met Glu Glu Met Met His Ser Glu Leu Asn

195 200 205

Lys Arg Ala Gln Ile Ala Asp Gly Met Tyr Lys Val Ile Gln Lys Glu

210 215 220

Val Glu Phe Tyr Glu His Lys Ala Lys Lys Gln Glu Gly Glu Ala Thr

225 230 235 240

Arg Tyr Ser Gln Ser Ala His Gln Ile Gly Ser Asp Phe Val Ser His

245 250 255

Ser Ala Ser Met Leu Lys Leu Ala Thr Ala Gly Glu Asp Ser Glu Ile

260 265 270

Glu His Asp Glu Glu Met Pro Tyr Asn Gln Lys Leu Lys Thr Met Pro

275 280 285

Phe Val Ala Gln Lys Asn Asp Arg Leu Ile Leu

290 295

<210> 49

<211> 208

<212> PRT

<213> Tenodera australasiae

<400> 49

Met Lys Thr Leu Ile Cys Leu Leu Leu Leu Ser Gly Thr Ala Leu Ala

1 5 10 15

Gly Lys Tyr Glu Glu Ile Glu Lys Val Lys Phe Ile Asp Arg Val Lys

20 25 30

Glu Val Lys Val Pro Val Ser Lys Pro Val Pro Tyr Ala Val Asn Gln

35 40 45

Ala Val Leu Tyr Pro Val Asp Val Pro Tyr Pro Ile Thr Val Glu Lys

50 55 60

Gln Val Pro Tyr Lys Val Ile His Glu Lys Glu His Ile Val Pro Ile

65 70 75 80

His His Ala Tyr Ser Ala Val Met Lys Glu Lys Gln Pro Phe Tyr Val

85 90 95

Glu Lys Lys Val Pro Tyr Pro Val Glu Val Pro Val Lys Ile Pro Glu

100 105 110

Pro Tyr Lys Val Glu Val Pro Val Pro Val Pro Arg Pro Tyr Val Val

115 120 125

Lys Glu Ile Glu Glu Arg Glu Val Lys Val Pro Val Lys Val Glu Lys

130 135 140

Ala Val Pro Tyr Pro Lys Pro Glu Pro Tyr Arg Val Tyr Leu Pro Lys

145 150 155 160

Gly Tyr Ser Asp Lys Leu Asp Thr Arg Leu Leu Asp Glu His Thr Ser

165 170 175

Tyr Pro Phe Lys Gly His Glu Gly Ile Glu Ala Ser His Val Asp Leu

180 185 190

His Glu Lys Ile Gly Glu Leu Asp Ile His Lys Lys Lys Glu Lys Arg

195 200 205

<210> 50

<211> 266

<212> PRT

<213> Tenodera australasiae

<400> 50

Met Ala Ile Leu Leu Asn Leu Ala Val Phe Cys Ile Phe Gly Ala Ile

1 5 10 15

Gln Val Ser Ser Glu Gln Pro Pro Lys Thr Thr Glu Leu Thr Val Val

20 25 30

Thr Arg Glu Leu Lys Cys Cys Asp Gly Ile Gln Glu Leu Phe Glu Tyr

35 40 45

Ile Lys Gly Arg His Thr Leu Asn Asp Ser Ile Glu Ala Lys Val Gly

50 55 60

Ser Ile Pro Ala Leu Asp Thr Glu Ile Arg Arg Leu Gln Arg Leu Gly

65 70 75 80

Cys Gly Lys Glu Leu Glu Glu Lys Lys Asn His Glu Glu Ala Ser Glu

85 90 95

Ala Val Lys Lys Ile Glu Glu Ile Leu Ile Lys Leu Leu Glu Val Ser

100 105 110

Asp Arg Leu Leu Thr Asp Thr Leu Lys Leu Leu Ser Leu Leu Gly Phe

115 120 125

Ala Ser Asn Asp Glu Arg Cys Ile Cys Lys Asn Tyr Ala Cys Val Gln

130 135 140

Lys Leu Glu Tyr Pro Glu Lys Lys Val Lys Glu Val Lys Glu Cys Lys

145 150 155 160

Glu Gly Lys Cys Asn Lys Arg Leu Arg Glu Leu Leu Asn Ser Leu Tyr

165 170 175

Arg Met Gln Ala Asp Leu Glu Asp Cys Val Gln Leu Tyr Glu Leu Ala

180 185 190

Glu Gln Asn Leu Gln Gln Val Ser Lys Leu Arg Glu Glu Val Gln Lys

195 200 205

Asp Thr Glu Cys Lys Asp Glu Lys Leu Ala Asn Asp Tyr Lys Glu Lys

210 215 220

Leu Ser Leu Ser Ser Leu Tyr Leu Ser Lys Ala Asn Tyr Leu Asp Tyr

225 230 235 240

Lys Val Val Tyr Glu Arg Ile Pro Tyr Leu Leu Lys Glu Cys Gln Glu

245 250 255

Thr Ile Lys Glu Ile Asn Ser Cys Gly Lys

260 265

<210> 51

<211> 201

<212> PRT

<213> Tenodera australasiae

<400> 51

Met Lys Lys Tyr Ser Val Gln Leu Ala Ala Leu Val Ile Leu Cys Ala
1 5 10 15

Ile Ser Phe Cys Ser Ala Gly Glu Glu Lys Lys Glu Glu Lys Lys Glu
20 25 30

Glu Lys Lys Glu Lys Lys Cys Pro Pro Glu Gly Val Glu Leu Lys Pro
35 40 45

Val Glu Leu Ser Glu Glu Asp Ile Lys His Leu Val His His Trp His
50 55 60

Leu Val Lys His Asn Asn Asn Leu Tyr Gly Lys Asp Leu Ser Cys Ser
65 70 75 80

Glu Tyr Thr Tyr Gly Lys Ser Glu Lys Gly Val Phe Phe His Gly Ser
85 90 95

Ala Tyr Asn Glu Ser Ser Asn Lys Val Trp Ala Thr Thr Gly Val Phe

100 105 110

Thr Gln Lys Glu Arg Gly Val Tyr Glu Gly Val Phe Glu Asp Lys Arg

115 120 125

Val Asn Gly Glu Phe Leu Phe Leu Tyr Val Ser Lys Lys Ile Tyr Ile

130 135 140

Leu Ala Gly Cys Ile Asn Asp Lys Glu Phe Ala Leu Trp Val Lys Thr

145 150 155 160

Val Asp Gly Lys Leu Asp Lys Glu Gln Glu Glu Ile Val Glu Ala Thr

165 170 175

Leu Lys Lys Asn Gly Leu Pro Asn Asp Leu Lys Gln Ala Pro Asn Asp

180 185 190

Asn Cys Ala Lys Lys Lys Ser Ser Val

195 200

<210> 52

<211> 384

<212> PRT

<213> Tenodera australasiae

<400> 52

Met Asp Ile Arg Phe Leu Leu Leu Phe Ile Leu Pro Ala Val Ala Thr
1 5 10 15

Ala Gly Thr Tyr Ser Asp Glu Thr Val Ser Leu Gly Met Asp Asn Glu
20 25 30

Asp Lys Glu Phe Asp Asn Ser Gly Ser Tyr Lys Thr Glu Gly Glu Tyr
35 40 45

Ser Val Ala Gly Ser Val Ser Ala Asp Ile Phe Asn Glu Gly Ser His
50 55 60

Leu Ala Gly Ala Ala Ala Glu Glu Trp Gly Met Lys Ala Lys Thr Val
65 70 75 80

Glu Gln Asn Ala Lys Trp Phe Arg Asn Leu Leu Asn Glu Lys Gln Met

85 90 95

Ile Leu Gly Gln Val Lys Lys Arg Leu Gly Leu Leu His Ala Leu Ala

100 105 110

Thr Glu Asn Leu Gln Asn Ala Glu Gln Tyr Arg Leu Ser Ala Ala Glu

115 120 125

Lys Cys Gln Ser Ser Lys Ala Leu Glu Thr Lys Ala Arg Glu Ser His

130 135 140

Asn Leu Asn Lys Asp Tyr Ala Val Gln Val His Glu Leu Arg Ala Lys

145 150 155 160

Leu Asn Asn Glu Leu Lys Gln His Leu Asp Ala Leu Ile Ser Ser Arg

165 170 175

Val Ala Glu Ala Asn Ala Asn Thr Asp Ile Ala Glu Ala Glu Ile Gln

180 185 190

Arg Ser Trp Ala Glu Phe Leu Glu Lys Gln Ala Gln Asp Met Leu Ser

195 200 205

Glu Ala Ala Lys Trp Glu Arg Ser Ser Thr Glu Asn Leu Asn Ser Ala

210 215 220

Lys Lys Phe Lys Leu Glu Ala Ala Ala Glu Lys Ala Lys Ala Lys Glu

225 230 235 240

Leu Glu Lys Lys Arg Lys Glu Ala Leu Asp Ala Ala Met Glu Ala Glu

245 250 255

Arg Lys Ala Leu Asp Phe Tyr Gln Phe His Leu Glu Phe His Met Lys

260 265 270

Ala Gln Ala Glu Glu His Gln Gly Gly Gln Ile Lys Val Lys Ala Lys

275 280 285

Tyr Leu Gln Met Met Ala Ala Ser Leu Leu Asp Ser Ala Ser Gln Leu

290 295 300

Arg Met Phe Ala Asn Leu Gln Asn Lys Asp Lys Ser Leu Trp Glu Lys

305 310 315 320

Lys Thr Ala Ser Ala Lys Ala Glu Ala Asn Leu Leu Glu Ala Ser Ala

325 330 335

Asp Lys Met Leu Gln Asn Ile Glu Lys Leu Lys Lys Ser Leu Gln Ala

340 345 350

Ser Ser Lys Val Ala Gln Lys Ser Asn Glu Leu Val Pro Gln Glu Glu

355 360 365

Lys Ser Val Lys Ser Glu Pro Arg Ala Ile Thr Pro Asp Lys Lys Val

370 375 380

<210> 53

<211> 269

<212> PRT

<213> Tenodera australasiae

<400> 53

Met Ala Ile Pro Ile Lys Leu Ala Leu Leu Phe Ser Ile Leu Gly Val
1 5 10 15

Ala Gln Val Cys Ser Ser Glu Tyr Lys Thr Asp Asp Leu Asp Ile Arg
20 25 30

Thr Tyr Glu Val Lys Cys His Gln Lys Val Glu Met Leu Val Gln Asn
35 40 45

Val Gln Asn Ser Ser Ser Ile Tyr Ala Arg Ala Ser Glu Ser Lys Glu
50 55 60

His Leu Ile Lys Ile Gly Thr His Leu Ala Glu Leu Leu Val Lys Tyr
65 70 75 80

Ala Cys Glu Lys Glu Ile Thr Asp Lys Lys Gly Ile Glu Glu Ile Ile
85 90 95

Ala Thr Leu Gln Glu Phe Arg Ser Ser Leu Glu Asn Tyr Ile Gln Glu

100 105 110

Val Ala Val Lys Tyr Phe Gly Ala Leu Tyr Ser Thr Thr Leu Tyr Ser

115 120 125

Thr Trp Ala Leu Glu Glu Ile His Cys Cys Cys Ser Lys Leu Asn Ile

130 135 140

Cys Leu Lys Lys Tyr Glu Tyr Pro Ile Ser Lys Asn Glu Val Lys Glu

145 150 155 160

Cys Lys Pro Asn Lys Cys Asn Lys Ala Val Val Gly Gly Gly Tyr Lys

165 170 175

Glu Leu Ile Thr Trp Leu Ser Lys Val Glu Glu Ser Glu Trp Leu Leu

180 185 190

Lys Glu Leu Ile Arg Leu Ile Glu Glu Leu Leu Lys Tyr Lys Asp Glu

195 200 205

Tyr Ile Lys Asp Pro Lys Cys Ala Asn Glu Lys Ala Ala Ala Leu Ile

210 215 220

Gly Glu Leu Asp Gln Arg Ile Val Leu Ile Val Cys Asp Leu Asn Tyr

225 230 235 240

Arg Gln His Glu Leu Val Lys Tyr Asp Leu Leu Met Leu His Arg Ala

245 250 255

Asp Glu Ile Leu Thr Gly Phe Lys Glu Cys Lys Lys Glu

260 265

<210> 54

<211> 64

<212> PRT

<213> Tenodera australasiae

<400> 54

Asp Asn Phe Ser Gln Ala Gln Val Phe Trp Asn Asn Val Leu Lys Pro

1 5 10 15

Asp Glu Lys Glu Arg Phe Val Glu Asn Val Tyr Glu Thr Leu Lys Leu

20 25 30

Val Glu Pro Glu Leu Gln Leu Asp Val Thr Lys Tyr Phe Gly Ala Val

35 40 45

His Pro Asp Ile Arg Glu Met Leu Ala Ala Lys Leu Glu Leu Lys Thr

50 55 60