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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 March 2018 (22.03.2018)

W!P O | PCT

(10) International Publication Number
WO 2018/053204 Al

(51) International Patent Classification:
C07K 14/435 (2006.01) *D01F 4/00* (2006.01)
D01F 5/00 (2006.01)

(21) International Application Number:
PCT/US20 17/05 1668

(22) International Filing Date:
14 September 2017 (14.09.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/394,683 14 September 2016 (14.09.2016) US

(71) Applicant: **BOLT THREADS, INC.** [US/US]; 5858 Horton Street, Suite 400, Emeryville, CA 94608 (US).

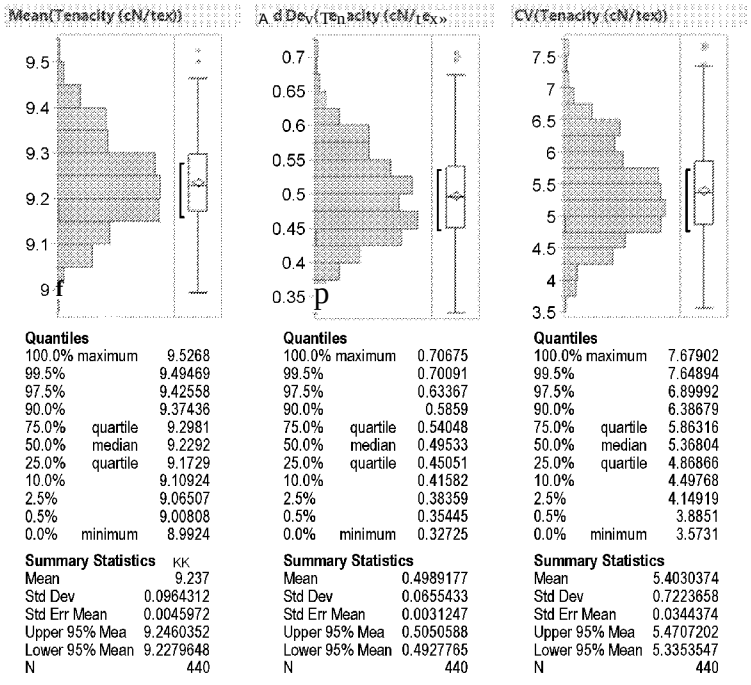
(72) Inventors: **WRAY, Lindsay;** c/o Bolt Threads, Inc., 5858 Horton Street, Suite 400, Emeryville, CA 94608 (US). **PERELMAN, Loren;** c/o Bolt Threads, Inc., 5858 Horton Street, Suite 400, Emeryville, CA 94608 (US). **KITTLESON, Joshua;** c/o Bolt Threads, Inc., 5858 Horton Street, Suite 400, Emeryville, CA 94608 (US). **BRESLAUER, David;** c/o Bolt Threads, Inc., 5858 Horton Street, Suite 400, Emeryville, CA 94608 (US).

(74) Agent: **HURT, Nicholas, S.** et al; Fenwick & West LLP, Silicon Valley Center, 801 California Street, Mountain View, CA 94041 (US).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, 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,

(54) Title: LONG UNIFORM RECOMBINANT PROTEIN FIBERS



Tenacity Data

95% of data is between:
Mean: 9.07 - 9.42
S.D.: 0.38 - 0.63
CV: 4.15 - 6.90

FIG. 2

(57) Abstract: The present disclosure provides improved long uniform recombinant protein fibers with desirable physical traits. The present disclosure also provides compositions derived from the long uniform recombinant protein fibers.

[Continued on next page]

WO 2018/053204 A1

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, KM, ML, MR, NE, SN, TD, TG).

Published:

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

LONG UNIFORM RECOMBINANT PROTEIN FIBERS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 62/394,683, filed September 14, 2016, the disclosure of which is incorporated herein by reference.

TECHNICAL FIELD

[0002] The present disclosure relates generally to recombinant protein fibers. Specifically, the present disclosure relates to increased uniformity of physical, mechanical and chemical properties of recombinant protein fibers.

BACKGROUND

[0003] Recombinant protein fibers, such as those synthesized from the polypeptides in spider silks, are not commercially available due to the difficulty in commercial scale fabrication and the technical challenges in producing fibers that are manufacturable into threads, yarns, and textiles.

[0004] There are many types of recombinant protein fibers that could be produced, with various useful properties.

[0005] One example is a recombinant protein fiber made from proteins designed by modifying spider silk proteins and protein fragments. Spider silk cannot be commercially farmed and harvested using the same methods that are applied to silkworm silk. This is due, in part, to the aggressive and territorial nature of spiders. Therefore, synthetically produced spider silk is the most likely cost-effective and viable path to commercialization.

[0006] A single species of spider creates a variety of fibers, each of which are utilized for different functions. Examples of these different functions include draglines, web capture spirals, prey immobilization, and silks to protect an egg sac. Dragline silks have exceptional mechanical properties. They are very strong for their weight and diameters, and also exhibit a combination of high extensibility in conjunction with high ultimate tensile strength.

[0007] Amino acid composition and protein structure vary considerably between types of silks and species of spiders. For example, orb weaving spiders have six unique types of glands that produce different silk polypeptide sequences that are polymerized into fibers tailored to fit an environmental or lifecycle niche. The fibers are named for the gland they originate from and the polypeptides are labeled with the gland abbreviation, for example "Sp" for spidroin (short for spider fibroin). In orb weaver spiders, examples include Major Ampullate (MaSp, also called dragline), Minor Ampullate (MiSp), Flagelliform (Flag), Aciniform (AcSp), Tubuliform (TuSp), and Pyriform (PySp).

[0008] There is a common class of orb weaver MaSp dragline silks (e.g., Nephila clavipes MaSpl) where the repeat domains contain glycine-rich regions, which are associated with amorphous regions of the fiber (possibly containing alpha-helices and/or beta-turns), and poly-alanine regions, which are associated with the beta-sheet crystalline regions of the fiber. The amino acid composition and sequence, as well as the fiber formation details both affect the mechanical properties of the fiber.

[0009] Currently, recombinant silk fibers are not commercially available and, with a handful of exceptions, are not produced in microorganisms outside of *Escherichia coli* and other gram-negative prokaryotes. Recombinant silks produced to date have largely consisted either of polymerized short silk sequence motifs or fragments of native repeat domains, sometimes in combination with NTDs and/or CTDs. While these methods are able to produce small scales of

recombinant silk polypeptides (milligrams at lab scale, kilograms at bioprocessing scale) using intracellular expression and purification by chromatography or bulk precipitation, they have not been scaled to volumes necessary for commercial manufacturing. Additional production hosts that have been utilized to make silk polypeptides include transgenic goats, transgenic silkworms, and plants. Similarly, these hosts have yet to enable commercial scale production of silk.

[0010] There are disclosures of continuous spinning methods for recombinant protein fibers. Several references generally disclose systems for continuous spinning of recombinant protein fibers, however none actually discloses working examples of fiber that are produced by continuous methods. *See* U.S. Pat. Nos. 7,868,146, 7,335,739, 8,979,992, 9,023,142, 9,051,453, PCT/JP2013/062429, and U.S. Pat. Pub. Nos. 2003/0201560, 2007/0256250, 2005/0054830, incorporated by reference herein in their entirety. All working examples (such as in PCT/JP2013/062429, and in U.S. Pat. Pub. No. 20050054830) are produced using spin dope dispensed from a syringe, not a larger vessel capable of dispensing the volumes needed for long continuous fibers. As a result, the fibers suffer from poor uniformity, poor reproducibility, or both.

[0011] The syringe-based approaches at the lab scale produce fibers with highly variable mechanical properties. For instance, collaboration between University of Wyoming, Arizona State University, Sandia National Laboratories and Utah State University published work where 4 different spider silk derived proteins were produced at small scales using *E. coli.* An *etal*, Biomacromolecules 2012, 13, 3938-3948. The cell suspension volumes used for purification were approximately 800 mL, and the spinning apparatus utilized 1 mL syringes from which to spin the fibers. This approach resulted in fibers that were 2-3 m long. These fibers were then examined by eye to exclude visible large defects, and sections 2 cm long were selected for analysis. The mechanical properties of the as-spun fibers produced by these small-scale methods

had average coefficient of variation (CV) of 40% for strength, 36% for extension, and 59% for toughness. The mechanical properties of the drawn (i.e. stretched) fibers produced by these small-scale methods had average coefficient of variation (CV) of 35% for strength, 88% for extension, and 97% for toughness. The average CV, in this instance, refers to the average CV of the 4 different proteins that were used in the spin dope. Furthermore, the average strength of these fibers was insufficient for commercial yarn and fabric production.

[0012] Another study, which produced spider silk derived proteins in small volumes using mammalian cells, also produced fibers with highly variable mechanical properties. Lazaris *etal*, Science 295, Jan 18, 2002, 472-476. Seven fibers from the same production methods were tested and had an average toughness of 0.895 gpd, and a CV of 61%.

[0013] There are a variety of test methods that have been developed for fiber, yarns and fabrics. The American Association of Textile Chemists and Colorists (AATCC) has developed a series of tests for fibers and textiles. The standard AATCC tests are known to persons of ordinary skill in the textile arts and can be found at in the 2016 AATCC Technical Manual (ISBN 978-1-942323-01-3) and are incorporated by reference in their entirety.

[0014] In order to manufacture goods comprising recombinant protein fibers, methods are required to produce large quantities of uniform fibers at low cost. What is needed, therefore, are large-scale methods to produce recombinant protein fibers with uniform properties, wherein those properties are adequate for commercial yarn spinning and textile production.

[0015] Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way. Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

[0016] The reagents employed in the examples are generally commercially available or can be prepared using commercially available instrumentation, methods, or reagents known in the art. The foregoing examples illustrate various aspects described herein and practice of the methods described herein. The examples are not intended to provide an exhaustive description of the many different embodiments of the invention. Thus, although the forgoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, those of ordinary skill in the art will realize readily that many changes and modifications can be made thereto without departing from the spirit or scope of the appended claims.

SUMMARY

[0017] In some embodiments, provided herein is a long uniform recombinant protein fiber, comprising a continuous fiber length of at least 600 m, wherein the mean properties of the fiber comprise: a tenacity greater than or equal to 12 cN/tex; a linear density less than or equal to 6 dtex; a coefficient of variation of tenacity less than 15% along the length; and a coefficient of variation of linear density less than 20% along the length, wherein the tenacity is measured using ASTM D3822- 14, and the linear density is measured using ASTM D 1577.

[0018] In some embodiments, the length of the recombinant protein fiber is at least 50 m. In some embodiments, the length of the recombinant protein fiber is at least 650 m.

[0019] In some embodiments, the tenacity of the recombinant protein fiber has a coefficient of variation less than 10% along the length of the recombinant protein fiber. In some embodiments, the linear density of the recombinant protein fiber has a coefficient of variation less than 15% along the length of the recombinant protein fiber.

[0020] In some embodiments, the mean elongation at break of the recombinant protein fiber is greater than 25% and the elongation at break of the recombinant protein fiber has a coefficient of variation of less than 35% along the length of the recombinant protein fiber.

[0021] In some embodiments, the mean initial modulus of the recombinant protein fiber is greater than 480 cN/tex and the initial modulus of the recombinant protein fiber has a coefficient of variation of less than 5% along the length.

[0022] In some embodiments, the mean elongation of the recombinant protein fiber is greater than 24% and the elongation of the recombinant protein fiber has a coefficient of variation of less than 45% along the length of the recombinant protein fiber.

[0023] In some embodiments, the mean work of rupture of the recombinant protein fiber is greater than 3 cN * cm and the work of rupture of the recombinant protein fiber has a coefficient of variation of less than 50% along the length of the recombinant protein fiber.

[0024] In some embodiments, the mean force at rupture of the recombinant protein fiber is greater than 7 cN and the force at rupture of the recombinant protein fiber has a coefficient of variation less than 25% along the length of the recombinant protein fiber.

[0025] In some embodiments, the recombinant protein fiber is produced by wet spinning a dope comprising a recombinant protein powder. In some embodiments, the recombinant protein powder is less than 65% proteinaceous block copolymer by mass.

[0026] In some embodiments, the recombinant protein fiber comprises a protein sequence comprising repeat units, wherein each repeat unit has at least 95% sequence identity to a sequence that comprises from 2 to 20 quasi-repeat units, each quasi-repeat unit having a composition comprising {GGY-[GPG-Xl]nl-GPS-(A)n2}, wherein for each quasi-repeat unit: XI is independently selected from the group consisting of SGGQQ, GAGQQ, GQGPY, AGQQ, and SQ; and nl is from 4 to 8, and n2 is from 6 to 10. In some embodiments, nl is from 4 to 5

for at least half of the quasi-repeat units. In some embodiments, n2 is from 5 to 8 for at least half of the quasi-repeat units.

[0027] In some embodiments, each quasi-repeat unit has at least 95% sequence identity to a MaSp2 dragline silk protein subsequence. In some embodiments, the repeat unit comprises SEQ ID NO: 1.

[0028] In some embodiments, the recombinant protein sequence comprises alanine-rich regions and glycine-rich regions, wherein: the alanine-rich regions form a plurality of nanocrystalline beta-sheets; and the glycine-rich regions form a plurality of beta-turn structures.

[0029] In some embodiments, the linear density and the tenacity of the recombinant protein fiber are measured using FAVIMAT fiber tensile test equipment model Favimat+ and Robot2.

[0030] In some embodiments, the recombinant fiber is not a MaSp2 dragline silk protein.

[0031] In some embodiments, also provided herein is a yarn comprising the recombinant protein fiber provided herein, wherein the yarn is a filament yarn. In some embodiments, the yarn is a spun yarn. In some embodiments, the yarn is a blended yarn.

[0032] In some embodiments, also provided herein is a textile comprising the yarn comprising the recombinant protein fiber provided herein, wherein the textile is a knitted textile. In some embodiments, the textile is a circular-knitted textile, a flat-knitted textile, or a warp-knitted textiles.

[0033] In some embodiments, also provided herein is a textile comprising the yarn comprising the recombinant protein fiber provided herein, wherein the textile is a woven textile. In some embodiments, the textile is a plain weave textile, a dobby weave textile, or a jacquard weave textile.

[0034] In some embodiments, also provided herein is a textile comprising the yarn comprising the recombinant protein fiber provided herein, wherein the textile is a non-woven textile. In

some embodiments, the textile is a needle punched textile, a spunlace textile, a wet-laid textile, a dry-laid textile, a melt-blown textile, or a 3-D printed non-woven textile..

BRIEF DESCRIPTION OF THE DRAWINGS

[0035] FIG. 1 schematically illustrates a molecular structure of a block copolymer of the present disclosure, in an embodiment.

[0036] FIG. 2 shows maximum tensile strength measured from fibers of the present disclosure, in embodiments.

[0037] FIG. 3 shows linear density measured from fibers of the present disclosure, in embodiments.

[0038] FIG. 4 shows stress-strain curves measured from fibers of the present disclosure, in embodiments.

[0039] The figures depict various embodiments of the present disclosure for purposes of illustration only. One skilled in the art will readily recognize from the following discussion that alternative embodiments of the structures and methods illustrated herein may be employed without departing from the principles described herein.

DEFINITIONS

[0040] Recombinant protein fibers (RPFs) are fibers that are produced from recombinant proteins. In some cases, the proteins making up the RPFs can contain concatenated repeat units and quasi-repeat units. Repeat units are defined as amino acid sequences that are repeated exactly within the polypeptide. Quasi-repeats are inexact repeats, i.e., there is some sequence

variation from quasi-repeat to quasi-repeat. Each repeat can be made up of concatenated quasi-repeats.

[0041] Amino acids can be referred to by their single-letter codes or by their three-letter codes. The single-letter codes, amino acid names, and three-letter codes are as follows: G - Glycine (Gly), P - Proline (Pro), A - Alanine (Ala), V - Valine (Val), L - Leucine (Leu), I - Isoleucine (He), M - Methionine (Met), C - Cysteine (Cys), F - Phenylalanine (Phe), Y - Tyrosine (Tyr), W - Tryptophan (Trp), H - Histidine (His), K - Lysine (Lys), R - Arginine (Arg), Q - Glutamine (Gin), N - Asparagine (Asn), E - Glutamic Acid (Glu), D - Aspartic Acid (Asp), S - Serine (Ser), T - Threonine (Thr).

[0042] Filament yarns are yarns that are composed of more than one fiber filaments that run the whole length of the yarn. Filament yarns can also be referred to as multi-filament yarns. The structure of a filament yarn is influenced by the amount of twist, and in some cases the fiber texturing. The properties of the filament yarn can be influenced by the structure of the yarn, fiber to fiber friction of the constituent fibers, and the properties of the constituent fibers. In some embodiments, the yarn structure and the RPF properties are chosen to impart various characteristics to the resulting yarns. The properties of the yarn can also be influenced by the number of fibers (i.e., filaments) in the yarn. The filament yarns in this application can be multifilament yarns. Throughout this disclosure "filament yarns" can refer to flat filament yarns, textured filament yarns, drawn filament yarns, undrawn filament yarns, or filament yarns of any structure.

[0043] Spun yarn is made by twisting staple fibers together to make a cohesive yarn (or thread, or "single"). The structure of a spun yarn is influenced by the spinning methods parameters. The properties of the spun yarn are influenced by parameters such as the structure of the yarn, fiber to fiber friction, and the properties of the constituent fibers.

[0044] Blended yarns are a type of yarn comprising various fibers being blended together. In different embodiments, the RPFs can be blended with cotton, wool, other animal fibers, polyamide, acrylic, nylon, linen, polyester, and/or combinations thereof. RPFs can be blended with non-recombinant protein fibers, or with more than one other type of non-recombinant protein fibers. RPFs can also be blended with a second type of RPF with different properties than the first type of RPFs. In this disclosure, blended yarns specifically refer to RPFs blended with non-RPFs or a second type of RPFs into a yarn. Even though spandex is generally incorporated into a yarn using somewhat different methods and structures than the other blended yarns described above (e.g., a wrapped RPF/spandex yarn has spandex core wrapped with RPF in order to hide the spandex from view in the textile), a composite RPF/spandex yarn therefore is another example of a blended yarn.

[0045] "Textured" fibers or yarns are fibers or yarns that have been subjected to processes that arrange the straight filaments into crimped, coiled or looped filaments. Some examples of methods used for processing textured fibers and yarns are air jet texturing, false twist texturing, or stuffer box texturing.

[0046] The standard test method for measuring tensile properties of single fibers is ASTM D3822-14. The standard test method for measuring tensile properties of yarns (or multiple fibers in a tow) by the single-strand method is ASTM D2256-10. All fiber and yarn mechanical properties measured in this disclosure are measured using one of these standards.

[0047] Some of the mechanical properties of the fibers in this disclosure are reported in units of MPa (i.e. 10^6 N/m^2, or force per unit area), and some are reported in units of cN/tex (force per linear density). The measurements of fibers mechanical properties reported in MPa were obtained using a custom instrument, which includes a linear actuator and calibrated load cell, and the fiber diameter was measured by light microscopy. The measurements of fibers mechanical

properties reported in cN/tex were obtained using FAVIMAT testing equipment (specifically, the Favimat+ and Robot2 models), which includes a measurement of the fiber linear density using a vibration method (e.g., according to ASTM D1577). To accurately convert measurements from MPa to cN/tex, an estimate of the bulk density (e.g. in g/cm^3) of the fiber is used. An expression that can be used to convert a force per unit area in MPa, "FA", to a force per linear density in cN/tex, "FLD", using the bulk density in g/cm^3, "BD", is FLD = FA/(10*BD). Since the bulk density of recombinant silk can vary, a given value of fiber tenacity in MPa does not translate to a given value of fiber tenacity in cN/tex. However, if the bulk density of the recombinant silk is assumed to be from 1.1 to 1.4 g/cm^3, then mechanical property values can be converted from one set of units into the other within a certain range of error. For example, a maximum tensile stress of 100 MPa is equivalent to about 9.1 cN/tex if the mass density of the fiber is 1.1 g/cm^3, and a maximum tensile stress of 100 MPa is equivalent to about 7.1 cN/tex if the mass density of the fiber is 1.4 g/cm^3.

[0048] The "work of rupture" of a fiber or yarn is the work done from the point of the pretension load to the point of the breaking load. The energy required to bring a fiber or yarn to the breaking load can be obtained from the area under the load-elongation curve. The units of work of rupture can therefore be cN*cm. The "toughness" of a fiber or yarn is the energy per unit mass required to rupture the fiber or yarn. The toughness is the integral of the stress-strain curve, and can be calculated by dividing the work of rupture by the mass of the sample of fiber or yarn being tested. The units of toughness can therefore be cN/tex.

[0049] Throughout this disclosure, and in the claims, when percentages of amino acids are recited, that percentage indicates a mole fraction percentage (not a weight fraction percentage).

[0050] Throughout this disclosure, and in the claims, where method steps are recited, the order in which the steps are carried out can be varied from the order in which they are described, so long as an operable method results.

[0051] Throughout this disclosure, "along the length of the fiber" refers to samples taken along the length of the fiber at certain intervals. In some embodiments, "along the length of the fiber" can refer to a samples taken at an interval of, e.g., 1 per meter, 1 per 2 meters, 1 per 5 meters, 1 per 20 meters, 1 per 50 meters, or 1 per 100 meters. If the fibers are sampled from a textile or garment, then "along the length of the fiber" can also refer to samples taken from different areas of a textile or garment at an interval of, e.g., 1 per 1 cm^2, or 1 per 2 cm^2, or 1 per 5 cm^2, or 1 per 10 cm^2, or 1 per 20 cm^2, or 1 per 50 cm^2, or 1 per 100 cm^2, or 1 per 200 cm^2, or 1 per 500 cm^2.

[0052] The coefficient of variation of a quantitative property of a population is known to those skilled in the art as the standard deviation of the property of the population divided by the mean of the property of the population. When discussing coefficient of variation, enough samples are taken from a fiber, yarn, or textile to sufficiently mitigate low sample bias towards an artificially low CV. In all embodiments described in this disclosure the total number of samples used to calculate the CV is greater than or equal to 20. In some embodiments, the total number of samples is 20, or 40, or 60, or 80, or 100, or more than 100.

[0053] When a range of values is recited in this disclosure, e.g., "from X to Y," the range includes the extremes of the range, i.e., the range includes X and Y.

DETAILED DESCRIPTION

RECOMBINANT PROTEIN FIBER ENGINEERING

[0054] Recombinant protein fibers (i.e., RPFs) can be engineered to have different mechanical, structural, chemical, and biological properties. Some methods to engineer long uniform RPFs

for different properties are protein sequence design (e.g., higher ratio of GPGto poly-alanine to improve elasticity, where glycine is between 25-50% of the polypeptide), and/or microorganism strain design and/or growth conditions and/or protein purification, and/or fiber spinning conditions (e.g., changing spinneret diameter to tune fiber diameter).

[0055] Embodiments of the present disclosure include long uniform RPFs. In some embodiments, a "long uniform RPF" has a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, and physical (e.g., linear density, diameter), mechanical (e.g., maximum tenacity, initial modulus, extensibility, toughness), chemical (e.g., moisture absorption, moisture regain) and/or biological (e.g., antimicrobial) properties that are uniform along the length of the fiber, wherein the physical, mechanical and/or chemical property has a CV along the length of the fiber less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%. In many embodiments, the long uniform RPFs are engineered to comprise various improved mechanical, structural, chemical and biological properties. In some embodiments, long uniform RPFs are used to create yarns, textiles and/or products. In embodiments, the yarn, textile and/or product structure and the long uniform RPF properties are

chosen to impart various characteristics to the resulting yarns, textiles and/or products fabricated from the long uniform RPFs.

[0056] In some embodiments, the hydrophilicity and/or moisture absorption of the long uniform RPFs can be engineered by changing the protein sequence. In some embodiments, the RPF hydrophilicity and/or moisture absorptivity is increased by increasing the ratio of substantially hydrophilic to substantially hydrophobic amino acids in the sequence, without disrupting fiber forming features such as poly-alanine stretches. Examples of relatively polar (relatively hydrophilic) amino acids in recombinant spider silk polypeptide sequences are glutamine, serine and tyrosine, while glycine and alanine are relatively hydrophobic. In some embodiments, a long uniform RPF comprising hydrophilic RPFs comprises greater than 25% glycine, or greater than 30% glycine, or greater than 35% glycine, or greater than 40% glycine, or greater than 45% glycine, or between 25% and 45% or between 25% and 40% or between 25% and 35% glycine, or between 35% and 45% glycine, or between 35% and 40% glycine, or between 40% and 45% glycine. In some embodiments, a long uniform RPF comprising hydrophilic RPFs comprises greater than 5% glutamine, or greater than 10% glutamine, or greater than 15% glutamine, or greater than 20% glutamine, or greater than 25% glutamine, or between 5% and 10% glutamine, or between 10% and 15% glutamine, or between 15% and 20% glutamine, or between 20% and 25% glutamine. In some embodiments, a long uniform RPF comprising highly moisture absorbing RPFs comprises greater than 25% glycine, or greater than 30% glycine, or greater than 35% glycine, or greater than 40% glycine, or greater than 45% glycine, or between 25% and 45% or between 25% and 40% or between 25% and 35% glycine, or between 35% and 45% glycine, or between 35% and 40% glycine, or between 40% and 45% glycine. In some embodiments, a long uniform RPF comprising highly moisture absorbing RPFs comprises greater than 5% glutamine, or greater than 10% glutamine, or greater than 15% glutamine, or

greater than 20% glutamine, or greater than 25% glutamine, or between 5% and 10% glutamine, or between 10% and 15% glutamine, or between 15% and 20% glutamine, or between 20% and 25% glutamine. In some embodiments, a highly moisture absorbing RPF, upon being submerged in water at a temperature of 21 °C +/- 1 °C, can have a median or mean diameter change greater than 10%, or greater than 15%, or greater than 20%, or greater than 25%, or greater than 30%, or greater than 35%, or greater than 40%, or greater than 45%, or greater than 50%, or greater than 60%, or greater than 70%, or greater than 80%, or greater than 90%, or from 10% to 20%, or from 20% to 30%, or from 30% to 40%, or from 40% to 50%, or from 50% to 60%, or from 60% to 70%, or from 70% to 80%, or from 80% and 90%, or from 90% to 100%, or from 20% to 35%, or from 15% to 40%, or from 15% to 35%.

[0057] In some embodiments, the wickability of textiles can be engineered by changing the spinning parameters of the fibers making up the textile. In some embodiments, the fiber cross-section shape can be changed by changing the residence time in the coagulation bath, or by changing the ratio of protein solvent to protein non-solvent in the coagulation bath. The long uniform RPFs of the present disclosure processed with residence times in coagulation baths at the longer end of the disclosed range (such as greater than 60 seconds) produce corrugated cross sections. That is, each fiber has a plurality of corrugations (or alternatively "grooves") disposed at an outer surface of a fiber. Each of these corrugations is parallel to a longitudinal axis of the corresponding fiber on which the corrugations are disposed. These corrugations can act as channels to assist in the wicking of liquids including water. Theses long uniform RPFs with tailored cross-sections can also be formed into filament yarns, or spun yarns, or blended yarns. Filament yarn, or spun yarn, or blended yarn containing long uniform RPFs with tailored cross-sections can be used to make textiles with tailored moisture transport properties, such as higher wicking rates.

[0058] In some embodiments, antimicrobial protein motifs are added to the protein sequence to impart antimicrobial properties to the resulting long uniform RPFs, as well as improve the antimicrobial properties of filament yarns, or spun yarns, or blended yarns, and fabrics comprising the long uniform RPFs. Some examples of antimicrobial protein sequence motifs are the human antimicrobial peptides human neutrophil defensin 2 (HNP-2), human neutrophil defensins 4 (HNP-4) and hepcidin. These antimicrobial amino acid sequences can be added to the spider silk-derived polypeptide sequence after every quasi-repeat unit, or every 2 quasi-repeat units, or every 3 quasi-repeat units, or every 4 quasi-repeat units, or every 5 quasi-repeat units, or every 6 quasi-repeat units, or every 7 quasi-repeat units, or every 8 quasi-repeat units, or every 9 quasi-repeat units, or every 10 quasi-repeat units, or every 12 quasi-repeat units, or every 14 quasi-repeat units, or every 16 quasi-repeat units, or every 18 quasi-repeat units, or every 20 quasi-repeat units, or every 30 quasi-repeat units, or every 40 quasi-repeat units, or every 50 quasi-repeat units, or every 60 quasi-repeat units, or every 70 quasi-repeat units, or every 80 quasi-repeat units, or every 90 quasi-repeat units, or every 100 quasi-repeat units. In some embodiments, a textile, comprising filament yarn, or spun yarn, or blended yarn, comprising long uniform RPFs with such antimicrobial amino acid sequences, is tested using AATCC test method 100-2012, and has an increase in colony forming units less than 100 times in 24 hours, or has an increase in colony forming units less than 500 times in 24 hours, or has an increase in colony forming units less than 1000 times in 24 hours, or has a change in colony forming units from a 100 times reduction to a 1000 times increase in 24 hours.

[0059] In some embodiments, the extensibility of the long uniform RPFs is increased by increasing the ratio of GPG to poly-alanine in the protein sequence. In some embodiments, a long uniform RPF with a high degree of extensibility (such as extensibility greater than 3%, or greater than 10%, or greater than 20%, or greater than 30%, or from 3 to 30%, or from 3 to

100%), comprises greater than 25% glycine, or greater than 30% glycine, or greater than 35% glycine, or greater than 40% glycine, or greater than 45% glycine, or from 25% and 45% glycine, or from 25% to 40% glycine, or from 25% to 35% glycine, or from 35% to 45% glycine, or from 35% to 40% glycine, or from 40% to 45% glycine.

[0060] In some embodiments, the maximum tensile strength of the long uniform RPFs is increased by increasing the monodispersity of the protein comprising the long uniform RPFs. In some embodiments, the monodispersity of the protein comprising the long uniform RPFs is improved by engineering the strain of the microorganism used to produce the recombinant protein to secrete the protein. In turn, improved monodispersity improves the maximum tensile strength of the long uniform RPFs. In some embodiments, the proteins of the spin dope (the synthesis of which is described in WO201 5042164 A2, especially at paragraphs 114-134, which are incorporated by reference herein) composed of any of the polypeptides of the present disclosure, that are used to produce the long uniform RPFs with a high tensile strength (such as greater than 10 cN/tex), are substantially monodisperse. In this disclosure, "substantially monodisperse" can be >50%, or >55%, or >60%, or >65%, or >70%, or >75%, or >80%, or >85%, or >90%, or >95%, or >99% of the protein in the spin dope (percentages here are mass percentages) having molecular weight >50%, or >55%, or >60%, or >65%, or >70%, or >75%, or >80%, or >85%, or >90%, or >95%, or >99% of the full-length molecular weight of the encoded protein. In this disclosure "substantially monodisperse" also encompasses spin dope mixtures in which from 50% to 100%, or from 60% to 100%, or from 70% to 100%, or from 80% to 100%, or from 90% to 100%, or from 50% to 99%, or from 60% to 99%, or from 70% to 99%, or from 80% to 99%, or from 90% to 99% of the protein in the spin dope (percentages here are mass percentages) having molecular weight from 50% to 100%, or from 60% to 100%, or from 70% to 100%, or from 80% to 100%, or from 90% to 100%, or from 50% to 99%, or from

60% to 99%, or from 70% to 99%, or from 80% to 99%, or from 90% to 99% of the full-length molecular weight of the encoded protein.

[0061] Work of rupture is a measure of toughness and combines elasticity and tenacity. Therefore, in some embodiments, the toughness of the long uniform RPFs is increased by combining protein sequence engineering and strain engineering to simultaneously increase the elasticity and the tenacity, as described in this disclosure. In some embodiments, long uniform RPFs with a high degree of toughness (such as greater than 100 cN/tex measured using ASTM D3822-14), comprise greater than 25% glycine, or greater than 30% glycine, or greater than 35% glycine, or greater than 40% glycine, or greater than 45% glycine, or from 25% to 45% or from 25% to 40% or from 25% to 35% glycine, or from 35% to 45% glycine, or from 35% to 40% glycine, or from 40% to 45% glycine. In some embodiments, the long uniform RPFs with a high work of rupture (such as greater than 0.5 cN*cm measured using ASTM D3822-14), comprises greater than 25% glycine, or greater than 30% glycine, or greater than 35% glycine, or greater than 40% glycine, or greater than 45% glycine, or between 25% and 45% or from 25% to 40% or from 25% to 35% glycine, or from 35% to 45% glycine, or from 35% to 40% glycine, or from 40% to 45% glycine. In some embodiments, the proteins of the spin dope (the synthesis of which is described in WO201 5042164 A2, especially at paragraphs 114-134, which are incorporated by reference herein), expressed from any of the polypeptides of the present disclosure, comprising the RPFs with a high degree of toughness (such as greater than 100 cN/tex measured using ASTM D3822-14) or a high work of rupture (such as greater than 0.5 cN*cm measured using ASTM D3822-14), are substantially monodisperse.

[0062] In some embodiments, the initial modulus of the long uniform RPFs is increased by engineering the proteins to have better intermolecular forces. In some embodiments, intermolecular forces are increased by adding protein blocks that provide hydrogen bonding and

cross-linking bonds between the molecules that comprise the fiber. One example of a protein motif that improves the intermolecular forces is by increasing the number of polyalanine segments for intermolecular crystallization. Another example of polypeptide engineering to increase intermolecular forces is through the addition of amino acids that are capable of covalently cross-linking such as the disulfide bridges of cysteine. A long uniform RPFs with tailored intermolecular forces can have high initial modulus. In some embodiments long uniform RPFs with engineered polypeptides described above can have a high initial modulus greater than 50 cN/tex, or greater than 115 cN/tex, or greater than 200 cN/tex, or greater than 400 cN/tex, or greater than 550 cN/tex, or greater than 600 cN/tex, or greater than 800 cN/tex, or greater than 1000 cN/tex, or greater than 2000 cN/tex, or greater than 3000 cN/tex, or greater than 4000 cN/tex, or greater than 5000 cN/tex, or from 200 to 900 cN/tex, or from 100 to 7000 cN/tex, or from 500 to 7000 cN/tex, or from 50 to 7000 cN/tex, or from 100 to 5000 cN/tex, or from 500 to 5000 cN/tex, or from 50 to 5000 cN/tex, or from 100 to 2000 cN/tex, or from 500 to 2000 cN/tex, or from 50 to 2000 cN/tex, or from 100 to 1000 cN/tex, or from 500 to 1000 cN/tex, or from 50 to 1000 cN/tex, or from 50 to 500 cN/tex, or from 100 to 1000 cN/tex, or from 500 to 1000 cN/tex, or from 100 to 700 cN/tex (measured using ASTM D3822-14).

[0063] In some embodiments, the initial modulus of the long uniform RPFs is increased by increasing the draw ratio of the fiber during spinning. In some embodiments, long uniform RPFs with a high initial modulus has a draw ratio of greater than 1.5X, or greater than 2X, or greater than 3X, or greater than 4X, or greater than 5X, or greater than 6X, or greater than 8X, or greater than 10X, or greater than 15X, or greater than 20X, or greater than 25X, or greater than 3OX, or from 1.5X to 3OX, or from 1.5X to 20X, or from 1.5X to 15X, or from 1.5X to 10X, or from 1.5X to 6X, or 1.5X to 4X, or from 2X to 3OX, or from 2X to 20X, or from 2X to 15X, or from 2X to 10X, or from 2X to 6X, or from 2X to 4X, or from 4X to 3OX, or from 4X to 20X, or from

4X to 15X, or from 4X to 10X, or from 4X to 6X, or from 6X to 30X, or from 6X to 20X, or from 6X to 15X, or from 6X to 10X, or from 10X to 30X, or from 10X to 20X, or from 10X to 15X.

[0064] In some embodiments the long uniform RPF cross-section shape is changed by changing the spinneret orifice shapes. In some embodiments, the long uniform RPF diameter or linear density is increased or decreased by increasing or decreasing the spinneret orifice diameter. The softness of a fiber is highly influenced by the diameter or linear density, and in some embodiments, the spinneret diameter can also be used to tune the softness of the long uniform RPFs by decreasing the fineness of the fibers. In some embodiments, the linear density of the long uniform RPFs can be tuned from less than 10 decitex (i.e., dtex), or less than 5 dtex, or less than 1 dtex, or from 1 to 20 dtex, or from 1 to 10 dtex by using a draw ratio during spinning of greater than 1.5X, or greater than 2X, or greater than 3X, or greater than 4X, or greater than 5X, or greater than 6X, or greater than 8X, or greater than 10X, or greater than 15X, or greater than 20X, or greater than 25X, or greater than 30X, or from 1.5X to 30X, or from 1.5X to 20X, or from 1.5X to 15X, or from 1.5X to 10X, or from 1.5X to 6X, or 1.5X to 4X, or from 2X to 30X, or from 2X to 20X, or from 2X to 15X, or from 2X to 10X, or from 2X to 6X, or from 2X to 4X, or from 4X to 30X, or from 4X to 20X, or from 4X to 15X, or from 4X to 10X, or from 4X to 6X, or from 6X to 30X, or from 6X to 20X, or from 6X to 15X, or from 6X to 10X, or from 10X to 30X, or from 10X to 20X, or from 10X to 15X. In some embodiments, a textile with good softness contains long uniform RPFs with fiber linear density less than 10 dtex, or less than 5 dtex, or less than 1 dtex, or from 1 to 20 dtex, or from 1 to 10 dtex. The drape of a fabric is highly influenced by the linear density or diameter of the fibers comprising the fabric, and in some embodiments, the spinneret diameter or the draw ratio can also be used to tune the drape of a fabric by increasing or decreasing the fineness of the long uniform RPFs comprising the fabric.

In some embodiments, a textile with desirable drape contains filament yarn, or spun yarn, or blended yarn comprising long uniform RPFs with fiber linear density less than 10 dtex, or less than 5 dtex, or less than 1 dtex, or from 1 to 20 dtex, or from 1 to 10 dtex.

[0065] In some embodiments, the long uniform RPF cross-section shape can be changed by changing the residence time in the coagulation bath, or by changing the ratio of protein solvent to protein non-solvent in the coagulation bath. The long uniform RPFs of the present disclosure processed with residence times in coagulation baths at the longer end of the disclosed range produce corrugated cross sections. That is, each long uniform RPFs has a plurality of corrugations (or alternatively "grooves") disposed at an outer surface of a fiber. Each of these corrugations is parallel to a longitudinal axis of the corresponding fiber on which the corrugations are disposed. The luster of a fiber is also highly influenced by the smoothness of the surface. A long uniform RPF with a smoother surface has a higher luster, and in some embodiments, the luster of the fiber can also be tuned by changing the coagulation bath residence time or chemistry. A filament yarn, or spun yarn, or blended yarn can also contain long uniform RPFs with tailored cross-sections to create a yarn with low or high luster.

RECOMBINANT PROTEIN FIBER PROTEIN DESIGN

[0066] Long uniform RPFs can be produced using the following proteins and methods.

[0067] Embodiments of the present disclosure include fibers synthesized from synthetic proteinaceous copolymers based on recombinant spider silk protein fragment sequences derived from MaSp2, such as from the species *Argiope bruennichi*. Each synthesized fiber contains protein molecules that include two to twenty repeat units, in which a molecular weight of each repeat unit is greater than about 20 kDal. Within each repeat unit of the copolymer are more than about 60 amino acid residues that are organized into a number of "quasi-repeat units." In

some embodiments, the repeat unit of a polypeptide described in this disclosure has at least 95% sequence identity to a MaSp2 dragline silk protein sequence.

[0068] Utilizing long polypeptides with fewer long exact repeat units has many advantages over utilizing polypeptides with a greater number of shorter exact repeat units to create a recombinant spider silk fiber. An important distinction is that a "long exact repeat" is defined as an amino acid sequence without shorter exact repeats concatenated within it. Long polypeptides with long exact repeats are more easily processed than long polypeptides with a greater number of short repeats because they suffer less from homologous recombination causing DNA fragmentation, they provide more control over the composition of amorphous versus crystalline domains, as well as the average size and size distribution of the nano-crystalline domains, and they do not suffer from unwanted crystallization during intermediate processing steps prior to fiber formation. Throughout this disclosure the term "repeat unit" refers to a subsequence that is exactly repeated within a larger sequence.

[0069] Throughout this disclosure, wherever a range of values is recited, that range includes every value falling within the range, as if written out explicitly, and further includes the values bounding the range. Thus, a range of "from X to Y" includes every value falling between X and Y, and includes X and Y.

[0070] The term percent "identity," in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that have a specified percentage of nucleotides or amino acid residues that are the same, when compared and aligned for maximum correspondence, as measured using one of the sequence comparison algorithms described below (e.g., BLASTP and BLASTN or other algorithms available to persons of skill) or by visual inspection. Depending on the application, the percent "identity" can exist over a region of the sequence being compared (i.e., subsequence), e.g., over a functional domain, or, alternatively,

exist over the full length of the two sequences to be compared. Within this disclosure, a "region" is considered to be 6 or more amino acids in a continuous stretch within a polypeptide.

[0071] For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

[0072] Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), or by visual inspection *(see* generally Ausubel et al., infra).

[0073] One example of an algorithm that is suitable for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul et al, J. Mol. Biol. 215:403-410 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. Such software also can be used to determine the mole percentage of any specified amino acid found within a polypeptide sequence or within a domain of such a sequence. As the person of ordinary skill will recognize such percentages also can be determined through inspection and manual calculation.

[0074] FIG. 1 schematically illustrates an example copolymer molecule of the present disclosure, in an embodiment. A block copolymer molecule of the present disclosure includes in

each repeat unit more than 60, or more than 100, or more than 150, or more than 200, or more than 250, or more than 300, or more than 350, or more than 400, or more than 450, or more than 500, or more than 600, or more than 700, or more than 800, or more than 900, or more than 1000 amino acid residues, or from 60 to 1000, or from 100 to 1000, or from 200 to 1000, or from 300 to 1000, or from 400 to 1000, or from 500 to 1000, or from 150 to 1000, or from 150 to 400, or from 150 to 500, or from 150 to 750, or from 200 to 400, or from 200 to 500, or from 200 to 750, or from 250 to 350, or from 250 to 400, or from 250 to 500, or from 250 to 750, or from 250 to 1000, or from 300 to 500, or from 300 to 750 amino acid residues. Each repeat unit of the polypeptide molecules of this disclosure can have a molecular weight from 20 kDal to 100 kDal, or greater than 20 kDal, or greater than 10 kDal, or greater than 5 kDal, or from 5 to 60 kDal, or from 5 to 40 kDal, or from 5 to 20 kDal, or from 5 to 100 kDal, or from 5 to 50 kDal, or from 10 to 20 kDal, or from 10 to 40 kDal, or from 10 to 60 kDal, or from 10 to 100 kDal, or from 10 to 50 kDal, or from 20 to 100 kDal, or from 20 to 80 kDal, or from 20 to 60 kDal, or from 20 to 40 kDal, or from 20 to 30 kDal. A copolymer molecule of the present disclosure can include in each repeat unit more than 300 amino acid residues. A copolymer molecule of the present disclosure can include in each repeat unit about 315 amino acid residues. These amino acid residues are organized within the molecule at several different levels. A copolymer molecule of the present disclosure includes from 2 to 20 occurrences of a repeat unit. After concatenating the repeat unit, the polypeptide molecules of this disclosure can be from 20 kDal to 2000 kDal, or greater than 20 kDal, or greater than 10 kDal, or greater than 5 kDal, or from 5 to 400 kDal, or from 5 to 300 kDal, or from 5 to 200 kDal, or from 5 to 100 kDal, or from 5 to 50 kDal, or from 5 to 500 kDal, or from 5 to 1000 kDal, or from 5 to 2000 kDal, or from 10 to 400 kDal, or from 10 to 300 kDal, or from 10 to 200 kDal, or from 10 to 100 kDal, or from 10 to 50 kDal, or from 10 to 500 kDal, or from 10 to 1000 kDal, or from 10 to 2000 kDal, or from 20 to 400 kDal, or

from 20 to 300 kDal, or from 20 to 200 kDal, or from 40 to 300 kDal, or from 40 to 500 kDal, or from 20 to 100 kDal, or from 20 to 50 kDal, or from 20 to 500 kDal, or from 20 to 1000 kDal, or from 20 to 2000 kDal. As shown in FIG. 1, each "repeat unit" of a copolymer fiber comprises from two to twenty "quasi-repeat" units (i.e., n3 is from 2 to 20). Quasi-repeats do not have to be exact repeats. Each repeat can be made up of concatenated quasi-repeats. Equation 1 shows the composition of a quasi-repeat unit according the present disclosure.

{GGY-**[GPG**-Xi]ni-**GPS**-(A)n2}n3. (Equation 1)

[0075] The variable compositional element Xi (termed a "motif) is according to any one of the following amino acid sequences shown in Equation 2 and Xi varies randomly within each quasi-repeat unit.

Xi = SGGQQ or GAGQQ or GQGPY or AGQQ or SQ (Equation 2)

[0076] Referring again to Equation 1, the compositional element of a quasi-repeat unit represented by "GGY-**[GPG**-Xi]ni-GPS" in Equation 1 is referred to a "first region." A quasi-repeat unit is formed, in part by repeating from 4 to 8 times the first region within the quasi-repeat unit. That is, the value of ni indicates the number of first region units that are repeated within a single quasi-repeat unit, the value of ni being any one of 4, 5, 6, 7 or 8. The compositional element represented by "(A)n2" is referred to a "second region" and is formed by repeating within each quasi-repeat unit the amino acid sequence "A" n_2 times. That is, the value of n_2 indicates the number of second region units that are repeated within a single quasi-repeat unit, the value of n_2 being any one of 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some embodiments, the repeat unit of a polypeptide of this disclosure has at least 95% sequence identity to a sequence containing quasi-repeats described by Equations 1 and 2. In some embodiments, the repeat unit of a polypeptide of this disclosure has at least 80%, or at least 90%,

or at least 95%, or at least 99% sequence identity to a sequence containing quasi-repeats described by Equations 1 and 2.

[0077] The first region described in Equation 1 is considered a glycine-rich region. A region can be glycine-rich if 6 or more consecutive amino acids within a sequence are more than 45% glycine. A region can be glycine-rich if 12 or more consecutive amino acids within a sequence are more than 45% glycine. A region can be glycine-rich if 18 or more consecutive amino acids within a sequence are more than 45% glycine. A region can be glycine-rich if 4 or more, or 6 or more, or 10 or more, or 12 or more, or 15 or more, or 20 or more, or 25 or more, or 30 or more, or 40 or more, or 50 or more, or 60 or more, or 70 or more, or 80 or more, or 100 or more, or 150 or more consecutive amino acids within a sequence are more than 30%, or more than 40%, or more than 45%, or more than 50%, or more than 55% glycine, or more than 60% glycine, or more than 70% glycine, or more than 80% glycine, or from 30% to 80%, or from 40% to 80%, or from 45% to 80%, or from 30% to 55%, or from 30% to 50%, or from 30% to 45%, or from 30% to 40%, or from 40% to 50%, or 40% to 55%, or 40% to 60% glycine. A region can be glycine-rich if from 5 to 150, or from 10 to 150, or from 12 to 150, or from 12 to 100, or from 12 to 80, or from 12 to 60, or from 20 to 60 consecutive amino acids within a sequence are more than 30%, or more than 40%, or more than 45%, or more than 50%, or more than 55% glycine, or more than 60% glycine, or more than 70% glycine, or more than 80% glycine, or from 30% to 80%, or from 40% to 80%, or from 45% to 80%, or from 30% to 55%, or from 30% to 50%, or from 30% to 45%, or from 30% to 40%, or from 40% to 50%, or 40% to 55%, or 40% to 60% glycine. In addition, a glycine-rich region can have less than 10%, or less than 20%, or less than 30%, or less than 40% alanine, or from about 0% to 10%, or from about 0% to 20%, or from about 0% to 30%, or from about 0% to 40%, or alanine. A region can be alanine-rich if 4 or more, or 6 or more, or 8 or more, or 10 or more consecutive amino acids within a sequence are

more than 70%, or more than 75%, or more than 80%, or more than 85%, or more than 90% alanine, or from 70% to about 100%, or from 75% to about 100%, or from 80% to about 100%, or from 85% to about 100%, or from 90% to about 100% alanine. A region can be alanine-rich if from 4 to 10, or from 4 to 12, or from 4 to 15, or from 6 to 10, or from 6 to 12, or from 6 to 15, or from 4 to 20, or from 6 to 20 consecutive amino acids within a sequence are more than 70%, or more than 75%, or more than 80%, or more than 85%, or more than 90% alanine, or from 70% to about 100%, or from 75% to about 100%, or from 80% to about 100%, or from 85% to about 100%, or from 90% to about 100% alanine. The repeats described in this disclosure can have 6, or more than 2, or more than 4 or more than 6, or more than 8, or more than 10, or more than 15, or more than 20, or from 2 to 25, or from 2 to 10, or from 4 to 10, or from 2 to 8, or from 4 to 8 alanine-rich regions. The repeats described in this disclosure can have 6, or more than 2, or more than 4 or more than 6, or more than 8, or more than 10, or more than 15, or more than 20, or from 2 to 25, or from 2 to 10, or from 4 to 10, or from 2 to 8, or from 4 to 8 glycine-rich regions.

[0078] In some embodiments, long uniform RPFs comprise proteins containing SEQ described by Equation 1 and Equation 2. In some embodiments, long uniform RPFs comprise proteins with repeat units, where each repeat unit has at least 95% sequence identity to a sequence that comprises from 2 to 20 quasi-repeat units, and each quasi-repeat unit has a composition of {GGY-[GPG-**Xi]ni** -GPS-(A)n2}, and for each quasi-repeat unit **X i** is independently selected from the group consisting of SGGQQ, GAGQQ, GQGPY, AGQQ, and SQ, and n1 is from 4 to 8, and n2 is from 6 to 10.

[0079] As further described below, one example of a copolymer molecule includes three "long" quasi-repeats followed by three "short" quasi-repeat units. A "long" quasi-repeat unit is comprised of quasi-repeat units that do not use the same **X i** constituent (as shown in Equation 2)

more than twice in a row, or more than two times in a repeat unit. Each "short" quasi-repeat unit includes any of the amino acid sequences identified in Equation 2, but regardless of the amino acid sequences used, the same sequences are in the same location within the molecule. Furthermore, in this example copolymer molecule, no more than 3 quasi-repeats out of 6 share the same **Xi.** "Short" quasi-repeat units are those in which nl=4 or 5 (as shown in Equation 1). Long quasi-repeat units are defined as those in which nl=6, 7 or 8 (as shown in Equation 1).

[0080] In some embodiments, the repeat unit of the copolymer is composed of X_{qr} quasi-repeat units, where X_{qr} is a number from 2 to 20, and the number of short quasi-repeat units is $\mathbf{X_{Sqr}}$ and the number of long quasi-repeat units is **Xiqr,** where

$$\mathbf{Xsqr + Xiqr = Xqr} \qquad \text{(Equation 3)}$$

and **Xsqr** is a number from 1 to (X_{qr}-1) and **Xiqr** is a number from 1 to (X_{qr}-1).

[0081] In another embodiment, nl is from 4 to 5 for at least half of the quasi-repeat units. In yet another embodiment, n2 is from 5 to 8 for at least half of the quasi-repeat units.

[0082] One feature of copolymer molecules of the present disclosure is the formation of nano-crystalline regions that, while not wishing to be bound by theory, are believed to form from the stacking of beta-sheet regions, and amorphous regions composed of alpha-helix structures, beta-turn structures, or both. Poly-alanine regions (or in some species $(GA)_n$ regions) in a molecule form crystalline beta-sheets within major ampullate (MA) fibers. Other regions within a repeat unit of major ampullate and flagelliform spider silks (for example containing GPGGX, GPGQQ, GGX where X = A, S or Y, GPG, SGGQQ, GAGQQ, GQGPY, AGQQ, and SQ, can form amorphous rubber-like structures that include alpha-helices and beta-turn containing structures. Furthermore, secondary, tertiary and quaternary structure is imparted to the morphology of the fibers via amino acid sequence and length, as well as the conditions by which the fibers are formed, processed and post-processed. Materials characterization techniques (such as NMR,

FUR and x-ray diffraction) have suggested that the poly-alanine crystalline domains within natural MA spider silks and recombinant silk derived from MA spider silk sequences are typically very small (<10 nm). Fibers can be highly crystalline or highly amorphous, or a blend of both crystalline and amorphous regions, but fibers with optimal mechanical properties have been speculated to be composed of 10-40% crystalline material by volume. In some embodiments, the repeat unit of a polypeptide described in this disclosure has at least 80%, or at least 90%, or at least 95%, or at least 99% sequence identity to a MA dragline silk protein sequence. In some embodiments, the repeat unit of a polypeptide described in this disclosure has at least 80%, or at least 90%, or at least 95%, or at least 99% sequence identity to a MaSp2 dragline silk protein sequence. In some embodiments, the repeat unit of a polypeptide described in this disclosure has at least 80%, or at least 90%, or at least 95%, or at least 99% sequence identity to a spider dragline silk protein sequence. In some embodiments, a quasi-repeat unit of a polypeptide described in this disclosure has at least 80%, or at least 90%, or at least 95%, or at least 99% sequence identity to a MA dragline silk protein sequence. In some embodiments, a quasi-repeat unit of a polypeptide described in this disclosure has at least 80%, or at least 90%, or at least 95%, or at least 99% sequence identity to a MaSp2 dragline silk protein sequence. In some embodiments, a quasi-repeat unit of a polypeptide described in this disclosure has at least 80%, or at least 90%, or at least 95%, or at least 99% sequence identity to a spider dragline silk protein sequence.

[0083] While not wishing to be bound by theory, the structural properties of the proteins within the spider silk are theorized to be related to fiber mechanical properties. Crystalline regions in a fiber have been linked with the tensile strength of a fiber, while the amorphous regions have been linked to the extensibility of a fiber. The major ampullate (MA) silks tend to have higher strengths and less extensibility than the flagelliform silks, and likewise the MA silks have higher

volume fraction of crystalline regions compared with flagelliform silks. Furthermore, theoretical models based on the molecular dynamics of crystalline and amorphous regions of spider silk proteins, support the assertion that the crystalline regions have been linked with the tensile strength of a fiber, while the amorphous regions have been linked to the extensibility of a fiber. Additionally, the theoretical modeling supports the importance of the secondary, tertiary and quaternary structure on the mechanical properties of RPFs. For instance, both the assembly of nano-crystal domains in a random, parallel and serial spatial distributions, and the strength of the interaction forces between entangled chains within the amorphous regions, and between the amorphous regions and the nano-crystalline regions, influenced the theoretical mechanical properties of the resulting fibers.

[0084] The repeat unit of the proteinaceous block copolymer that forms fibers with good mechanical properties can be synthesized using a portion of a silk polypeptide. Some exemplary sequences that can be used as repeats in the proteinaceous block copolymers of this disclosure are shown in Table 1. These polypeptide repeat units contain alanine-rich regions and glycine-rich regions, and are 150 amino acids in length or longer. These exemplary sequences were demonstrated to express using a *Pichia* expression system as taught in co-owned PCT Publication WO 2015042164.

Table 1: Exemplary sequences that can be used as repeat units

Seq. ID No.	AA
1	GGYGPGAGQQGPGSGGQQGPGGQGPYGSGQQGPGGAGQQGPGGQGPYGPGAAAAAAAA AGGYGPGAGQQGPGGAGQQGPGSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGGQG PYGPSAAAAAAAAAGGYGPGAGQRSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGG QGPYGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPGAAAAAAAVG GYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPGSQGP GSGGQQGPGGQGPYGPSAAAAAAAA

Seq. ID No.	ΛΛ
2	GGQGGRGGFGGLGSQGAGGAGQGGAGAAAAAAAAGGDGGSGLGGYGAGRGHGVGLGGA GGAGAASAAAAAGGQGGRGGFGGLGSQGAGGAGQGGAGAAAAAAAAGGDGGSGLGGYG AGRGHGAGLGGAGGAGAASAAAAAGGQGGRGGFGGLGSQGSGGAGQGGSGAAAAAAAA GGDGGSGLGGYGAGRGYGAGLGGAGGAGAASAAAAAGGQGGRGGFGGLGSQGAGGAGQ GGSGAAAAAAAAVADGGSGLGGYGAGRGYGAGLGGAGGAGAASAAAAT
3	GSAPQGAGGPAPQGPSQQGPVSQGPYGPGAAAAAAAAGGYGPGAGQQGPGSQGPGSGG QQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQ GPGGQGPYGPGAAAAAAAVGGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAA AAAGGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPG SGGQQGPGGQGPYGSGQQGPGGAGQQGPGGQGPYGPGAAAAAAAAA
4	GGYGPGAGQQGPGSGGQQGPGGQGPYGSGQQGPGGAGQQGPGGQGPYGPGAAAAAAAA AGGYGPGAGQQGPGGAGQQGPGSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGGQG PYGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGY GPGAGQQGPGSGGQQGPGGQGPYGSGQQGPGGAGQQGPGGQGPYGGGYGPGAGQQGPG SQGPGSGGQQGPGGQGPYGPSAAAAAAAA
5	GPGARRQGPGSQGPGSGGQQGPGGQGPYGSGQQGPGGAGQQGPGGQGPYGPGAAAAAA AAAGGYGPGAGQQGPGGAGQQGPGSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGG QGPYGPSAAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPGAAAAAAAV GGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPGSQG PGSGGQQGPGGQGPYGPSAAAAAAAA
6	GPGARRQGPGSQGPGSGGQQGPGGQGPYGSGQQGPGGAGQQGPGGQGPYGPGAAAAAA AAAGGYGPGAGQQGPGGAGQQGPGSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGG QGPYGPSAAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPGAAAAAAAVG GYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPGSQGP GSGGQQGPGGQGPYGPSAAAAAAAA
7	GGYGPGAGQQGPGSGGQQGPGGQGPYGSGQQGPGGAGQQGPGGQGPYGPGAAAAAAAA AGGYGPGAGQQGPGGAGQQGPEGPGSQGPGSGGQQGPGGQGPYGPGAAAAAAAVGGYG PGAGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSG GQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPGSGGQQGPGGQGPYGSGQQGPGGA GQQGPGGQGPYGPGAAAAAAAAA
8	GVFSAGQGATPWENSQLAESFISRFLRFIGQSGAFSPNQLDDMSSIGDTLKTAIEKMA QSRKSSKSKLQALNMAFASSMAEIAVAEQGGLSLEAKTNAIASALSAAFLETTGYVNQ QFVNEIKTLIFMIAQASSNEISGSAAAAGGSSGGGGGSGQGGYGQGAYASASAAAAYG SAPQGTGGPASQGPSQQGPVSQPSYGPSATVAVTAVGGRPQGPSAPRQQGPSQQGPGQ QGPGGRGPYGPSAAAAAAAA
9	GAGAGAGAGAGAGAGAGSGASTSVSTSSSSGSGAGAGAGSGAGSGAGAGSGAGAGAGA GGAGAGFGSGLGLGYGVGLSSAQAQAQAQAAAQAQAQAQAQAYAAAQAQAQAQAQAQA AAAAAAAAAAGAGAGAGAGAGAGAGAGS GASTSVSTSSSSGSGAGAGAGSGAGSGAGA GSGAGAGAGAGGAGAGFGSGLGLGYGVGLSSAQAQAQAQAAAQAQAQAQAQAYAAAQA QAQAQAQAQAAAAAAAAAAA

Seq. ID No.	ΛΛ
10	GAGAGAGAGAGAGAGAGSGASTSVSTSSSSGSGAGAGAGSGAGSGAGAGSGAGAGAGA GGAGAAFGSGLGLGYGVGLSSAQAQAQAQAAAQAQADAQAQAYAAAQAQAQAQAQAQA AAAAAAAAAAGAGAGAGAGSGAGAGAGSGASTSVSTSSSSGSGAGAGAGSGAGSGAGA GSGAGAGAGAGGAGAGFGSGLGLGYGVGLSSAQAQAQAQAAAQAQADAQAQAYAAAQA QAQAQAQAQAAAAAAAAAAA
11	GAGAGAGAGSGAGAGAGSGASTSVSTSSSSGSGAGAGAGSGAGSGAGAGSGAGAGAGA GGAGAGFGSGLGLGYGVGLSSAQAQAQSAAAARAQADAQAQAYAAAQAQAQAQAQAQA AAAAAAAAAAGAGAGAGAGAGAGAGAGS GASTSVSTSSSSASGAGAGAGSGAGSGAGA GSGAGAGAGAGGAGAGFGSGLGLGYGVGLSSAQAQAQAQAAAQAQAQAQAQALAAAQA QAQAQAQAQAAAATAAAAAA
12	GGYGPGAGQQGPGGAGQQGPGSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGGQGP YGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQGPGSQGPGSGGQQGPGGQGPYGP SAAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPGAAAAAAAVGGYGPG AGQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQ QGPGGQGPYGPSAAAAAAAA
13	GGYGPGAGQQGPGGAGQQGPGSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGGQGP YGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQGPGSQGPGSGGQQGPGGQGPYGP SAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQGPGGQGPYGPGAAAAAAAVGGYGPGA GQQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQ GPGGQGPYGPSAAAAAAAA
14	GHQGPHRKTPWETPEMAENFMNNVRENLEASRIFPDELMKDMEAITNTMIAAVDGLEA QHRSSYASLQAMNTAFASSMAQLFATEQDYVDTEVIAGAIGKAYQQITGYENPHLASE VTRLIQLFREEDDLENEVEISFADTDNAIARAAAGAAAGSAAASSSADASATAEGASG DSGFLFSTGTFGRGGAGAGAGAAAASAAAASAAAAGAEGDRGLFFSTGDFGRGGAGAG AGAAAASAAAASAAAA
15	GGAQKHPSGEYSVATASAAATSVTSGGAPVGKPGVPAPIFYPQGPLQQGPAPGPSNVQ PGTSQQGPIGGVGESNTFSSSFASALGGNRGFSGVISSASATAVASAFQKGLAPYGTA FALSAASAAADAYNSIGSGASASAYAQAFARVLYPLLQQYGLSSSADASAFASAIASS FSTGVAGQGPSVPYVGQQQPSIMVSAASASAAASAAAVGGGPVVQGPYDGGQPQQPNI AASAAAAATATSS
16	GGQGGRGGFGGLGSQGEGGAGQGGAGAAAAAAAAGADGGFGLGGYGAGRGYGAGLGGA GGAGAASAAAAAGGQGGRSGFGGLGSQGAGGAGQGGAGAAAAAAAAGADGGSGLGGYG AGRGYGASLGGADGAGAASAAAAAGGQGGRGGFGGLGSQGAGGAGQGGAGAAAAAAAA SGDGGSGLGGYGAGRGYGAGLGGAGGAGAASAAAAAGGEGGRGGFGGLGSQGAGGAGQ GGSLAAAAAAAA
17	GPGGYGGPGQPGPGQGQYGPGPGQQGPRQGGQQGPASAAAAAAAGPGGYGGPGQQGPR QGQQQGPASAAAAAAAAAGPRGYGGPGQQGPVQGGQQGPASAAAAAAAAGVGGYGGP GQQGPGQGQYGPGTGQQGQGPSGQQGPAGAAAAAAGGAAGPGGYGGPGQQGPGQGQYG PGTGQQGQGPSGQQGPAGAAAAAAAAAGPGGYGGPGQQGPGQGQYGPGAGQQGQGPGS QQGPASAAAAAA

Seq. ID No.	ΛΛ
18	GSGAGQGTGAGAGAAAAAAGAAGS GAGQGAGS GAGAAAAAAAAS AAGAGQGAGS G SGA GAAAAAAAAAGAGQGAGSGSGAGAAAAAAAAAAAAQQQQQQQAAAAAAAAAAAAAGSG QGAS FGVTQQFGAPSGAAS SAAAAAAAAAAAAAGSGAGQEAGTGAGAAAAAAAAGAAG SGAGQGAGSGAGAAAAAAAAAS AAGAGQGAGSGSGAGAAAAAAAAAAAAQQQQQQQAA AAAAAAAAAA
19	GGAQKQPSGESSVATASAAATSVTSAGAPVGKPGVPAPIFYPQGPLQQGPAPGPSYVQ PATSQQGPIGGAGRSNAFSSSFASALSGNRGFSEVISSASATAVASAFQKGLAPYGTA FALSAASAAADAYNSIGSGANAFAYAQAFARVLYPLVQQYGLSSSAKASAFASAIASS FSSGAAGQGQSIPYGGQQQPPMTISAASASAGASAAAVKGGQVGQGPYGGQQQSTAAS ASAAATTATA
20	GADGGSGLGGYGAGRGYGAGLGGADGAGAASAAAAAGGQGGRGGFGRLGSQGAGGAGQ GGAGAAAAVAAAGGDGGSGLGGYGAGRGYGAGLGGAGGAGAASAAAAAGGQGGRGGFG GLGSQGAGGAGQGGAGAAASGDGGSGLGGYGAGRGYGAGLGGADGAGAASAASAAGGQ GGRGGFGGLGSQGAGGAGQGGAGAAAAAATAGGDGGSGLGGYGAGRGYGAGLGGAGGA GAASAAAAA
21	GAGAGQGGRGGYGQGGFGGQGSGAGAGASAAAGAGAGQGGRGGYGQGGFGGQGSGAGA GASAAAGAGAGQGGRGGYGQGGFGGQGSGAGAGASAAAAAGAGQGGRGGYGQGGLGGS GSGAGAGAGAAAAAAAGAGGYGQGGLGGYGQGAGAGQGGLGGYGSGAGAGASAAAAAG AGGAGQGGLGGYGQGAGAGQGGLGGYGSGAGAGAAAAAAAGAGGSGQGGLGGYGSGGG AGGASAAAA
22	GAYAYAYAIANAFASILANTGLLSVSSAASVASSVASAIATSVSSSSAAAAASASAAA AASAGASAASSASASSSASAAAGAGAGAGAGASGASGAAGGSGGFGLSSGFGAGIGGL GGYPSGALGGLGIPSGLLSSGLLSPAANQRIASLIPLILSAISPNGVNFGVIGSNIAS LASQISQSGGGIAASQAFTQALLELVAAFIQVLSSAQIGAVSSSSASAGATANAFAQS LSSAFAG
23	GAAQKQPSGESSVATASAAATSVTSGGAPVGKPGVPAPIFYPQGPLQQGPAPGPSNVQ PGTSQQGPIGGVGGSNAFSSSFASALSLNRGFTEVISSASATAVASAFQKGLAPYGTA FALSAASAAADAYNSIGSGANAFAYAQAFARVLYPLVRQYGLSSSGKASAFASAIASS FSSGTSGQGPSIGQQQPPVTISAASASAGASAAAVGGGQVGQGPYGGQQQSTAASASA AAATATS
24	GAAQKQPSGESSVATASAAATSVTSGGAPVGKPGVPAPIFYPQGPLQQGPAPGPSNVQ PGTSQQGPIGGVGGSNAFSSSFASALSLNRGFTEVISSASATAVASAFQKGLAPYGTA FALSAASAAADAYNSIGSGANAFAYAQAFARVLYPLVRQYGLSSSGKASAFASAIASS FSSGTSGQGPSIGQQQPPVTISAASASAGASAAAVGGGQVGQGPYGGQQQSTAASASA AAATATS
25	GAAQKQPSGESSVATASAAATSVTSGGAPVGKPGVPAPIFYPQGPLQQGPAPGPSNVQ PGTSQQGPIGGVGGSNAFSSSFASALSLNRGFTEVISSASATAVASAFQKGLAPYGTA FALSAASAAADAYNSIGSGANAFAYAQAFARVLYPLVQQYGLSSSAKASAFASAIASS FSSGTSGQGPSIGQQQPPVTISAASASAGASAAAVGGGQVGQGPYGGQQQSTAASASA AAATATS

Seq. ID No.	ΛΛ
26	GGAQKQPSGESSVATASAAATSVTSAGAPVGKPGVPAPIFYPQGPLQQGPAPGPSNVQ PGTSQQGPIGGVGGSNAFSSSFASALSLNRGFTEVISSASATAVASAFQKGLAPYGTA FALSAASAAADAYNSIGSGANAFAYAQAFARVLYPLVQQYGLSSSAKASAFASAIASS FSSGTSGQGPSNGQQQPPVTISAASASAGASAAAVGGGQVSQGPYGGQQQSTAASASA AAATATS
27	GGAQKQPSGESSVATASAAATSVTSAGAPGGKPGVPAPIFYPQGPLQQGPAPGPSNVQ PGTSQQGPIGGVGGSNAFSSSFASALSLNRGFTEVISSASATAVASAFQKGLAPYGTA FALSAASAAADAYNSIGSGANAFAYAQAFARVLYPLVQQYGLSSSAKASAFASAIASS FSSGTSGQGPSIGQQQPPVTISAASASAGASAAAVGGGQVGQGPYGGQQQSTAASASA AAATATS
28	GPGGYGGPGQQGPGQGQQQGPASAAAAAAAAGPGGYGGPGQQGPGQGQQQGPASAAAA AAAAAGPGGYGGPGQQRPGQAQYGRGTGQQGQGPGAQQGPASAAAAAAAGAGLYGGPG QQGPGQGQQQGPASAAAAAAAAAAAGPGGYGGPGQQGPGQAQQQGPASAAAAAAAGPG GYSGPGQQGPGQAQQQGPASAAAAAAAAAGPGGYGGPGQQGPGQGQQQGPASAAAAAA ATAA
29	GAGGDGGLFLSSGDFGRGGAGAGAGAAAASAAAASSAAAGARGGSGFGVGTGGFGRGG AGDGASAAAASAAAASAAAAGAGGDSGLFLSSGDFGRGGAGAGAGAAAASAAAASAAA AGTGGVGGLFLSSGDFGRGGAGAGAGAAAASAAAASSAAAGARGGSGFGVGTGGFGRG GPGAGTGAAAASAAAASAAAAGAGGDSGLFLSSEDFGRGGAGAGTGAAAASAAAASAA AA
30	GAGRGYGGGYGGGAAAGAGAGAGAGRGYGGGYGGGAGSGAGSGAGAGGGSGYGRGAGA GAGAGAAAAAGAGAGGAGGYGGGAGAGAGASAAAGAGAGAGGAGGYGGGYGGGAGAGA GAGAAAAAGAGAGAGAGRGYGGGFGGGAGSGAGAGAGAGGGSGYGRGAGGYGGGYGGG AGTGAGAAAATGAGAGAGAGRGYGGGYGGGAGAGAGAGAGGGSGYGRGAGAGASVA A
31	GALGQGASVWSSPQMAENFMNGFSMALSQAGAFSGQEMKDFDDVRDIMNSAMDKMIRS GKSGRGAMRAMNAAFGSAIAEIVAANGGKEYQIGAVLDAVTNTLLQLTGNADNGFLNE ISRLITLFSSVEANDVSASAGADASGSSGPVGGYSSGAGAAVGQGTAQAVGYGGGAQG VASSAAAGATNYAQGVSTGSTQNVATSTVTTTTNVAGSTATGYNTGYGIGAAAGAAA
32	GGQGGQGGYDGLGSQGAGQGGYGQGGAAAAAAAASGAGSAQRGGLGAGGAGQGYGAGS GGQGGAGQGGAAAATAAAAGGQGGQGGYGGLGSQGSGQGGYGQGGAAAAAAAASGDGG AGQEGLGAGGAGQGYGAGLGGQGGAGQGGAAAAAAAAAGGQGGQGGYGGLGSQGAGQG GYGQGGAAAAAAAASGAGGAGQGGLGAAGAGQGYGAGSGGQGGAGQGGAAAAAAAAA
33	GGQGGQGGYGGLGSQGAGQGGYGQGGVAAAAAAASGAGGAGRGGLGAGGAGQEYGAVS GGQGGAGQGGEAAAAAAAAGGQGGQGGYGGLGSQGAGQGGYGQGGAAAAAAAASGAGG ARRGGLGAGGAGQGYGAGLGGQGGAGQGSASAAAAAAAGGQGGQGGYGGLGSQGSGQG GYGQGGAAAAAAAASGAGGAGRGSLGAGGAGQGYGAGLGGQGGAGQGGAAAAASAAA
34	GPGGYGGPGQQGPGQGQYGPGTGQQGQGPGGQQGPVGAAAAAAAAVSSGGYGSQGAGQ GGQQGSGQRGPAAAGPGGYSGPGQQGPGQGGQQGPASAAAAAAAAAGPGGYGGSGQQG PGQGRGTGQQGQGPGGQQGPASAAAAAAAGPGGYGGPGQQGPGQGQYGPGTGQQGQGP ASAAAAAAAGPGGYGGPGQQGPGQGQYGPGTGQQGQGPGGQQGPGGASAAAAAAA

Seq. ID No.	ΛΛ
35	GGYGPGAGQQGPGSGGQQGPGGQGPYGSGQQGPGGAGQQGPGGQGPYGPGAAAAAAAA AGGYGPGAGQQGPGGAGQQGPGSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGGQG PYGPSAAAAAAAAAGGYGPGAGQRSQGPGGQGPYGPGAGQQGPGSQGPGSGGQQGPGG QGPYGPSAAAAAAAAGPGAGRQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAA
36	GQGGQGGQGGLGQGGYGQGAGSSAAAAAAAAAAAAAAGRGQGGYGQGSGGNAAAAAAA AAAAASGQGSQGGQGGQGQGGYGQGAGSSAAAAAAAAAAAAASGRGQGGYGQGAGGNA AAAAAAAAAAAAAGQGGQGGYGGLGQGGYGQGAGSSAAAAAAAAAAAAGGQGGQGQGG YGQGSGGSAAAAAAAAAAAAAAGRGQGGYGQGSGGNAAAAAAAAAAAAAA
37	GRGPGGYGPGQQGPGGPGAAAAAAGPGGYGPGGYGPGQQGPGGPGAAAAAAAGRGPGG YGPGQQGPGQQGPGGSGAAAAAAGRGPGGYGPGQQGPGGPGAAAAAAGPGGYGPGQQG PGAAAAAAAAGRGPGGYGPGQQGPGGPGAAAAAAAGRGPGGYGPGQQGPGQQGPGGSG AAAAAAGRGPGGYGPGQQGPGGPGAAAAAAGPGGYGPGQQGPGAAAAAAAA
38	GRGPGGYGPGQQGPGGSGAAAAAAGRGPGGYGPGQQGPGGPGAAAAAAGPGGYGPGQQ GTGAAAAAAAAGSGAGGYGPGQQGPGGPGAAAAAAGPGGYGPGQQGPGAAAAAAAGSGP GGYGPGQQGPGGS SAAAAAAGPGRYGPGQQGPGAAAAASAGRGPGGYGPGQQGPGGPG AAAAAAGPGGYGPGQQGPGAAAAAAAGSGPGGYGPGQQGPGGPGAAAAAAA
39	GAAATAGAGASVAGGYGGGAGAAAGAGAGGYGGGYGAVAGSGAGAAAAAS S GAGGAAG YGRGYGAGS GAGAGAGTVAAYGGAGGVATSS S SATAS GS R I VT S GGYGYGTS AAAGAG VAAGSYAGAVNRLSSAEAASRVSSNIAAIASGGASALPSVI SNIYSGWASGVSSNEA L IQALLELLSALVHVLSSAS IGNVS SVGVDSTLNWQDSVGQYVG
40	GGQGGFSGQGQGGFGPGAGSSAAAAAAAAAAARQGGQGQGGFGQGAGGNAAAAAAAAA AAAAAQQGGQGGFSGRGQGGFGPGAGSSAAAAAAGQGGQGQGGFGQGAGGNAAAAAAA AAAAAAAAGQGGQGRGGFGQGAGGNAAAAAAAAAAAAAAQQGGQGGFGGRGQGGFGP GAGSSAAAAAAGQGGQGRGGFGQGAGGNAAAASAAAAASAAAAGQ
41	GGYGPGAGQQGPGGAGQQGPGSQGPGGAGQQGPGGQGPYGPGAAAAAAAVGGYGPGAG QQGPGSQGPGSGGQQGPGGQGPYGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQG PGGLGPYGPSAAAAAAAAGGYGPGAGQQGPGSQGPGSGGQQRPGGLGPYGPSAAAAAA AAGGYGPGAGQQGPGSQGPGSGGQQRPGGLGPYGPSAAAAAAAA
42	GAGAGGGYGGGYSAGGGAGAGSGAAAGAGAGRGGAGGYSAGAGTGAGAAAGAGTAGGY S GGYGAGAS SSAGSS F I S S SSMSSS QATGYS S S S GYGGGAAS AAAGAGAAAGGYGGGY GAGAGAGAAAASGATGRVANS LGAMAS GGI NALPGVFSNI FSQVSAASGGAS GGAVLV QALTEVIALLLHI LSSAS I GNVSSQGLEGSMAIAQQAIGAYAG
43	GAGAGGAGGYAQGYGAGAGAGAGAGTGAGGAGGYGQGYGAGSGAGAGGAGGYGAGAGA GAGAGDASGYGQGYGDGAGAGAGAAAAAGAAAGARGAGGYGGGAGAGAGAGAGAAGGY GQGYGAGAGEGAGAGAGAGAVAGAGAAAAAGAGAGAGGAEGYGAGAGAGGAGGYGQSY GDGAAAAAGSGAGAGGSGGYGAGAGAGSGAGAAGGYGGGAGA
44	GPGGYGPGQQGPGGYGPGQQGPGRYGPGQQGPSGPGSAAAAAAGSGQQGPGGYGPRQQ GPGGYGQGQQGPSGPGSAAAASAAASAESGQQGPGGYGPGQQGPGGYGPGQQGPGGYG PGQQGPSGPGSAAAAAAAASGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSAAAAAA AASGPGQQGPGGYGPGQQGPGGYGPGQQGLSGPGSAAAAAAA

Seq. ID No.	ΛΛ
45	GRGPGGYGQGQQGPGGPGAAAAAAGPGGYGPGQQGPGAAAAAAAGSGPGGYGPGQQGP GRSGAAAAAAAAGRGPGGYGPGQQGPGGPGAAAAAAGPGGYGPGQQGPGAAAAASAGR GPGGYGPGQQGPGGSGAAAAAAGRGPGGYGPGQQGPGGPGAAAAAAAGRGPGGYGPGQ QGPGQQGPGGSGAAAAAAGRGPGGYGPGQQGPGGPGAAAAAA
46	GVGAGGEGGYDQGYGAGAGAGSGGGAGGAGGYGGGAGAGSGGGAGGAGGYGGGAGAGA GAGAGGAGGYGGGAGAGTGARAGAGGVGGYGQSYGAGASAAAGAGVGAGGAGAGGAGG YGQGYGAGAGIGAGDAGGYGGGAGAGASAGAGGYGGGAGAGAGGVGGYGKGYGAGSGA GAAAAAGAGAGSAGGYGRGDGAGAGGASGYGQGYGAGAAA
47	GYGAGAGRGYGAGAGAGAGAVAASGAGAGAGYGAGAGAGAGAGYGAGAGRGYGAGAGA GAGSGAASGAGAGAGYGAGAGAGAGYGAGAGSGYGTGAGAGAGAAAAGGAGAGAGYGA GAGRGYGAGAGAGAASGAGAGAGAGAASGAGAGSGYGAGAAAAGGAGAGAGGGYGAGA GRGYGAGAGAGAGAGSGSGSAAGYGQGYGSGSGAGAAA
48	GQGTDSSASSVSTSTSVSSSATGPDTGYPVGYYGAGQAEAAASAAAAAAASAAEAATI AGLGYGRQGQGTDSSASSVSTSTSVSSSATGPDMGYPVGNYGAGQAEAAASAAAAAAA SAAEAATIASLGYGRQGQGTDSSASSVSTSTSVSSSATGPGSRYPVRDYGADQAEAAA SAAAAAAAAASAAEEIASLGYGRQ
49	GQGTDSVASSASSSASASSSATGPDTGYPVGYYGAGQAEAAASAAAAAAASAAEAATI AGLGYGRQGQGTDSSASSVSTSTSVSSSATGPGSRYPVRDYGADQAEAAASATAAAAA AASAAEEIASLGYGRQGQGTDSVASSASSSASASSSATGPDTGYPVGYYGAGQAEAAA SAAAAAAASAAEAATIAGLGYGRQ
50	GQGGQGGYGGLGQGGYGQGAGSSAAAAAAAAAAAAAGGQGGQGQGRYGQGAGSSAAAA AAAAAAAAAGRGQGGYGQGSGGNAAAAAAAAAAAASGQGSQGGQGGQGQGGYGQGAG SSAAAAAAAAAAAAASGRGQGGYGQGAGGNAAAAAAAAAAAAAAGQGGQGGYGGLGQG GYGQGAGSSAAAAAAAAAAA
51	GGLGGQGGLGGLGSQGAGLGGYGQGGAGQGGAAAAAAAAGGLGGQGGRGGLGSQGAGQ GGYGQGGAGQGGAAAAAAAAGGLGGQGGLGALGSQGAGQGGAGQGGYGQGGAAAAAAG GLGGQGGLGGLGSQGAGQGGYGQGGAGQGGAAAAAAAAGGLGGQGGLGGLGSQGAGPG GYGQGGAGQGGAAAAAAAA
52	GGQGRGGFGQGAGGNAAAAAAAAAAAAAAQQVGQFGFGGRGQGGFGPFAGSSAAAAAA ASAAAGQGGQGQGGFGQGAGGNAAAAAAAAAAAARQGGQGQGGFSQGAGGNAAAAAAA AAAAAAAAQQGGQGGFGGRGQGGFGPGAGSSAAAAAAATAAAGQGGQGRGGFGQGAGS NAAAAAAAAAAAAAAAGQ
53	GGQGGQGGYGGLGSQGAGQGGYGAGQGAAAAAAAAGGAGGAGRGGLGAGGAGQGYGAG LGGQGGAGQAAAAAAAGGAGGARQGGLGAGGAGQGYGAGLGGQGGAGQGGAAAAAAAA GGQGGQGGYGGLGSQGAGQGGYGAGQGGAAAAAAAAGGQGGQGGYGGLGSQGAGQGGY GGRQGGAGAAAAAAAA
54	GGAGQRGYGGLGNQGAGRGGLGGQGAGAAAAAAAGGAGQGGYGGLGNQGAGRGGQGAA AAAGGAGQGGYGGLGSQGAGRGGQGAGAAAAAAVGAGQEGIRGQGAGQGGYGGLGSQG SGRGGLGGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGVRQGGYGGLGSQGA GRGGQGAGAAAAAA

Seq. ID No.	AA
55	GGAGQGGLGGQGAGQGAGASAAAAGGAGQGGYGGLGSQGAGRGGEGAGAAAAAAGGAG QGGYGGLGGQGAGQGGYGGLGSQGAGRGGLGGQGAGAAAAGGAGQGGLGGQGAGQGAG AAAAAAGGAGQGGYGGLGSQGAGRGGLGGQGAGAVAAAAAGGAGQGGYGGLGSQGAGR GGQGAGAAAAAA
56	GAGAGAGAGSGAGAAGGYGGGAGAGVGAGGAGGYDQGYGAGAGAGSGAGAGGAGGYGG GAGAGADAGAGGAGGYGGGAGAGAGARAGAGGVGGYGQSYGAGAGAGAGVGAGGAGAG GADGYGQGYGAGAGTGAGDAGGYGGGAGAGASAGAGGYGGGAGAGGVGVYGKGYGSGS GAGAAAAA
57	GGAGGYGVGQGYGAGAGAGAAAGAGAGGAGGYGAGQGYGAGAGVGAAAAAGAGAGVGG AGGYGRGAGAGAGAGAGAAAGAGAGAAAGAGAGGAGGYGAGQGYGAGAGVGAAAAAGA GAGVGGAGGYGRGAGAGAGAGGAGGYGRGAGAGAGAGAGGAGGYGAGQGYGAGA GAGAAAAA
58	GEAFSASSASSAVVFESAGPGEEAGSSGDGASAAASAAAAAGAGSGRRGPGGARSRGG AGAGAGAGSGVGGYGSGSGAGAGAGAGAGGEGGFGEGQGYGAGAGAGFGSGAGAGA GAGSGAGAGEGVGSGAGAGAGAGFGVGAGAGAGAGAGFGSGAGAGSGAGAGYGAGRAG GRGRGGRG
59	GEAFSASSASSAVVFESAGPGEEAGSSGGGASAAASAAAAAGAGSGRRGPGGARSRGG AGAGAGAGSGVGGYGSGSGAGAGAGAGAGGEGGFGEGQGYGAGAGAGFGSGAGAGA GAGSGAGAGEGVGSGAGAGAGAGFGVGAGAGAGAGAGFGSGAGAGSGAGAGYGAGRAG GRGRGGRG
60	GNGLGQALLANGVLNSGNYLQLANSLAYSFGSSLSQYSSSAAGASAAGAASGAAGAGA GAASSGGSSGSASSSTTTTTTTSTSAAAAAAAAAAAAASAAASTSASASASASASASAF SQTFVQTVLQSAAFGSYFGGNLSLQSAQAAASAAAQAAAQQIGLGSYGYALANAVASA FASAGANA
61	GNGLGQALLANGVLNSGNYLQLANSLAYSFGSSLSQYSSSAAGASAAGAASGAAGAGA GAASSGGSSGSASSSTTTTTTTSTSAAAAAAAAAAAAASAAASTSASASASASASASAF SQTFVQTVLQSAAFGSYFGGNLSLQSAQAAASAAAQAAAQQIGLGSYGYALANAVASA FASAGANA
62	GNGLGQALLANGVLNSGNYLQLANSLAYSFGSSLSQYSSSAAGASAAGAASGAAGAGA GAASSGGSSGSASSSTTTTTTTSTSAAAAAAAAAAAAASAAASTSASASASASASASAF SQTFVQTVLQSAAFGSYFGGNLSLQSAQAAASAAAQAAAQQIGLGSYGYALANAVASA FASAGANA
63	GASGAGQGQGYGQQGQGGSSAAAAAAAAAAAAAAAQGQGQGYGQQGQGSAAAAAAAAA AGASGAGQGQGYGQQGQGSAAAAAAAAAAGASGAGQGQGYGQQGQGGSSAAAAAAAAA AAAAAAAQGQGYGQQGQGSAAAAAAAAAGASGAGQGQGYGQQGQGGSSAAAAAAAAA AAAAAA
64	GRGQGGYGQGSGGNAAAAAAAGQGGFGGQEGNGQGAGSAAAAAAAAAAAAGGSGQGRY GGRGQGGYGQGAGAAASAAAAAAAAAGQGGFGGQEGNGQGAGSAAAAAAAAAAAGG SGQGGYGGRGQGGYGQGAGAAAAAAAAAAAAAGQGGQGGFGSQGGNGQGAGSAAAAA AAAAA

Seq. ID No.	ΛΛ
65	GQNTPWSSTELADAFINAFMNEAGRTGAFTADQLDDMSTIGDTIKTAMDKMARSNKSS KGKLQALNMAFASSMAEIAAVEQGGLSVDAKTNAIADSLNSAFYQTTGAANPQFVNEI RSLINMFAQSSANEVSYGGGYGGQSAGAAASAAAAGGGGQGGYGNLGGQGAGAAAAAA ASAA
66	GQNTPWSSTELADAFINAFLNEAGRTGAFTADQLDDMSTIGDTLKTAMDKMARSNKSS QSKLQALNMAFASSMAEIAAVEQGGLSVAEKTNAIADSLNSAFYQTTGAVNVQFVNEI RSLISMFAQASANEVSYGGGYGGGQGGQSAGAAAAAASAGAGQGGYGGLGGQGAGSAA AAAA
67	GGQGGQGGYGGLGSQGAGQGGYGQGGAAAAAASAGGQGGQGGYGGLGSQGAGQGGYGG GAFSGQQGGAASVATASAAASRLSSPGAASRVSSAVTSLVSSGGPTNSAALSNTISNV VSQISSSNPGLSGCDVLVQALLEIVSALVHILGSANIGQVNSSGVGRSASIVGQSINQ AFS
68	GGAGQGGYGGLGGQGAGAAAAAAGGAGQGGYGGQGAGQGAAAAAASGAGQGGYEGPGA GQGAGAAAAAAGGAGQGGYGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGA GAAAAAAGGAGQGGYGGQGAGQGAAAAAAGGAGQGGYGGLGSGQGGYGRQGAGAAAAA AAA
69	GASSAAAAAAATATSGGAPGGYGGYGPGIGGAFVPASTTGTGSGSGSGAGAAGSGGLG GLGSSGGSGGLGGGNGGSGASAAASAAAASSSPGSGGYGPGQGVGSGSGSGAAGGSGT GSGAGGPGSGGYGGPQFFASAYGGQGLLGTSGYGNGQGGASGTGSGGVGGSGSGAGSN S
70	GQPIWTNPNAAMTMTNNLVQCASRSGVLTADQMDDMGMMADSVNSQMQKMGPNPPQHR LRAMNTAMAAEVAEWATSPPQSYSAVLNTIGACLRESMMQATGSVDNAFTNEVMQLV KMLSADSANEVSTASASGASYATSTSSAVSSSQATGYSTAAGYGNAAGAGAGAAAAVS
71	GQKIWTNPDAAMAMTNNLVQCAGRSGALTADQMDDLGMVSDSVNSQVRKMGANAPPHK IKAMSTAVAAGVAEWASSPPQSYSAVLNTIGGCLRESMMQVTGSVDNTFTTEMMQMV NMFAADNANEVSASASGSGASYATGTSSAVSTSQATGYSTAGGYGTAAGAGAGAAAAA
72	GSGYGAGAGAGAGSGYGAGAGAGSGYGAGAGAGAGSGYVAGAGAGAGAGSGYGAGAGA GAGSSYSAGAGAGAGSGYGAGSSASAGSAVSTQTVSSSATTSSQSAAAATGAAYGTRA STGSGASAGAAASGAGAGYGGQAGYGQGGGAAAYRAGAGSQAAYGQGASGSSGAAAAA
73	GGQGGRGGFGGLSSQGAGGAGQGGSGAAAAAAAAGGDGGSGLGDYGAGRGYGAGLGGA GGAGVASAAASAAASRLSSPSAASRVSSAVTSLISGGGGPTNPAALSNTFSNWYQISV SSPGLSGCDVLIQALLELVSALVHILGSAIIGQVNSSAAGESASLVGQSVYQAFS
74	GVGQAATPWENSQLAEDFINSFLRFIAQSGAFSPNQLDDMSSIGDTLKTAIEKMAQSR KSSKSKLQALNMAFASSMAEIAVAEQGGLSLEAKTNAIANALASAFLETTGFVNQQFV SEIKSLIYMIAQASSNEISGSAAAAGGGSGGGGGSGQGGYGQGASASASAAAA
75	GGGDGYGQGGYGNQRGVGSYGQGAGAGAAATSAAGGAGSGRGGYGEQGGLGGYGQGAG AGAASTAAGGGDGYGQGGYGNQGGRGSYGQGSGAGAGAAVAAAAGGAVSGQGGYDGEG GQGGYGQGSGAGAAVAAASGGTGAGQGGYGSQGSQAGYGQGAGFRAAAATAAA
76	GAGAGYGGQVGYGQGAGASAGAAAAGAGAGYGGQAGYGQGAGGSAGAAAAGAGAGRQA GYGQGAGASARAAAAGAGTGYGQGAGASAGAAAAGAGAGSQVGYGQGAGASSGAAAAA GAGAGYGGQVGYEQGAGASAGAEAAASSAGAGYGGQAGYGQGAGASAGAAAA

Seq. ID No.	ΛΛ
77	GGAGQGGYGGLGGQGAGQGGLGGQRAGAAAAAAGGAGQGGYGGLGSQGAGRGGYGGVG SGASAASAAASRLSS PEAS SRVSSAVSNLVS SGPTNSAALSST I SNWSQI SASNPGL S GCDVLVQALLEVVS AL I QI LGS S S I GQVNYGTAGQAAQI VGQS VYQALG
78	GGYGPGSGQQGPGGAGQQGPGGQGPYGPGSS SAAAVGGYGPSSGLQGPAGQGPYGPGA AASAAAAAGASRLSS PQAS SRVSSAVSSLVS SGPTNSAALTNT I SSWSQI SASNPGL S GCDVL I QALLE I VSALVH I LGYS S I GQI NYDAAAQYAS LVGQS VAQALA
79	GGAGAGQGSYGGQGGYGQGGAGAATATAAAAGGAGSGQGGYGGQGGLGGYGQGAGAGA AAAAAAAAGGAGAGQGGYGGQGGQGGYGQGAGAGAAAAAAGGAGAGQGGYGGQGGYGQ GGGAGAAAAAAAAASGGSGSGQGGYGGQGGLGGYGQGAGAGAGAAASAAAA
80	GQGGQGGYGRQSQGAGSAAAAAAAAAAAAAAGSGQGGYGGQGQGGYGQS SASASAAAS AASTVANSVSRLS S PSAVSRVSSAVSSLVSNGQVNMAALPNI I SNI SSSVSASAPGAS GCEVI VQALLEVI TALVQIVSSSSVGYINPS AVNQI TNWANAMAQVMG
81	GGAGQGGYGGLGGQGSGAAAAGTGQGGYGSLGGQGAGAAGAAAAAVGGAGQGGYGGVG SAAASAAASRLSS PEASSRVSSAVSNLVSSGPTNSAALSNTI SNWSQI SSSNPGLSG CDVLVQALLEWSALIHI LGSSS IGQVNYGSAGQATQ IVGQSVYQALG
82	GAGAGGAGGYGAGQGYGAGAGAGAAAGAGAGGARGYGARQGYGSGAGAGAGARAGGAG GYGRGAGAGAAAASGAGAGGYGAGQGYGAGAGAVASAAAGAGSGAGGAGGYGRGAGAV AGAGAGGAGGYGAGAGAAAGVGAGGSGGYGGRQGGYSAGAGAGAAAAA
83	GQGGQGGYGGLGQGGYGQGAGSSAAAAAAAAAAAGRGQGGYGQGSGGNAAAAAAAAAA AAS GQGGQGGQGGQGGQGGYGQGAGS SAAAAAAAAAAAAAAAGRGQGGYGQGAGGNAAA AAAAAAAAASGQGGQGGQGGQGGQGGYGQGAGSSAAAAAAAAAAAAAA
84	GGYGPGSGQQGPGQQGPGQQGPGQQGPYGAGASAAAAAAGGYGPGSGQQGPGVRVAAP VASAAASRLSSSAASSRVS SAVSSLVSSGPTTPAALSNTI SSAVSQI SASNPGLSGCD VLVQALL EVVS ALVH I LGS S SVGQ I NYGAS AQYAQMVGQSVTQALV
85	GAGAGGAGYGRGAGAGAGAAAGAGAGAAAGAGAGAGGYGGQGGYGAGAGAGAAAAAGA GAGGAAGYSRGGRAGAAGAGAGAAAGAGAGAGGYGGQGGYGAGAGAGAAAAAGAGSGG AGGYGRGAGAGAAAGAGAAAGAGAGAGGYGGQGGYGAGAGAAAAA
86	GAGAGRGGYGRGAGAGGYGGQGGYGAGAGAGAAAAAGAGAGGYGDKHACWSRCRYTV ASTTSRLSSAEAS SRI SSAASTLVSGGYLNTAALPSVI SDLFAQVGASS PGVSDS EVL I QVLLE IVSSLIHI LSSSSVGQVDFSSVGSSAAAVGQSMQWMG
87	GAGAGAGGAGGYGRGAGAGAGAGAGAAAGQGYGSGAGAGAGASAGGAGSYGRGAGAGA AAASGAGAGGYGAGQGYGAGAGAVASAAAGAGSGAGGAGGYGRGAVAGSGAGAGAGAG GAGGYGAGAGAGAAAGAVAGGSGGYGGRQGGYSAGAGAGAAAAA
88	GPGGYGPVQQGPSGPGSAAGPGGYGPAQQGPARYGPGSAAAAAAAAGSAGYGPGPQAS AAASRLAS PDSGARVASAVSNLVSSGPTSSAALSSVI SNAVSQ IGASNPGLSGCDVLI QALLE IVSACVTI LSSSS I GQVNYGAASQFAQWGQSVLSAFS
89	GTGGVGGLFLSSGDFGRGGAGAGAGAAAASAAAASSAAAGARGGSGFGVGTGGFGRGG AGAGTGAAAASAAAASAAAAGAGGDGGLFLSSGDFGRGGAGAGAGAAAASAAAASSAA AGARGGSGFGVGTGGFGRGGAGDGASAAAASAAAASAAAA

Seq. ID No.	ΛΛ
90	GGYGPGAGQQGPGGAGQQGPGGQGPYGPSVAAAASAAGGYGPGAGQQGPVASAAVSRL SSPQASSRVSSAVSSLVSSGPTNPAALSNAMSSWSQVSASNPGLSGCDVLVQALLE I VSALVH I LGSSS I GQI NYAASSQYAQMVGQSVAQALA
91	GGAGQGGYGGLGSQGAGRGGYGGQGAGAAAAATGGAGQGGYGGVGSGASAASAAASRL SSPQASSRVSSAVSNLVASGPTNSAALSSTI SNAVSQ IGASNPGLSGCDVLIQALLEV VSALIHI LGSSS I GQVNYGSAGQATQIVGQSVYQALG
92	GGAGQGGYGGLGSQGAGRGGYGGQGAGAAVAAIGGVGQGGYGGVGSGASAASAAASRL SSPEASSRVSSAVSNLVSSGPTNSAALSSTI SNWSQ IGASNPGLSGCDVLIQALLEV VSALVHI LGSSS I GQVNYGSAGQATQIVGQSVYQALG
93	GASGGYGGGAGEGAGAAAAAGAGAGGAGGYGGGAGSGAGAVARAGAGGAGGYGSGIGG GYGSGAGAAAGAGAGGAGAYGGGYGTGAGAGARGADSAGAAAGYGGGVGTGTGSSAGY GRGAGAGAGAGAAAGSGAGAAGGYGGGYGAGAGAGA
94	GAGSGQGGYGGQGGLGGYGQGAGAGAAAGASGSGSGGAGGQGGLGGYGQGAGAGAAAAA AGASGAGQGGFGPYGSSYQSSTSYSVTSQGAAGGLGGYGQGSGAGAAAAGAAGQGGQG GYGQGAGAGAGAGAGQGGLGGYGQGAGSSAASAAAA
95	GGAGQGGYGGLGGQGVGRGGLGGQGAGAAAAGGAGQGGYGGVGSGASAASAAASRLSS PQASSRLSSAVSNLVATGPTNSAALSSTI SNWSQIGASNPGLSGCDVL IQALLEWS ALIQI LGSSS I GQVNYGSAGQATQI VGQSVYQALG
96	GAGSGGAGGYGRGAGAGAGAAAGAGAGAGSYGGQGGYGAGAGAGAAAAAGAGAGAGGY GRGAGAGAGAGAGAAARAGAGAGGAGYGGQGGYGAGAGAGAAAAAGAGAGGAGGYGRG AGAGAGAAAGAGAGAGGYGGQSGYGAGAGAAAAA
97	GASGAGQGQGYGQQGQGGSSAAAAAAAAAAAQGQGQGYGQQGQGYGQQGQGGSSAAAA AAAAAAAAAQGQGQGYGQQGQGSAAAAAAAAAGASGAGQGQGYGQQGQGGSSAAAAAA AAAAAAAAAQGQGYGQQGQGSAAAAAAAAAAAA

[0085] In an embodiment a block copolymer polypeptide repeat unit that forms fibers with good mechanical properties is synthesized using SEQ ID NO. 1. This repeat unit contains 6 quasi-repeats, each of which includes motifs that vary in composition, as described herein. This repeat unit can be concatenated 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 times to form polypeptide molecules from 20 kDal to 535 kDal, or greater than 20 kDal, or greater than 10 kDal, or greater than 5 kDal, or from 5 to 400 kDal, or from 5 to 300 kDal, or from 5 to 200 kDal, or from 5 to 100 kDal, or from 5 to 50 kDal, or from 5 to 600 kDal, or from 5 to 800 kDal, or from 5 to 1000 kDal, or from 10 to 400 kDal, or from 10 to 300 kDal, or from 10 to 200 kDal,

or from 10 to 100 kDal, or from 10 to 50 kDal, or from 10 to 600 kDal, or from 10 to 800 kDal, or from 10 to 1000 kDal, or from 20 to 400 kDal, or from 20 to 300 kDal, or from 20 to 200 kDal, or from 20 to 100 kDal, or from 20 to 50 kDal, or from 40 to 300 kDal, or from 40 to 500 kDal, or from 20 to 600 kDal, or from 20 to 800 kDal, or from 20 to 1000 kDal. This polypeptide repeat unit also contains poly-alanine regions related to nanocrystalline regions, and glycine-rich regions related to beta-turn containing less-crystalline regions. In other embodiments the repeat is selected from any of the sequences listed as SEQ ID Nos: 2-97.

[0086] In some embodiments, a long uniform RPFs comprises proteins containing one or more sequences from the list of SEQ ID Nos: 1-97.

[0087] In some embodiments, the quasi-repeat unit of the polypeptide can be described by the formula {GGY-[GPG-**Xi**]$_{n1}$**i**-GPS-(A)$_{n2}$}, where **Xi** is independently selected from the group consisting of SGGQQ, GAGQQ, GQGPY, AGQQ and SQ, n1 is a number from 4 to 8, and n2 is a number from 6 to 20. The repeat unit is composed of multiple quasi-repeat units. In additional embodiments, 3 "long" quasi repeats are followed by 3 "short" quasi-repeat units. As mentioned above, short quasi- repeat units are those in which nl=4 or 5. Long quasi-repeat units are defined as those in which nl=6, 7 or 8. In some embodiments, all of the short quasi-repeats have the same **Xi** motifs in the same positions within each quasi-repeat unit of a repeat unit. In some embodiments, no more than 3 quasi-repeat units out of 6 share the same **Xi** motifs.

[0088] In additional embodiments, a repeat unit is composed of quasi-repeat units that do not use the same **Xi** more than two occurrences in a row within a repeat unit. In additional embodiments, a repeat unit is composed of quasi-repeat units where at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 of the quasi-repeats do not use the same **Xi** more than 2 times in a single quasi-repeat unit of the repeat unit.

[0089] In some embodiments, the structure of fibers formed from the described polypeptides form beta-sheet structures, beta-turn structures, or alpha-helix structures. In some embodiments, the secondary, tertiary and quaternary protein structures of the formed fibers are described as having nanocrystalline beta-sheet regions, amorphous beta-turn regions, amorphous alpha helix regions, randomly spatially distributed nanocrystalline regions embedded in a non-crystalline matrix, or randomly oriented nanocrystalline regions embedded in a non-crystalline matrix.

[0090] In some embodiments, the polypeptides utilized to form fibers with mechanical properties as described herein include glycine-rich regions from 20 to 100 amino acids long concatenated with poly-alanine regions from 4 to 20 amino acids long. In some embodiments, polypeptides utilized to form fibers with good mechanical properties comprise 5-25% poly-alanine regions (from 4 to 20 poly-alanine residues). In some embodiments, polypeptides utilized to form fibers with good mechanical properties comprise 25-50% glycine. In some embodiments, polypeptides utilized to form fibers with good mechanical properties comprise 15-35% GGX, where X is any amino acid. In some embodiments, polypeptides utilized to form fibers with good mechanical properties comprise 15-60% GPG. In some embodiments, polypeptides utilized to form fibers with good mechanical properties comprise 10-40% alanine. In some embodiments, polypeptides utilized to form fibers with good mechanical properties comprise 0-20% proline. In some embodiments, polypeptides utilized to form fibers with good mechanical properties comprise 10-50% beta-turns. In some embodiments, polypeptides utilized to form fibers with good mechanical properties comprise 10-50% alpha-helix composition. In some embodiments all of these compositional ranges will apply to the same polypeptide. In some embodiments two or more of these compositional ranges will apply to the same polypeptide.

RECOMBINANT PROTEIN FIBER SPIN DOPE AND SPINNING PARAMETERS

[0091] Long uniform RPFs can be produced using the following processing conditions and methods.

[0092] In some embodiments, a spin dope is synthesized containing proteins expressed from any of the polypeptides of the present disclosure. The spin dope is prepared using published techniques such as those found in WO2015042164 A2, especially at paragraphs 114-134. In some embodiments, a fiber spinning solution was prepared by dissolving a powder comprising the purified and dried block copolymer polypeptide (hereinafter "recombinant polypeptide powder") in a formic acid-based spinning solution, using standard mixing techniques. Depending on the embodiment, the recombinant polypeptide powder can comprise various impurities and the purity of the recombinant protein powder as expressed by the amount of recombinant protein (i.e., proteinaceous block copolymer that forms protein fibers) by mass can range from 30-100%, 40-90%, 50-90%, 60-90%, 30-70%, 30-40% and/or 30-65%. Spin dopes were mixed until the polypeptide was completely dissolved as determined by visual inspection. Spin dopes were degassed and undissolved particulates were removed by centrifugation.

[0093] In an embodiment the fraction of protein that is at least some percentage (e.g., 80%) of the intended length is determined through quantitative analysis of the results of a size-separation process. In an embodiment, the size-separation process can include size-exclusion chromatography. In an embodiment, the size-separation process can include gel electrophoresis. The quantitative analysis can include determining the fraction of total protein falling within a designated size range by integrating the area of a chromatogram or densitometric scan peak. For example, if a sample is run through a size-separation process, and the relative areas under the peaks corresponding to full-length, 60% full-length and 20% full length are 3:2:1, then the

fraction that is full length corresponds to 3 parts out of a total of 6 parts by mass = 50% mass ratio.

[0094] In some embodiments, the proteins of the spin dope, expressed from any of the polypeptides of the present disclosure, are substantially monodisperse. In some embodiments, the proteins of the spin dope, expressed from any of the polypeptides of the present disclosure, have from 5% to 99%, or from 5% to 50%, or from 50% to 99%, or from 20% to 80%, or from 40% to 60%, or from 5% to 30%, or from 70% to 99%, or from 5% to 20%, or from 5% to 10%, or from 80% to 99%, or from 90% to 99% of the protein in the spin dope having molecular weight from 5% to 99%, or from 5% to 50%, or from 50% to 99%, or from 20% to 80%, or from 40% to 60%, or from 5% to 30%, or from 70% to 99%, or from 5% to 20%, or from 5% to 10%, or from 80% to 99%, or from 90% to 99% of the molecular weight of the encoded proteins. The "encoded proteins" are defined as the polypeptide amino acid sequences that are encoded by the DNA utilized in protein expression. In other words, the "encoded proteins" are the polypeptides that would be produced if there were no imperfect processes (e.g., transcription errors, protein degradation, homologous recombination, truncation, protein fragmentation, protein agglomeration) at any stage during protein production. A higher monodispersity of proteins in the spin dopes, in other words a higher purity, can have the advantage of producing fibers with better mechanical properties, such as higher initial modulus, higher extensibility, higher ultimate tensile strength, and higher maximum tensile strength.

[0095] In other embodiments, fibers with low monodispersity, <10%, or <15%, or <20%, or <25%, or <30%, or <35%, or <40%, or <45%, or <50% of the protein in the spin dope having molecular weight >50%, or >55%, or >60%, or >65%, or >70%, or >75%, or >80%, or >85%, or >90%, or >95%, or >99% of the molecular weight of the proteins encoded by the DNA utilized in protein expression, were still able to create fibers with good mechanical properties. The

mechanical properties described herein (e.g., high initial modulus and/or extensibility), from fibers formed from low purity spin dopes was achieved through the use of the long polypeptide repeat units, suitable polypeptide compositions and spin dope and fiber spinning parameters described elsewhere in the present disclosure.

[0096] In other embodiments, the proteins are produced via secretion from a microorganism such as *Pichia pastoris, Escherichia coli, Bacillus subtilis,* or mammalian cells. Optionally, the secretion rate is at least 20 mg /g DCW / hr (DCW = dry cell weight). Optionally, the proteins are then recovered, separated, and spun into fibers using spin dopes containing solvents. Some examples of the classes of solvents that can be used in spin dopes are aqueous, inorganic or organic, including but not limited to ethanol, methanol, isopropanol, t-butyl alcohol, ethyl acetate, and ethylene glycol. Various methods for synthesizing recombinant proteinaceous block copolymers have been published such as those found in WO201 50421 64 A2, especially at paragraphs 114-134.

[0097] In some embodiments, the fibers are extruded through a spinneret to form long uniform RPFs, for example greater than 20 m long. Continuous fiber manufacturing includes the following processes: pumping, filtration, fiber forming, and optionally, fiber treatment. The spin dope is pumped through a filter and subsequently through the spinneret, which contains small holes. Resistance in the fluid paths through the filter and the spinneret produces a pressure drop across each of these elements. The pumping pressure and type of pump required is dictated by the system elements' intrinsic fluid dynamic properties, the pathways used to interconnect them, and the viscosity of the spin dope liquid. Filtration is used to screen out particles that would lead to defects in the fiber, or lead to an obstruction of one of the spinneret holes. In some embodiments, screen filtration or deep bed type filtration systems is used. RPFs are formed using wet spinning, and the spin dope coagulates in a coagulation bath upon leaving the

spinneret holes. Due to the friction between the coagulated fiber and the coagulant, continuous fiber manufacture employs lower spinning speeds than those used for other spinning processes (such as melt spinning or dry spinning). In some embodiments post-spinning fiber treatments, such as cold drawing or hot drawing are used. Drawing imparts a higher degree of polymer orientation in the fiber, which leads to improved mechanical properties.

[0098] In some embodiments, a solution of polypeptide is spun into fibers using elements of processes known in the art. These processes include, for example, wet spinning, dry-jet wet spinning, and dry spinning. In preferred wet-spinning embodiments, the filament is extruded through an orifice into a liquid coagulation bath. In one embodiment, the filament can be extruded through an air gap prior to contacting the coagulation bath. In a dry-jet wet spinning process, the spinning solution is attenuated and stretched in an inert, non-coagulating fluid, e.g., air, before entering the coagulating bath. Suitable coagulating fluids are the same as those used in a wet-spinning process.

[0099] In other embodiments, the coagulation bath conditions for wet spinning are chosen to promote fiber formation with certain mechanical properties. Optionally, the coagulation bath is maintained at temperatures of 0-90° C, more preferably 20-60° C. Optionally, the coagulation bath comprises about 60%, 70%, 80%, 90%, or even 100% alcohol, preferably isopropanol, ethanol, or methanol. Optionally, the coagulation bath is 95:5, 90:10, 85:15, 80:20, 75:25, 70:30, 65:35, 60:40, 55:45 or 50:50 by volume methanol: water. Optionally, the coagulation bath contains additives to enhance the fiber mechanical properties, such as additives comprising ammonium sulfate, sodium chloride, sodium sulfate, or other protein precipitating salts at temperature from 20 to 60 °C.

[00100] In some embodiments, the extruded filament or fiber is passed through more than one bath. For embodiments in which more than one bath is used, the different baths have either

different or same chemical compositions. In some embodiments, the extruded filament or fiber is passed through more than one coagulation bath. For embodiments in which more than one coagulation bath is used, the different coagulation baths have either different or same chemical compositions. The residence time can be tuned to improve mechanical properties, such as from 2 seconds to 100 minutes in the coagulant bath. The reeling/drawing rate can be tuned to improve fiber mechanical properties, such as a rate from 0.1 to 100 meters/minute.

[00101] Optionally, the filament or fiber is also passed through one or more rinse baths to remove residual solvent and/or coagulant. Rinse baths of decreasing salt or alcohol concentration up to, preferably, an ultimate water bath, preferably follow salt or alcohol baths.

[00102] Following extrusion, the filament or fiber can be drawn. Drawing can improve the consistency, axial orientation and toughness of the filament. Drawing can be enhanced by the composition of a coagulation bath. Drawing may also be performed in a drawing bath containing a plasticizer such as water, glycerol or a salt solution. Drawing can also be performed in a drawing bath containing a crosslinker such as gluteraldehyde or formaldehyde. Drawing can be performed at temperature from 25-100 °C to alter fiber properties, preferably at 60 °C. As is common in a continuous process, drawing can be performed simultaneously during the coagulation, wash, plasticizing, and/or crosslinking procedures described previously. Drawing ratio depends on the filament being processed. In some embodiments, the drawing rate is about 4x, or 5x, or 6x, or 7x, or 8x, or 9x, or 10x, or llx, or 12x, or 13x, or 14x, or 15x the rate of reeling from the coagulation bath.

[00103] In certain embodiments of the invention, the filament is wound onto a spool after extrusion or after drawing. Winding rates are generally 1 to 500 m/min, preferably 10 to 50 m/min.

[00104] The draw ratio can also be tuned to improve fiber mechanical properties. In different embodiments the draw ratio was 1.5X to 6X. In one embodiment, lower draw ratios improved the fiber extensibility. In one embodiment, higher draw ratios improved the fiber maximum tensile strength. Drawing can also be done in different environments, such as in solution, in humid air, or at elevated temperatures.

[00105] The fibers of the present disclosure processed with residence times in coagulation baths at the longer end of the disclosed range produce corrugated cross sections. That is, each fiber has a plurality of corrugations (or alternatively "grooves") disposed at an outer surface of a fiber. Each of these corrugations is parallel to a longitudinal axis of the corresponding fiber on which the corrugations are disposed. The fibers of the present disclosure processed with higher ethanol content in the coagulation bath produce hollow core fibers. That is, the fiber includes an inner surface and an outer surface. The inner surface defines a hollow core parallel to the longitudinal axis of the fiber.

[00106] In some embodiments a coagulation bath or the first coagulation bath is prepared using combinations of one or more of water, acids, solvents and salts, including but not limited to the following classes of chemicals of Bronsted-Lowry acids, Lewis acids, binary hydride acids, organic acids, metal cation acids, organic solvents, inorganic solvents, alkali metal salts, and alkaline earth metal salts. Some examples of acids used in the preparation of a coagulation bath or the first coagulation bath are dilute hydrochloric acid, dilute sulfuric acid, formic acid and acetic acid. Some examples of solvents that are used in the preparation of the first coagulation bath are ethanol, methanol, isopropanol, t-butyl alcohol, ethyl acetate, and ethylene glycol. Examples of salts used in the preparation of a coagulation bath or the first coagulation bath include LiCl, KC1, BeCb, MgCb, CaCk, NaCl, ammonium sulfate, sodium sulfate, and other salts of nitrates, sulfates or phosphates.

[00107] In some embodiments, the chemical composition and extrusion parameters of a coagulation bath or the first coagulation bath are chosen so that the fiber remains translucent in a coagulation bath or the first coagulation bath. In some embodiments the chemical composition and extrusion parameters of a coagulation bath or the first coagulation bath are chosen to slow down the rate of coagulation of the fiber in a coagulation bath or the first coagulation bath, which improves the ability to draw the resulting fiber in subsequent drawing steps. In various embodiments, these subsequent drawing steps are done in different environments, including wet, dry, and humid air environments. Examples of wet environments include one or more additional baths or coagulation baths. In some embodiments, the fiber travels through one or more baths after the first coagulation bath. The one or more additional baths, or coagulation baths, are prepared, in embodiments, using combinations of one or more of water, acids, solvents and salts, including but not limited to the following classes of chemicals of Bronsted-Lowry acids, Lewis acids, binary hydride acids, organic acids, metal cation acids, organic solvents, inorganic solvents, alkali metal salts, and alkaline earth metal salts. Some examples of acids that are used in the preparation of the second baths or coagulant baths are dilute hydrochloric acid, dilute sulfuric acid, formic acid and acetic acid. Some examples of solvents that are used in the preparation of the second coagulant baths are ethanol, methanol, isopropanol, t-butyl alcohol, ethyl acetate, and ethylene glycol. Some examples of salts used in the preparation of a second bath or coagulation bath include LiCl, KC1, MgCb, CaCk, NaCl, ammonium sulfate, sodium sulfate, and other salts of nitrates, sulfates, or phosphates. In some embodiments, there are two coagulation baths, where the first coagulation bath has a different chemical composition than the second coagulation bath, and the second coagulation bath has a higher concentration of solvents than the first coagulation bath. In some embodiments, there are more than two coagulation baths, and the first coagulation bath has a different chemical composition than the second

coagulation bath, and the second coagulation bath has a lower concentration of solvents than the first coagulation bath. In some embodiments, there are two baths, the first being a coagulation bath and the second being a wash bath. In some embodiments, the first coagulation bath has a different chemical composition than the second wash bath, and the second wash bath has a higher concentration of solvents than the first bath. In some embodiments, there are more than two baths, and the first bath has a different chemical composition than the second bath, and the second bath has a lower concentration of solvents than the first bath.

[00108] In some embodiments a spin dope is further prepared using combinations of one or more of water, acids, solvents and salts, including but not limited to the following classes of chemicals of Bronsted-Lowry acids, Lewis acids, binary hydride acids, organic acids, metal cation acids, organic solvents, inorganic solvents, alkali metal salts, and alkaline earth metal salts. Some examples of acids that are used in the preparation of spin dopes are dilute hydrochloric acid, dilute sulfuric acid, formic acid and acetic acid. Some examples of solvents that are used in the preparation of spin dopes are ethanol, methanol, isopropanol, t-butyl alcohol, ethyl acetate, and ethylene glycol. Some examples of salts that are used in the preparation of spin dopes are LiCl, KC1, MgCk, CaCb, NaCl, ammonium sulfate, sodium sulfate, and other salts of nitrates, sulfates or phosphates.

[00109] In some embodiments, a spinneret is chosen to enhance the fiber mechanical properties. The dimensions of the spinneret can be from 0.001 cm to 5 cm long, and from 25 to 200 um in diameter. In some embodiments, a spinneret includes multiple orifices to spin multiple fibers simultaneously. In some embodiments, the cross-section of a spinneret gradually tapers to the smallest diameter at the orifice, is straight-walled and then quickly tapers to the orifice, or includes multiple constrictions. An extrusion pressure of a spin dope from a spinneret can also be varied to affect the fiber mechanical properties in a range from 10 to 1000 psi. The interaction

between fiber properties and extrusion pressure can be affected by spin dope viscosity, drawing/reeling rate, and coagulation bath chemistry.

[00110] The concentration of protein to solvent in the spin dope is also an important parameter. In some embodiments, the concentration of protein weight for weight is 20%, or 25%, or 30%, or 35%, or 40%, or 45% or 50%, or 55%, or from 20% to 55%, or from 20% to 40%, or from 30% to 40%, or from 30% to 55%, or from 30% to 50% in solution with solvents and other additives making up the remainder.

LONG UNIFORM RECOMBINANT PROTEIN FIBER SPIN DOPE AND SPINNING PARAMETERS

[00111] Methods to process long uniform RPFs are described. In some embodiments, the long uniform RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, and physical, mechanical and/or chemical properties that are uniform along the length of the fiber. In some embodiments, the physical, mechanical and/or chemical property has a CV along the length of the fiber less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%. When spinning long uniform RPFs, the

following process parameters and conditions are useful guidelines. The spin dope should be mixed to minimize viscosity gradients/inhomogeneities. Spin dope viscosity changes should be minimized so as to not vary significantly during the duration of the spinning event. The spin dope should be devoid of bubbles and particulates that blind the filter over time, clog the spinneret orifice, or introduce defects into the fiber that become breaking points during subsequent processing steps (e.g., during the drawing steps). The extrusion pressure should be consistent for the duration of the spinning event (e.g., the extrusion pressure should be non-pulsatile). The coagulation bath formulation also should not change appreciably during the spinning event (e.g., due to preferential evaporation of one of the bath components, water absorption, or chemical reactions between bath components).

[00112] The spin dope can be mixed to minimize viscosity gradients/inhomogeneities using elevated temperatures as well as high shear mixing approaches (as opposed to gentle agitation). Elevated temperatures include those above 22 °C and below the flash point of the dope solvent. Some high shear mixing methods include impeller-based mixing, homogenization, sonic mixing, and planetary mixing.

[00113] Spin dope viscosity changes can be minimized so as to not vary significantly during the duration of the spinning event by using methods such as continual and/or in-line agitation, temperature control of the dope to reduce the molecular chain mobility (e.g., from 4 °C to 22 °C, or at approximately 15 °C, or at approximately 7 °C), and the addition of one or more chaotropic additive that disrupts the silk chain hydrogen bonding (e.g., urea, MgCb, LiCl, LiBr, and/or sodium dodecyl sulfate). In some embodiments, one or more chaotropic additive can be added to the spin dope in a concentration of less than 15 wt% (i.e., percentage by weight), less than 10 wt%, less than 5 wt%, less than 4 wt%, less than 3 wt%, less than 2 wt%, or less than 1 wt%, or

from 0.1 wt% to 15 wt%, from 0.1 wt% to 10 wt%, from 0.1 wt% to 5 wt%, from 0.1 wt% to 2 wt%, or from 0.1 wt% to 1 wt%.

[00114] Various methods can be used to avoid bubbles and particulates that blind the filter over time, clog the spinneret orifice, or introduce defects into the fiber that become breaking points during subsequent processing steps (e.g., during the drawing steps). Bubbles can be removed using methods such as degassing the spin dope under vacuum, spin dope centrifugation, and spin dope sonication. Some methods to minimize the creation of bubbles in the first place include minimizing the introduction of bubbles by subsurface addition of powder, and removal of the overhead air gap over the spin dope during mixing. Particulates (non-silk particulates, undissolved silk powder) can be removed using methods such as filtering, and centrifugation at high relative centrifugal force (RCF). In some embodiments, the centrifugation is performed from 5000 to 20000 RCF, for a duration from 15 min to 30 min, at a temperature from 4 °C to 22 °C. In some embodiments, the centrifugation is performed at 16000 RCF for a duration greater than 15 min at a temperature from 4 °C to 22 °C, or at 7000 RCF for a duration greater than 15 min at a temperature from 4 °C to 22 °C.

[00115] The flow of dope through the spinneret can be made continuous (i.e., non-pulsatile) for the duration of the spinning event by using a positive displacement pump, rather than a pump that is pulsatile (e.g., a peristaltic pump). Some example of positive displacement pumps are screw pumps, gear pumps, and piston pumps.

[00116] The coagulation bath formulation can be held appreciably constant during the spinning event (e.g., due to preferential evaporation of one of the bath components, water absorption, or chemical reactions between bath components) using methods such as monitoring the bath components with quantitative techniques such as chromatography or spectroscopic techniques, and compensating changes in coagulation bath components using inline adjustment

of the coagulation bath formulation. In some embodiments, the tolerable variation of the concentrations of the components of the coagulation bath formulation are +/- 10% absolute concentration, or +/- 5% absolute concentration. In an example embodiment, for a coagulation bath with a target concentration of 20% formic acid and 80% ethanol, the tolerable range of formic acid concentration can be from 10% to 40% with the remainder of the formulation comprising ethanol, or be from 15% to 25% with the remainder of the formulation comprising ethanol. In some embodiments, the tolerable concentration of absorbed water can be less than 10%, or less than 5%, or less than 2%, or less than 1%.

LONG UNIFORM RECOMBINANT PROTEIN FIBERS

[00117] Embodiments of the present disclosure include RPFs with lengths greater than 20 m. In some embodiments, the RPFs are engineered to comprise various improved mechanical, physical, chemical, and biological properties, as compared to prior art fibers. The long fibers can have uniform mechanical properties, as described by low coefficient of variation (CV) of the mechanical properties along the length of the fibers. Some examples of physical properties of long uniform RPFs are linear density, cross-sectional shape and diameter. Some examples of mechanical properties of of long uniform RPFs are maximum tenacity, initial modulus, extensibility, toughness and work of rupture. Some examples of chemical properties of of long uniform RPFs are moisture absorption, moisture regain and moisture content. An example of biological properties of long uniform RPFs is antimicrobial action.

[00118] When discussing coefficient of variation, enough samples are taken from a fiber, yarn, or textile to sufficiently mitigate low sample bias towards an artificially low CV. In some embodiments, the samples are taken are regular intervals along the length of a fiber, or length of a yarn, or across the area of a textile, in a sufficient quantity to eliminate low sample bias

towards an artificially low CV. In some embodiments, the total number of samples is 10, or 20, or 40, or 60, or 80, or 100, or more than 100.

[00119] In some embodiments, a RPF has improved mechanical properties such as high initial modulus, high extensibility, high tenacity, and high toughness, and one or more of these properties are uniform along the length of the fiber (e.g., with coefficient of variation less than 30%). In some embodiments the RPFs with improved mechanical properties, which are uniform along the length have improved physical properties such as low linear density (low dtex), small diameter, engineered cross-section shapes and low porosity. In some embodiments, the RPFs with improved mechanical properties, which are uniform along the length, have improved chemical properties such as hydrophilicity. In some embodiments, the RPFs with improved mechanical properties, which are uniform along the length have improved biological properties such as being antimicrobial.

[00120] In some embodiments, a long RPF (e.g., with length greater than 20 m) has improved mechanical properties, such as high initial modulus, high extensibility, high tenacity, and high toughness, which are uniform along the length of the fiber (e.g., with coefficient of variation less than 30%). In some embodiments the long RPF (e.g., with lengths greater than 20 m) has improved physical properties, such as high fineness (low dtex), engineered cross-section shapes and porosity, which are uniform along the length of the fiber (e.g., with coefficient of variation less than 30%). In some embodiments, the long RPFs (e.g., with lengths greater than 20 m) have improved chemical properties, such as absorbing moisture effectively (e.g., with a diameter change greater than 10% upon being submerged in water), which are uniform along the length of the fiber (e.g., with coefficient of variation less than 30%). In some embodiments, the long RPFs (e.g., with lengths greater than 20 m) have improved biological properties, such as being

antimicrobial, which are uniform along the length of the fiber (e.g., with coefficient of variation less than 30%).

[00121] In some embodiments, a long RPF has a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m.

Engineering Long Uniform RPF Physical Properties

[00122] In some embodiments, the long uniform RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, and have a mean or median linear density less than 20 dtex, or less than 15 dtex, or less than 10 dtex, or less than 5 dtex, or less than 3 dtex, or less than 2 dtex, or less than 1.5 dtex, or greater than 1.5 dtex, or greater than 1.7 dtex, or greater than 2 dtex, or from 1 to 30 dtex, or from 1 to 20 dtex, or from 1 to 15 dtex, or from 1 to 10 dtex, or from 1 to 5 dtex, or from 1 to 3 dtex, or from 1.5 to 2 dtex, or from 1.5 to

2.5 dtex, or from 0.1 to 30 dtex, or from 0.1 to 20 dtex, or from 0.1 to 10 dtex, or from 0.1 to 5 dtex, or from 0.1 to 3 dtex, or from 0.1 to 2 dtex, and the linear density has a CV along the length of the fiber less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%. Yarns produced from long uniform RPFs are useful in a myriad of applications, such as construction into ropes, textiles and garments, upholstery or linens.

[00123] In some embodiments, the long uniform RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, and have a mean or median diameter less than 100 microns, or less than 75 microns, or less than 50 microns, or less than 25 microns, or less than 20 microns, or less than 15 microns, or less than 10 microns, or less than 5 microns, or less than 2 microns, or greater than 100 microns, or greater than 75 microns, or greater than 50 microns, or greater than 25 microns, or greater than 20 microns, or greater than 15 microns, or greater than 10 microns, or greater than 5 microns, or greater than 1 micron, or from 1 to 100 microns, or from 1 to 75 microns, or from 1 to 50 microns, or from 1 to 25 microns, or from 1 to 20 microns, or from 1 to 15 microns, or from 1 to 10 microns, or from 1 to 5 microns, or from 5 to 100 microns, or from 5 to 75 microns, or from 5 to 50 microns, or from 5 to 25 microns, or

from 5 to 20 microns, or from 5 to 15 microns, or from 5 to 10 microns, and the diameter has a CV along the length of the fiber less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from l% to 10%.

[00124] Microfibers are a classification of fibers having a fineness of less than about 1 dtex (i.e., about 10 μm in diameter). H.K., Kaynak and O. Babaarslan, Woven Fabrics, Croatia: InTech, 2012. In some embodiments, the median or mean linear density of the long uniform RPFs with one or more of the long lengths, mechanical properties, linear density, and low coefficient of variation discussed herein is less than 1 dtex (i.e., about 15 microns in diameter). In some embodiments, the median or mean linear density of the RPFs comprising the filament yarn, or spun yarn, or blended yarn is less than about 0.5 dtex (about 10 microns in diameter). The small diameter of microfibers imparts a range of qualities and characteristics to microfiber yarns and fabrics that are desirable to consumers. Microfibers are inherently more flexible (bending is inversely proportional to fiber diameter) and thus have a soft feel, low stiffness, and high drapeability. Microfibers can also be formed into filament yarns having high fiber density (greater fibers per yarn cross-sectional area), giving microfiber yarns a higher strength compared to other yarns of similar dimensions. Microfibers also contribute to discrete stress relief within the yarn, resulting in anti-wrinkle fabrics. Furthermore, microfibers have high compaction efficiency within the yarn, which improves fabric waterproofness and windproofness while maintaining breathability compared to other waterproofing and windproofing techniques (such as polyvinyl coatings). The high density of fibers within microfiber fabrics results in microchannel structures between fibers, which promotes the capillary effect and imparts a wicking and quick

drying characteristic. The high surface area to volume of microfiber yarns allows for brighter and sharper dyeing, and printed fabrics have clearer and sharper pattern retention as well. Currently, recombinant silk fibers do not have a fineness that is small enough to result in silks having microfiber type characteristics. U.S. Pat. App. Pub. No. 2014/0058066 generally discloses fiber diameters between 5-100 μηι, but does not actually disclose any working examples of any fiber having a diameter as small as 5 μπι.

[00125] In some embodiments, the long uniform RPFs have a longitudinal axis, an inner surface and an outer surface, the inner surface defining a hollow core parallel to the longitudinal axis of the fiber. In some embodiments, the long uniform RPFs have a longitudinal axis and an outer surface, the outer surface including a plurality of corrugations, each corrugation of the plurality parallel or substantially parallel to the longitudinal axis of the fiber. By substantially parallel, we mean an angular deviation between a line defining the longitudinal fiber axis and a line defining the axis of corrugation of less than 25° or less than 20° or less than 15° or less than 10° or less than 5°. In some embodiments, the long uniform RPFs have a longitudinal axis and cross-sectional shape transverse to the longitudinal axis that is substantially circular, or that is substantially triangular, or that is substantially bilobal, or that is substantially trilobal, or that is substantially ovular, or that is substantially c-shaped.

[00126] Surface area to volume ratios are relatively small when the fiber has a smooth surface and a circular cross-section. In some embodiments, the long uniform RPFs have a surface area to volume ratio greater than 1000 cm^{-1}, or from 1000 to $3x10^5$ cm^{-1}, or greater than $lxl\ 0^4$ cm^{-1}, or greater than $lxl\ 0^5$ cm^{-1}. Surface area to volume ratios can be substantially larger when the fiber has a rough surface and/or a non-circular cross-section, for instance if the fiber is striated. In some embodiments, the long uniform RPFs have a surface area to volume ratio from 1000 to

$3x10^7$ cm^{-1}, or greater than $1x10^6$ cm^{-1}, or greater than $1x10^7$ cm^{-1}. Fibers with high surface area to volume ratios could be useful in biomedical applications, filters, and garments.

Engineering Long Uniform RPF Mechanical Properties

[00127] In some embodiments, RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, wherein the mean or median properties of the fibers comprise an initial modulus greater than 50 cN/tex, or greater than 115 cN/tex, or greater than 200 cN/tex, or greater than 400 cN/tex, or greater than 550 cN/tex, or greater than 600 cN/tex, or greater than 800 cN/tex, or greater than 1000 cN/tex, or greater than 2000 cN/tex, or greater than 3000 cN/tex, or greater than 4000 cN/tex, or greater than 5000 cN/tex, or from 200 to 900 cN/tex, or from 100 to 7000 cN/tex, or from 500 to 7000 cN/tex, or from 50 to 7000 cN/tex, or from 100 to 5000 cN/tex, or from 500 to 5000 cN/tex, or from 50 to 5000 cN/tex, or from 100 to 2000 cN/tex, or from 500 to 2000 cN/tex, or from 50 to 2000 cN/tex, or from 100 to 1000 cN/tex, or from 500 to 1000 cN/tex, or from 50 to 1000 cN/tex, or from 50 to 500 cN/tex, or from 100 to 1000 cN/tex, or from 500 to 1000 cN/tex, or from 100 to 700 cN/tex, and/or a maximum tensile strength greater than 0.5 cN/tex, or greater than 1 cN/tex, or greater than 2 cN/tex, or greater than 4 cN/tex, or greater than 6 cN/tex, or greater than 7.7 cN/tex, or greater than 8 cN/tex, or a greater than 10 cN/tex, or greater than 15 cN/tex, or greater than 20 cN/tex, or

greater than 25 cN/tex, or greater than 30 cN/tex, or greater than 40 cN/tex, or greater than 50 cN/tex, or greater than 60 cN/tex, or greater than 70 cN/tex, or greater than 80 cN/tex, or greater than 90 cN/tex, or greater than 100 cN/tex, or from 0.5 cN/tex to 120 cN/tex, or from 1 cN/tex to 120 cN/tex, or from 6 cN/tex to 120 cN/tex, or from 6 cN/tex to 50 cN/tex, or from 6 cN/tex to 30 cN/tex, or from 6 cN/tex to 20 cN/tex, and/or an extensibility greater than 1%, or greater than 3%, or greater than 5%, or greater than 10%, or greater than 20%, or greater than 30%, or greater than 100%, or greater than 200%, or greater than 300%, or greater than 400%, or from 1% to 400%, or from 1 to 200%, or from 1 to 100%, or from 1 to 20%, or from 1 to 30%, or from 1 to 40%, or from 10 to 200%, or from 10 to 100%, or from 10 to 50%, or from 10 to 20%, or from 10% to 20%, or from 50% to 150%, or from 100% to 150%, or from 300% to 400%, and the CV of the initial modulus, maximum tensile strength and/or extensibility along the length of the fiber is less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%. The standard test method for measuring tensile properties of single fibers is ASTM D3822-14. These fiber mechanical properties also enable use of the fibers in industrial fiber drawing and yarn forming methods. Yarns produced from long uniform RPFs are useful in a myriad of applications, such as construction into ropes, textiles and garments, upholstery or linens. Filament yarns, or spun yarns, or blended yarns comprising long uniform RPFs with high modulus, maximum tenacity, and/or extensibility can be used in many applications, including: carpeting and carpet backing, industrial textile products (such as tire cord and tire fabric, seat belts, industrial webbing and tape, tents, fishing line and nets, rope, and tape reinforcement), apparel fabrics (such as women's sheer hosiery, underwear, nightwear, anklets and socks, and a

variety of apparel fabrics), and interior and household products (such as bed ticking, furniture upholstery, curtains, bedspreads, sheets, and draperies).

[00128] In some embodiments, the long uniform RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, wherein the mean or median properties of the fibers comprise a linear density less than 10 dtex, or less than 5 dtex, or less than 3 dtex, or less than 2 dtex, or less than 1.5 dtex, or greater than 1.5 dtex, or greater than 1.7 dtex, or greater than 2 dtex, or from 1 to 15 dtex, or from 1 to 10 dtex, or from 1 to 5 dtex, or from 1 to 3 dtex, or from 1.5 to 2 dtex, or from 1.5 to 2.5 dtex, and a maximum tensile strength greater than 0.5 cN/tex, or greater than 1 cN/tex, or greater than 2 cN/tex, or greater than 4 cN/tex, or greater than 6 cN/tex, or greater than 7.7 cN/tex, or greater than 8 cN/tex, or a greater than 10 cN/tex, or greater than 15 cN/tex, or greater than 20 cN/tex, or greater than 25 cN/tex, or greater than 30 cN/tex, or greater than 40 cN/tex, or greater than 50 cN/tex, or greater than 60 cN/tex, or greater than 70 cN/tex, or greater than 80 cN/tex, or greater than 90 cN/tex, or greater than 100 cN/tex, or from 0.5 cN/tex to 120 cN/tex, or from 1 cN/tex to 120 cN/tex, or from 6 cN/tex to 120 cN/tex, or from 6 cN/tex to 50 cN/tex, or from 6 cN/tex to 30 cN/tex, or from 6 cN/tex to 20 cN/tex, and the CV of the linear density and/or the maximum tensile strength along the length of the fiber is less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or

from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%.

[00129] In some embodiments, RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, wherein the mean or median properties of the fibers comprise a work of rupture greater than 0.1 cN*cm, or greater than 0.2 cN*cm, or greater than 0.3 cN*cm, or greater than 0.4 cN*cm, or greater than 0.5 cN*cm, or greater than 0.6 cN*cm, or greater than 0.7 cN*cm, or greater than 0.8 cN*cm, or greater than 0.9 cN*cm, or greater than 1 cN*cm, or greater than 1.3 cN*cm, or greater than 2 cN*cm, or greater than 5 cN*cm, or greater than 10 cN*cm, or from 0.1 to 10 cN*cm, or from 0.1 to 5 cN*cm, or from 0.1 to 2 cN*cm, or from 0.2 to 5 cN*cm, or from 0.2 to 10 cN*cm, or from 0.2 to 2 cN*cm, or from 0.3 to 2 cN*cm, or from 0.4 to 10 cN*cm, or from 0.4 to 5 cN*cm, or from 0.4 to 2 cN*cm, or from 0.4 to 1 cN*cm, or from 0.5 to 2 cN*cm, or from 0.5 to 1.3 cN*cm, 0.6 to 2 cN*cm, or from 0.7 to 1.1 cN*cm, and the CV of the work or rupture along the length of the fiber is less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%.

[00130] In some embodiments, RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, wherein the mean or median properties of the fibers comprise a toughness greater than 2 cN/tex, or from 0.5 to 70 cN/tex, or greater than 3 cN/tex, or greater than 4 cN/tex, or greater than 5 cN/tex, or greater than 7.5 cN/tex, or greater than 10 cN/tex, or greater than 20 cN/tex, or greater than 30 cN/tex, or greater than 40 cN/tex, or greater than 50 cN/tex, greater than 60 cN/tex, or greater than 70 cN/tex, or from 2 to 3 cN/tex, or from 3 to 4 cN/tex, or from 4 to 5 cN/tex, or from 5 to 7.5 cN/tex, or from 7.5 to 10 cN/tex, or from 10 to 20 cN/tex, or from 20 to 30 cN/tex, or from 30 to 40 cN/tex, or from 40 to 50 cN/tex, or from 50 to 60 cN/tex, or from 60 to 70 cN/tex, and the CV of the toughness along the length of the fiber is less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%. Filament yarns, or spun yarns, or blended yarns comprising long uniform RPFs with high toughness can be used in many applications, including: carpeting and carpet backing, industrial textile products (such as tire cord and tire fabric, seat belts, industrial webbing and tape, tents, fishing line and nets, rope, and tape reinforcement), apparel fabrics (such as women's sheer hosiery, underwear, nightwear, anklets and socks, and a variety of apparel fabrics), interior

and household products (such as bed ticking, furniture upholstery, curtains, bedspreads, sheets, and draperies).

Engineering Long Uniform RPF Moisture Properties

[00131] Another moisture-related characteristic of a fiber is the degree of swelling when submerged in water. In some embodiments, the long uniform RPF have high moisture absorption properties. In some embodiments, RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, wherein the mean or median properties of the fibers comprise diameter change upon being submerged in water at a temperature of 21 °C +/- 1 °C of greater than 5%, or from 0.1% to 100%, or greater than 1%, or greater than 2%, or greater than 4%, or greater than 6%, or greater than 8%, or greater than 10%, or greater than 15%, or greater than 20%, or greater than 25%, or greater than 30%, or greater than 35%, or greater than 40%, or greater than 45%, or greater than 50%, or greater than 60%, or greater than 70%, or greater than 80%, or greater than 90%, or from 5% to 10%, or from 10% to 20%, or from 20% to 30%, or from 30% to 40%, or from 40% to 50%, or from 50% to 60%, or from 60% to 70%, or from 70% to 80%, or from 80% and 90%, or from 90% to 100%, or from 20% to 35%, or from 15% to 40%, or from 15% to 35%, and the CV of the diameter change upon being submerged in water at a temperature of 21 °C +/- 1 °C along the length of the fiber is

less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%. Such long uniform RPFs can be made into filament yarns, spun yarns or blended yarns, and such filament yarns, spun yarns, or blended yarns are useful in textiles and garments such as skin knits or woven fabrics where transfer of moisture away from the skin is desired, such as active wear apparel. In some embodiments, these filament yarns, or spun yarns, or blended yarns can be constructed into plaited yarn or textile, or double knit textiles. In some embodiments, these textiles can be located in a position towards the outer surface of a textile and/or garment to allow the absorbed moisture to easily evaporate. Fiber diameter change can be directly measured using optical microscopy.

[00132] Two other moisture-related characteristics of fibers are moisture regain and moisture content, which measure the uptake of water vapor from the environment. In one type of measurement a sample is allowed to equilibrate in an environment with a known relative humidity (e.g., 60-70% relative humidity) and temperature (e.g., 20-25 °C), and then heated to drive out the water (e.g., at a temperature slightly above 100 °C). Using a tool, such as a thermogravimetric analysis (TGA) system, the initial conditioned mass (containing some water), the final dry mass, and the mass change can be measured over time. The moisture regain of the fiber is defined as the lost water mass divided by the dry mass. The moisture content of the RPF is defined as the lost water mass divided by the conditioned mass. In some embodiments, RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than

5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, wherein the mean or median properties of the fibers comprise a moisture regain or moisture content, when measured from equilibrium conditioned mass at 65% relative humidity environment at 22 °C and heated at 110 °C until approximately equilibrium dry mass is achieved, of greater than 1%, or greater than 2%, or greater than 3% or greater than 4%, or greater than 5%, or greater than 6%, or greater than 7%, or greater than 8%, or greater than 9%, or greater than 10%, or greater than 12%, or greater than 14%, or greater than 16%, or greater than 18%, or greater than 20%, or from 1% to 30%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%, or from 5% to 15%, or from 5% to 10%, and the CV of the moisture regain or moisture content along the length of the fiber is less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%.

[00133] There are many different metrics by which to characterize the interaction between a fiber and water. One such method is measuring the hydrophilicity of the surface of the fiber, characterized by the contact angle with water. In some embodiments, the long uniform RPF when measured with a fiber tensiometer, have a median or mean tensiometer contact angle of less than 90 degrees, or less than 80 degrees, or less than 70 degrees, or less than 60 degrees, or between 60 and 90 degrees or 60 and 80 degrees, or from 60 and 70 degrees, or from 70 and 90 degrees, or from 70 and 80 degrees, or from 80 and 90 degrees when tested using a standard

assay with a water-filled tensiometer, and the CV of the tensiometer contact angle along the length of the fiber is less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%. Such long uniform RPFs can be made into yarns, and such yarns are useful in textiles which use fiber properties and yarn constructions used to pull moisture away from the skin in order to create more comfort for the wearer. In some embodiments, these filament yarns, or spun yarns, or blended yarns can be constructed into plaited yarn or textile, or double knit textiles. In some embodiments, these textiles can be located in a position towards the outer surface of a textile and/or garment to allow the absorbed moisture to easily evaporate.

[00134] In some embodiments, the long uniform RPFs have high moisture wicking properties. A standard method of measuring wicking rate is the AATCC test method 197-2011 for vertical wicking of textiles, and AATCC test method 198-2011 for horizontal wicking of textiles. In some embodiments, a plain weave 1/1 textile with warp density of 72 warps/cm and pick density of 40 picks/cm, comprising filament yarn, or spun yarn, or blended yarn, comprising long uniform RPFs, is tested using AATCC test method 197-2011, and has a median or mean horizontal wicking rate greater than 1 mm/s, or from 0.1 to 100 mm/s, or greater than 0.1 mm/s, or greater than 0.2 mm/s, or greater than 0.4 mm/s, or greater than 0.6 mm/s, or greater than 0.8 mm/s, or greater than 2 mm/s, or greater than 4 mm/s, or greater than 6 mm/s, or greater than 8 mm/s, or greater than 10 mm/s, or greater than 15 mm/s, or greater than 20 mm/s, or greater than 40 mm/s, or greater than 60 mm/s, or greater than 80 mm/s, or greater than 100 mm/s, or from 0.1 mm/s to 1 mm/s, or from 1 mm/s to 10 mm/s, or from 10 mm/s to 20 mm/s, or from 20 mm/s to 30 mm/s, or from 30 mm/s to 40 mm/s, or from 40 mm/s to 50 mm/s, or from 50 mm/s to 60

mm/s, or from 60 mm/s to 70 mm/s, or from 70 mm/s to 80 mm/s, or from 80 mm/s to 90 mm/s, or from 90 mm/s to 100 mm/s, and the CV of the horizontal wicking rate across the area of the textile is less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to 15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%. Such filament yarns, or spun yarns, or blended yarns containing long uniform RPFs are useful in textiles and garments such as skin knits or woven fabrics where wicking of moisture away from the skin is desired, such as active wear apparel. In some embodiments, these filament yarns, or spun yarns, or blended yarns can be constructed into plaited yarn or textile, or double knit textiles. In some embodiments, these textiles are located in a position towards the outer surface of a textile and/or garment to allow the absorbed moisture to easily evaporate.

Combinations of Long Uniform RPF Properties

[00135] In different embodiments, fibers, yarns and textiles characteristics can be grouped together. For example, fibers can be engineered to have high moisture absorption and have high extensibility. In fact, all of the fibers, yarns and textiles properties discussed in this disclosure can be combined with each other. However, in some cases the quantification of the fibers, yarns or textiles property and the method by which the property is obtained, are both important, and may change which properties can be combined. For example, moisture absorption can be imparted to the fibers by increasing the ratio of poly-alanine to glycine-rich regions in the protein sequence, however, increasing the ratio of poly-alanine regions in the protein sequence tends to the make the fiber less extensible. Table 2 illustrates combinations of fibers, yarns and

textiles properties that are not mutually exclusive (Y), and fibers properties that are mutually exclusive (N).

Table 2: Fibers, yarns and textiles properties, viable combinations

	moisture absorption	wickability	antimicrobial	extensibility	tenacity	initial modulus	toughness	cross-section	linear density (or diameter)
moisture absorption		Y	Y	Y	Y	Y	Y	Y	Y
wickability			Y	Y	Y	Y	Y	Y	Y
antimicrobial				Y	Y	Y	Y	Y	Y
extensibility					Y	Y	Y	Y	Y
tenacity						Y	Y	Y	Y
initial modulus							Y	Y	Y
toughness								Y	Y
cross-section									Y
linear density (or diameter)									

[00136] One example of a combination of physical, mechanical and moisture properties of long uniform RPFs is linear density, maximum tensile strength and diameter change upon being submerged in water. In some embodiments, the long uniform RPFs have a length greater than 20 m, or greater than 50 m, or greater than 100 m, or greater than 200 m, or greater than 300 m, or greater than 400 m, or greater than 500 m, or greater than 750 m, or greater than 1000 m, or greater than 1500 m, or greater than 2000 m, or greater than 5000 m, or greater than 10000 m, or from 20 to 2000 m, or from 50 to 2000 m, or from 100 to 2000 m, or from 200 to 2000 m, or from 500 to 2000 m, or from 20 to 5000 m, or from 50 to 5000 m, or from 100 to 5000 m, or from 200 to 5000 m, or from 500 to 5000 m, or from 20 to 10000 m, or from 50 to 10000 m, or from 100 to 10000 m, or from 200 to 10000 m, or from 500 to 10000 m, wherein the mean or

median properties of the fibers comprise a linear density less than 10 dtex, or less than 5 dtex, or less than 3 dtex, or less than 2 dtex, or less than 1.5 dtex, or greater than 1.5 dtex, or greater than 1.7 dtex, or greater than 2 dtex, or from 1 to 15 dtex, or from 1 to 10 dtex, or from 1 to 5 dtex, or from 1 to 3 dtex, or from 1.5 to 2 dtex, or from 1.5 to 2.5 dtex, and a maximum tensile strength greater than 0.5 cN/tex, or greater than 1 cN/tex, or greater than 2 cN/tex, or greater than 4 cN/tex, or greater than 6 cN/tex, or greater than 7.7 cN/tex, or greater than 8 cN/tex, or a greater than 10 cN/tex, or greater than 15 cN/tex, or greater than 20 cN/tex, or greater than 25 cN/tex, or greater than 30 cN/tex, or greater than 40 cN/tex, or greater than 50 cN/tex, or greater than 60 cN/tex, or greater than 70 cN/tex, or greater than 80 cN/tex, or greater than 90 cN/tex, or greater than 100 cN/tex, or from 0.5 cN/tex to 120 cN/tex, or from 1 cN/tex to 120 cN/tex, or from 6 cN/tex to 120 cN/tex, or from 6 cN/tex to 50 cN/tex, or from 6 cN/tex to 30 cN/tex, or from 6 cN/tex to 20 cN/tex, and a diameter change upon being submerged in water at a temperature of 21 °C +/- 1 °C of greater than 5%, or from 0.1% to 100%, or greater than 1%, or greater than 2%, or greater than 4%, or greater than 6%, or greater than 8%, or greater than 10%, or greater than 15%, or greater than 20%, or greater than 25%, or greater than 30%, or greater than 35%, or greater than 40%, or greater than 45%, or greater than 50%, or greater than 60%, or greater than 70%, or greater than 80%, or greater than 90%, or from 5% to 10%, or from 10% to 20%, or from 20% to 30%, or from 30% to 40%, or from 40% to 50%, or from 50% to 60%, or from 60% to 70%, or from 70% to 80%, or from 80% and 90%, or from 90% to 100%, or from 20% to 35%, or from 15% to 40%, or from 15% to 35%, and the CV of the linear density and/or the maximum tensile strength and/or diameter change upon being submerged in water at a temperature of 21 °C +/- 1 °C along the length of the fiber is less than 50%, or less than 40%, or less than 30%, or less than 20%, or less than 15%, or less than 10%, or less than 5%, or from 0.1% to 50%, or from 0.1% to 40%, or from 0.1% to 30%, or from 0.1% to 20%, or from 0.1% to

15%, or from 0.1% to 10%, or from 1% to 50%, or from 1% to 40%, or from 1% to 30%, or from 1% to 20%, or from 1% to 15%, or from 1% to 10%.

METHODS OF FORMING RECOMBINANT PROTEIN FIBER YARNS AND TEXTILES

[00137] Individual RPFs are made into yarns to be used in textiles. There are different methods of forming yarns from ling uniform RPFs and there are different methods of forming textiles from yarns comprising long uniform RPFs, which produce yarns and textiles with different structures and properties.

[00138] Depending on the type of yarn desired, several filament yarn forming methods can be used to make filament yarns containing long uniform RPFs. These methods may include simple twisting of flat filament fibers using a silk throwing apparatus or continuous spinning. Textured filament yarns comprising long uniform RPFs can be further subjected to processes that arrange the straight filaments into crimped, coiled or looped filaments to create bulk, texture or stretch. Some examples of methods used for processing textured filament yarns comprising long uniform RPFs are air jet texturing, false twist texturing, or stuffer box texturing. Filament yarns may also be texturized during the spinning using false twist texturizing, air jet texturizing or stuffer box apparatus. Heating, chemically bonding or plying may also be employed.

[00139] In some embodiments, the yarns comprising long uniform RPFs are manufactured using a ring spinning apparatus. In some embodiments, the yarns comprising long uniform RPFs are manufactured using an open end spinning apparatus. In some embodiments, the yarns comprising long uniform RPFs are manufactured using an air-jet spinning apparatus. In certain embodiments, twist is applied resulting in a twist angle optimized for desired mechanical, structural or other properties of the yarn. In certain embodiments, the twist applied to the inner core of the yarn has a different twist angle compared with the outer skin of the yarn. Throughout

this disclosure "spun" yarns can refer to ring spun yarns, open end spun yarns, air-jet spun yarns, vortex spun yarns, or any other method of producing a yarn where the yarn comprises staple fibers.

[00140] In some embodiments, the blended yarn comprising long uniform RPFs and/or non-RPFs is manufactured by spinning. The structure of a spun yarn is influenced by the spinning methods parameters. The properties of the spun yarn are influenced by the structure of the yarn, as well as the constituent fibers. In embodiments, the blended yarn structure and the long uniform RPFs properties and the type of non-RPFs blended with the RPFs are all chosen to impart various characteristics to the resulting yarns. In some embodiments, the blended yarns are manufactured using a ring spinning apparatus. In some embodiments, the blended yarns are manufactured using an open end spinning apparatus. In some embodiments, the blended yarns are manufactured using an air-jet spinning apparatus. In many embodiments, twist is applied of a certain twist angle to optimize the mechanical properties of the blended yarn. In many embodiments, the twist applied to the inner core of the yarn has a different twist angle compared with the outer skin of the blended yarn.

[00141] In some embodiments, a method of making a spun yarn is employed, wherein a plurality of RPFs is provided, the fibers are cut into staple, the fibers are conveyed the fibers to a spinning apparatus, and twist is provided to spin the fibers into a yarn. In some embodiments, the spinning apparatus is a ring spinning apparatus. In some embodiments, the spinning apparatus is an open end spinning apparatus. In some embodiments, the spinning apparatus is an air jet spinning apparatus. In some embodiments, the fibers are carded prior to spinning. In some embodiments, the fibers are combed prior to spinning.

[00142] In some embodiments, a method of making a blended spun yarn is employed, wherein a plurality of long uniform RPFs and non-RPFs is provided, the fibers are cut into staple, the

fibers are loaded in to a spinning apparatus, and twist is provided to spin the fibers into a yarn. In some embodiments, the spinning apparatus is a ring spinning apparatus. In some embodiments, the spinning apparatus is an open end spinning apparatus. In some embodiments, the spinning apparatus is an air jet spinning apparatus. In some embodiments, the fibers are carded prior to spinning. In some embodiments, the fibers are combed prior to spinning.

[00143] In some embodiments, the yarns comprising long uniform RPFs are manufactured into textiles, for example by weaving or knitting. In some embodiments, yarns containing long uniform RPFs are manufactured into textiles by knitting using a circular knitting apparatus, a warp knitting apparatus, a flat knitting apparatus, a one piece knitting apparatus, or a 3-D knitting apparatus. In some embodiments, yarns containing long uniform RPFs are manufactured into textiles by weaving using a plain weave loom, a dobby loom or a jacquard loom. In some embodiments, long uniform RPFs are manufactured into textiles using a 3d printing method. In some embodiments, long uniform RPFs or yarns containing long uniform RPFs are manufactured into non-woven textiles using techniques such as wet laying, spin bonding, stitch bonding, spunlacing (i.e., hydroentaglement), or needlepunching. In embodiments, the textile construction, the yarn structure and the long uniform RPFs properties are chosen to impart various characteristics to the resulting yarns and textiles.

EXAMPLES

EXAMPLE 1: RECOMBINANT PROTEIN FIBER SPINNING

[00144] Copolymers in this example were secreted from *Pichia pastoris* commonly used for the expression of recombinant DNA using published techniques, i.e., those described in WO2015042164 A2, at paragraphs 114-134. The copolymer polypeptide utilized for fiber

spinning in this Example was SEQ ID NO. 1, concatenated 3 times, with a 3X FLAG sequence: GDYKDDDDKDYKDDDDKDYKDDDDK (SEQ ID NO: 98) bound to the C-terminal end of the polypeptide. The secreted proteins were purified and dried using standard techniques. The dried polypeptide powder was dissolved in a formic acid-based spinning solvent to generate a homogenous spin dope.

[00145] The RPFs in this example were spun by extruding the spin dope through a 125 μιη diameter orifice with 1:1 ratio of length to diameter into a room temperature alcohol-based coagulation bath comprising 20% formic acid with a residence time of approximately 15 seconds. Fibers were pulled out of the coagulation bath under tension, and then drawn to three times their length, and subsequently allowed to dry. The volumetric flow rate out of spinneret was approximately 45 uL/min. The duration of the spinning event (i.e., spin run) was approximately 10 minutes. The volume of spin dope consumed in the spinning event was approximately 0.45 mL. The length of fibers produced in a single spinning event was approximately 200 m.

EXAMPLE 2: LONG UNIFORM RECOMBINANT PROTEIN FIBERS MECHANICAL PROPERTIES

[00146] Figs. 2 and 3 show tenacity and linear density data from 1800 m of RPF. The fiber in this Example was produced by the methods described in Example 1. Two different batches of spin dopes were used, in 9 separate spin runs of 200 m each. 23-25 samples were collected at regular intervals along the length of the 200 m fiber from a given spin run and measured for tenacity and linear density.

[00147] To determine the mean, standard deviation and coefficient of variation (CV) for each parameter, a Monte Carlo method was used, where the data was randomly grouped into sets of 25 data points 440 different times. The resulting data was plotted in Figs. 2 and 3. Fig. 2 shows

the distribution of means, standard deviations and CVs of each group of 25 for maximum tenacity. Fig. 3 shows the distribution of means, standard deviations and CVs of each group of 25 for linear density.

[00148] Using this approach, the average mean for maximum tenacity was 9.24 cN/tex, the average standard deviation was 0.50 cN/tex, and the average CV was 5.40%. It can also be observed from the data that the mean maximum tenacity from all 440 randomized groups of 25 measurements was from 9 cN/tex to 9.5 cN/tex. Similarly, the standard deviation of the maximum tenacity for the 440 randomized groups of 25 measurements was from 0.3 cN/tex to 0.7 cN/tex, and the CV of the maximum tenacity for the 440 randomized groups of 25 measurements was from 3.5% to 7.7%.

[00149] Using the above approach, the average mean for linear density was 8.84 dtex, the average standard deviation was 0.93 dtex, and the average CV was 10.54%. It can also be observed from the data that the mean linear density from all 440 randomized groups of 25 measurements was from 8.3 dtex to 9.3 dtex. Similarly, the standard deviation of the linear density for the 440 randomized groups of 25 measurements was from 0.6 dtex to 1.4 dtex, and the CV of the linear density for the 440 randomized groups of 25 measurements was from 6.5% to 16%.

[00150] Stated another way, due to sampling statistics and depending on the groupings of the 25 data points within the 1800 m of fiber, the average measured parameters can vary. However, using proper statistics the values of the mean parameters can be determined to a high degree of confidence. For instance, for the long uniform RPF data in this Example, 95% of the time the CV for tenacity was from 4.15% and 6.90%, and 95% of the time the CV for the linear density was from 7.45% to 13.72%.

[00151] This Example illustrates that the process described in this disclosure to produce the long uniform fibers was robust and reproducible, since multiple spin dopes were used for the different spinning events, and the average CV for linear density (a physical property of the fibers) and the average CV for tenacity (a mechanical property of the fibers) were about 10.4% and 5.4%, respectively.

EXAMPLE 3: LONG UNIFORM RECOMBINANT PROTEIN FIBERS MECHANICAL PROPERTIES

[00152] The spin dope in this Example was produced by the methods described in Example 1. The RPFs in this example were spun by extruding the spin dope through a 50-hole spinneret in which each orifice was 127 μm diameter with 1:1 ratio of length to diameter into a room temperature alcohol-based coagulation bath comprising 20% formic acid with a residence time of approximately 28 seconds. Fibers were pulled out of the coagulation bath under tension, and then drawn to four times their length in a pure alcohol wash bath. The fibers were subsequently passed through a 200 °C oven and collected onto a spool under tension. The volumetric flow rate out of spinneret was approximately 0.8 mL/min. The duration of the spinning event was approximately 30 minutes. The total volume of material consumed in the spinning event was approximately 23 mL.

[00153] FIG. 4 shows stress strain curves of 23 fibers with compositions described herein, and produced by the methods described above. The fibers in this Example have maximum tensile stress greater than 20 cN/tex, and the average of the maximum tensile stresses of the 23 fibers was about 18.6 cN/tex. This set of fibers was sampled from a 50-filament tow that was 800 m long, and therefore produced 40,000 m of fiber in a single spinning event. The maximum tensile stress ranges from about 17 to 21 cN/tex, and the standard deviation of the maximum tensile stress in this example was about 1.0 cN/tex. The average initial modulus of the 23 fibers was

about 575 cN/tex, and the standard deviation in this example was about 6.7 cN/tex. The average maximum elongation of the 23 fibers was about 10.2%, and the standard deviation in this example was about 3.6%. The average linear density of the 23 fibers was about 3.1 dtex, and the standard deviation in this example was about 0.11 dtex.

[00154] For this specific spinning event, and the conditions and the batch of spin dope material used, the coefficient of variation of the maximum tensile strength for this set of fibers was about 5.4%, the coefficient of variation of the initial modulus for this set of fibers was about 1.2%, and the coefficient of variation of the linear density for this set of fibers was about 3.5%.

[00155] This Example illustrates that a long uniform RFP with tenacity above 20 cN/tex was produced using the methods described in this disclosure.

EXAMPLE 4: PURITY OF RECOMBINANT 18B POLYPEPTIDE POWDER

[00156] Copolymers in this example were produced and dried into powder ("18B powder") as discussed above with respect to Example 1. Data characterizing the relative amounts of high, low and intermediate molecular weight impurities as compared with monomeric 18B and aggregate 18B (i.e., proteinaceous block copolymers) was collected using Size Exclusion Chromatography. 18B powder was dissolved in 5M Guanidine Thiocyanate and injected onto a Yarra SEC-3000 SEC-HPLC column to separate constituents on the basis of molecular weight. Refractive index was used as the detection modality. 18B aggregates, 18B monomer, low molecular weight (1-8 kDa) impurities, intermediate molecular weight impurities (8-50 kDa) and high molecular weight impurities (110-150 kDa) were quantified. Relevant composition was reported as mass % and area%. BSA was used as a general protein standard with the assumption that >90% of all proteins demonstrate dn/dc values (the response factor of refractive index)

within -7% of each other. Poly(ethylene oxide) was used as a retention time standard, and a BSA calibrator was used as a check standard to ensure consistent performance of the method.

[00157] Table 3 (below) lists the area% of the different components quantified using SEC and the mass% of 18B in its aggregate and monomeric forms. As shown, the overall purity by mass of the 18B powder was 59.86%

Table 3: Size Exclusion Chromatography

Source Ref	High MW Impurity [area %]	IMW Impurity [area %]	LMW Impurity [area%]
148	1.98	31.73	6.10

Source Ref	18B Aggregate [area %]	18B Aggregate [mass %]	18B Monomer [area %]	18B Monomer [mass %]	18B Aggregate + Monomer [area%]	18B Aggregate + Monomer [mass%]
148	6.07	6.04	54.11	53.82	60.18	59.86

EXAMPLE 5: SPIN DOPE PREPARATION AND RHEOLOGY OF THE SPIN DOPES

[00158] The 18B powder were dissolved in formic acid and mixed using a Thinky Planetary Centrifugal Mixer 400ARE-TWIN at 1600 RPM to generate spin dopes. Prior to dissolution, the 18B powder was baked to reduce the moisture content of the powder down to less than 4%.

[00159] A Malvern Kinexus Lab+ Rotational Rheometer was to measure the complex viscosity and the phase angle of the spin dopes. Parameters were set to a temperature of 22°C, a frequency of 100-0. 1 Hz, and a strain of 1%. An interval of 3 points/decade was used to determine an average value for a given frequency.

[00160] The table below includes the concentration by weight of the 18B powder in the spin dope, the complex viscosity and the phase angle as measured at 10 Hz. Data was not collected for 125-FACU.

Table 4: 18B Powder Concentration in Spin Dope, Complex Viscosity and Phase Angle

Source ID	18B Powder Concentration by weight [%]	Complex Viscosity [Pa s]	Phase Angle [o]
144	36	42.67	65.26

EXAMPLE 6: DRAWING CONDITIONS

[00161] The 18B protein powder referenced in Example 4 was wet-spun into fiber using traditional techniques. **A** spin dope was prepared using 67% formic acid (by weight) and 33% 18B powder (by weight). The spin dope was mixed using a FlackTek SpeedMixer **DAC** 600.2 **VAC-LR** vacuum mixer.

[00162] The spin dope was extruded directly into a coagulation bath comprised of 100% ethanol at room temperature through a spinneret that is 50μm in diameter at a rate of 1.25 ml/minute to form a precursor fiber. Both the spinneret and the coagulation bath were maintained at room temperature. Precursor fiber is then collected on a set of uptake godets at a reel rate of 3.2 meters/minute. The precursor fiber was then drawn between the uptake godets and a heated godet spaced 81 inches apart. The reel rate of the heated godet was 19.8 meters/minute, providing a draw ratio of 6.19X. The drawn fiber was then drawn between the heated godet and a final godet that were spaced 139 inches apart. The uptake rate of the final godet was 22 meters/minute providing a draw ration of 1.12X.

[00163] Between the heated godet and the final godet, the drawn fiber was passed through a 40-inch tube furnace that was heated to **200°C. A** lubricant comprising 200% proof ethanol at 99% by weight and Setol® was applied to the drawn, heat-treated fiber at a rate of 1.1 mL/minute. The fiber was then wound on a spool for further analysis.

EXAMPLE 7: LONG UNIFORM RECOMBINANT PROTEIN FIBERS MECHANICAL PROPERTIES

[00164] Data characterizing tensile properties and linear density was collected from an approximately 700M spool (plus or minus 50M) of fiber produced in a single spin run according to the methods described in Examples 5 and 6. To produce the data, the fiber was divided into sixteen (16) segments of approximately 50m each representing different regions of the spool. 11 of the 16 segments representing a random sample of the length of the spool were tested using a FAVIMAT fiber tensile test equipment model Favimat+ and Robot2 using 5 tows per segment and 4 filaments per tow to produce a total of 220 samples, from which 12 outliers were subtracted to produce 208 samples. Linear density was tested in accordance with ASTM D1577. Tensile properties were tested in accordance with ASTM D3822-14. Table 5 below lists the properties calculated the 208 samples taken from the 11 segments.

Table 5 - Data collected from all Samples

	Mean	**Std Dev**	**Coefficient of Variance ("CV")%**	**Min**	**Max**	**N**
Tenacity (cN/tex)	12.75	1.69	13.28	7.96	18.14	208.00
Linear. Den. (dtex)	5.98	1.04	17.40	2.90	8.56	208.00
Initial Modulus (cN/tex)	488.88	22.99	4.70	400.97	571.33	208.00
Elongation (%)	24.88	10.12	40.70	2.78	46.69	208.00
Elongation at Break (%)	26.15	8.67	33.15	3.91	47.11	208.00
Work of rupture (cN*cm)	3.40	1.68	49.26	0.13	7.26	208.00
Force at rupture (cN)	7.62	1.73	22.71	3.09	12.56	208.00

Table 6 - Data by segment

	Tenacity (cN/tex)			Linear Den. (dtex)			Initial Modulus (cN/tex)			Elongation (%)		
Segment	**Mean**	**Std Dev**	**CV%**	**Mean**	**Std Dev**	**CV%**	**Mean**	**Std Dev**	**CV%**	**Mean**	**Std Dev**	**CV%**
1	12.66	1.71	13.51	5.92	0.94	15.79	492.24	21.47	4.36	25.27	10.03	39.68
2	12.38	2.01	16.27	6.00	1.40	23.36	483.22	23.08	4.78	27.23	11.33	41.59
4	13.02	1.48	11.40	5.34	1.26	23.49	494.08	22.81	4.62	23.99	5.68	23.67
6	12.52	1.39	11.11	5.87	0.84	14.28	483.86	24.33	5.03	26.42	8.15	30.84
8	12.48	1.20	9.59	6.45	1.06	16.48	487.68	14.89	3.05	23.45	11.62	49.55
9	12.93	1.51	11.69	6.22	0.88	14.12	498.16	15.20	3.05	20.73	13.79	66.52
11	13.17	1.43	10.84	5.84	0.82	13.98	501.40	16.81	3.35	21.90	8.48	38.72
12	14.11	1.99	14.07	6.11	1.04	17.00	507.46	29.76	5.86	24.74	7.29	29.45
14	12.70	1.84	14.47	5.98	1.16	19.40	477.68	24.57	5.14	28.33	8.07	28.49
15	12.56	1.72	13.66	6.33	0.76	12.07	477.82	17.02	3.56	24.00	13.94	58.08
16	11.79	1.55	13.19	5.87	0.92	15.64	476.86	17.92	3.76	26.50	10.97	41.38

Table 7 - Data by segment

	Elongation at Break (%)			Work of rupture (cN*cm)			Force at rupture(cN)		
Region	Mean	Std Dev	CV%	Mean	Std Dev	CV%	Mean	Std Dev	CV%
1	25.62	9.88	38.59	3.28	1.36	41.43	7.46	1.46	19.55
2	27.93	10.58	37.90	3.65	1.99	54.57	7.49	2.21	29.46
4	24.30	5.66	23.28	2.92	0.99	33.81	6.91	1.64	23.76
6	26.72	8.08	30.25	3.43	1.37	39.88	7.36	1.53	20.74
8	25.51	8.92	34.98	3.57	2.04	57.02	8.08	1.72	21.31
9	26.65	8.29	31.13	3.08	2.35	76.32	8.06	1.62	20.07
11	22.68	7.57	33.39	3.05	1.29	42.28	7.68	1.38	17.94
12	25.00	7.28	29.13	3.87	1.67	43.11	8.66	2.13	24.61
14	28.63	8.09	28.26	3.84	1.63	42.35	7.52	1.63	21.68
15	25.88	12.15	46.93	3.40	2.10	61.74	7.96	1.61	20.19
16	28.53	7.67	26.90	3.32	1.57	47.18	6.94	1.64	23.65

ADDITIONAL CONSIDERATIONS

[00165] The foregoing description of the embodiments of the disclosure has been presented for the purpose of illustration; it is not intended to be exhaustive or to limit the claims to the precise forms disclosed. Persons skilled in the relevant art can appreciate that many modifications and variations are possible in light of the above disclosure.

[00166] The language used in the specification has been principally selected for readability and instructional purposes, and it may not have been selected to delineate or circumscribe the inventive subject matter. It is therefore intended that the scope of the disclosure be limited not by this detailed description, but rather by any claims that issue on an application based hereon. Accordingly, the disclosure of the embodiments is intended to be illustrative, but not limiting, of the scope of the invention, which is set forth in the following claims.

What is claimed is:

1. A long uniform recombinant protein fiber, comprising:

a continuous fiber length of at least 600 m, wherein the mean properties of the fiber comprise:

a tenacity greater than or equal to 12 cN/tex;

a linear density less than or equal to 6 dtex;

a coefficient of variation of tenacity less than 15% along the length; and

a coefficient of variation of linear density less than 20% along the length,

wherein the tenacity is measured using ASTM D3822-14, and the linear density is measured using ASTM D1577.

2. The recombinant protein fiber of claim 1, wherein the length is at least 50 m.

3. The recombinant protein fiber of claim 1, wherein the length is at least 650 m.

4. The recombinant protein fiber of any of claims 1-3, wherein the tenacity has a coefficient of variation less than 10% along the length.

5. The recombinant protein fiber of any of claims 1-4, wherein the linear density has a coefficient of variation less than 15% along the length.

6. The recombinant protein fiber of any of claims 1-5, wherein the mean elongation at break is greater than 25% and the elongation at break has a coefficient of variation of less than 35% along the length. .

7. The recombinant protein fiber of any of claims 1-6, wherein the mean initial modulus is greater than 480 cN/tex and the initial modulus has a coefficient of variation of less than 5% along the length.

8. The recombinant protein fiber of any of claims 1-7, wherein the mean elongation is greater than 24% and the elongation has a coefficient of variation of less than 45% along the length.

9. The recombinant protein fiber of any of claims 1-8, wherein the mean work of rupture is greater than 3 cN * cm and the work of rupture has a coefficient of variation of less than 50% along the length.

10. The recombinant protein fiber of any of claims 1-9, wherein the mean force at rupture is greater than 7 cN and the force at rupture has a coefficient of variation less than 25% along the length.

11. The recombinant protein fiber of any of claims 1-10, wherein the recombinant protein fiber is produced by wet spinning a dope comprising a recombinant protein powder.

12. The recombinant protein fiber of claim 11, wherein the recombinant protein powder is less than 65% proteinaceous block copolymer by mass.

13. The recombinant protein fiber of any of claims 1-12, wherein the recombinant protein fiber comprises a protein sequence comprising repeat units, wherein

each repeat unit has at least 95% sequence identity to a sequence that comprises from 2 to 20 quasi-repeat units,

each quasi-repeat unit having a composition comprising {GGY-[GPG-**Xi]ni**-GPS-(A)n2}, wherein for each quasi-repeat unit:

Xi is independently selected from the group consisting of SGGQQ, GAGQQ, GQGPY, AGQQ, and SQ; and

n1 is from 4 to 8, and n2 is from 6 to 10.

14. The recombinant protein fiber of claiml3, wherein n1 is from 4 to 5 for at least half of the quasi-repeat units.

15. The recombinant protein fiber of any one of claims 13-14, wherein n2 is from 5 to 8 for at least half of the quasi-repeat units.

16. The recombinant protein fiber of any one of claims 13-15, wherein each quasi-repeat unit has at least 95% sequence identity to a MaSp2 dragline silk protein subsequence.

17. The recombinant protein fiber of any one of claims 13-16, wherein the recombinant protein sequence comprises alanine-rich regions and glycine-rich regions, wherein:

the alanine-rich regions form a plurality of nanocrystalline beta-sheets; and

the glycine-rich regions form a plurality of beta-turn structures.

18. The recombinant protein fiber of any one of claims 13-17, wherein the repeat unit comprises SEQ ID NO: 1.

19. The recombinant protein fiber of any one of claims 1-18, wherein the linear density and the tenacity are measured using FAVIMAT fiber tensile test equipment model Favimat+ and Robot2.

20. The recombinant protein fiber of any one of claims 1-20, with the proviso that the recombinant fiber is not a MaSp2 dragline silk protein.

21. A yarn comprising the recombinant protein fiber of any of claims 1-20, wherein the yarn is a filament yarn.

22. A yarn comprising the recombinant protein fiber of any of claims 1-20, wherein the yarn is a spun yarn.

23. A yarn comprising the recombinant protein fiber of any of claims 1-20, wherein the yarn is a blended yarn.

24. A textile comprising the yarn of any of claims 21-23, wherein the textile is a knitted textile.

25. The textile of claim 24, wherein the textile is selected from the group consisting of a circular-knitted textile, a flat-knitted textile, and a warp-knitted textiles.

26. A textile comprising the yarn of any of claims 21-23, wherein the textile is a woven textile.

27. The textile of claim 26, wherein the textile is selected from the group consisting of a plain weave textile, a dobby weave textile, and a jacquard weave textile.

28. A textile comprising the yarn of any of claims 21-23, wherein the textile is a non-woven textile.

29. The textile of claim 28, wherein the textile is selected from the group consisting of a needle punched textile, a spunlace textile, a wet-laid textile, a dry-laid textile, a melt-blown textile, and a 3-D printed non-woven textile.

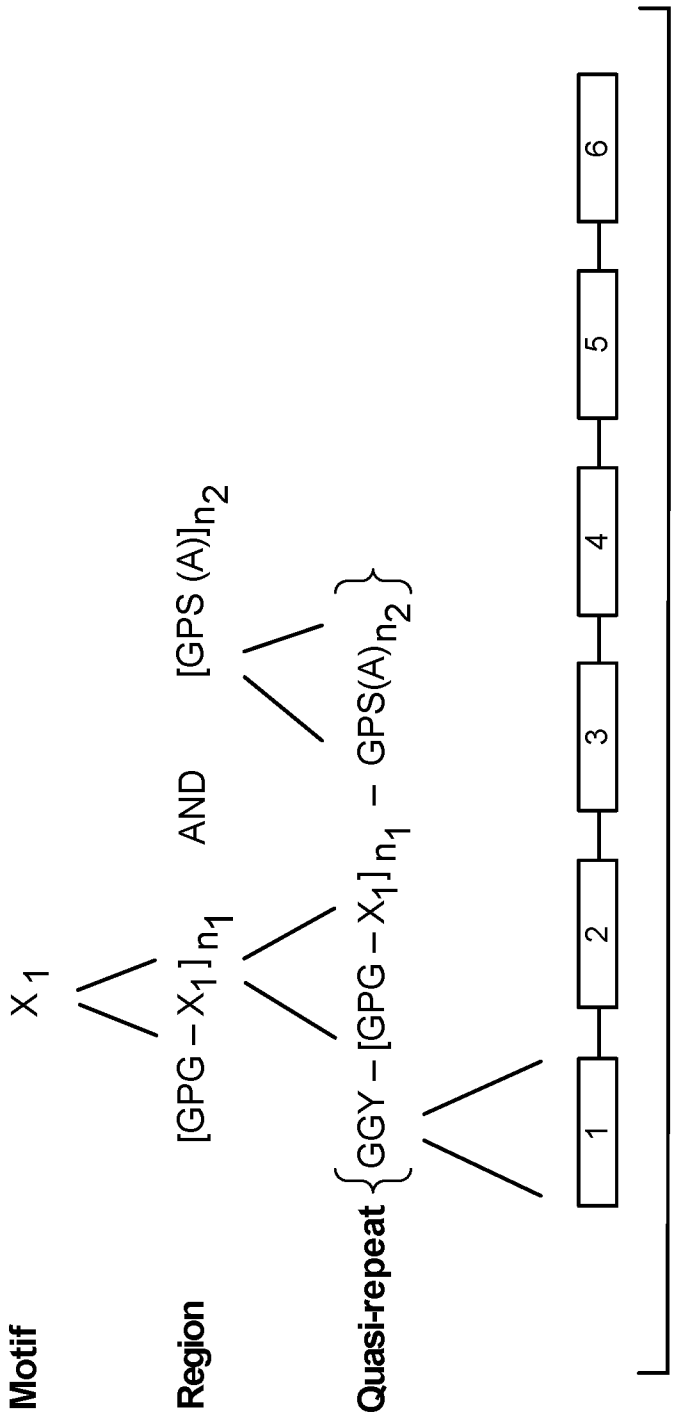

Motif
X 1
Region
[GPG – X1]n1
AND
[GPS (A)]n2
Quasi-repeat
{GGY – [GPG – X1]n1 – GPS(A)n2 }
1
2
3
4
5
6
Repeat Unit

FIG. 1

Tenacity Data

95% of data is between:
Mean: 9.07 – 9.42
S.D.: 0.38 – 0.63
CV: 4.15 – 6.90

Quantiles			
100.0% maximum	9.5268		0.70675
99.5%	9.49469		0.70091
97.5%	9.42558		0.63367
90.0%	9.37436		0.5859
75.0%	quartile	9.2981	0.54048
50.0%	median	9.2292	0.49533
25.0%	quartile	9.1729	0.45051
10.0%		9.10924	0.41582
2.5%		9.06507	0.38359
0.5%		9.00808	0.35445
0.0%	minimum	8.9924	0.32725

Summary Statistics	
Mean	9.237
Std Dev	0.0964312
Std Err Mean	0.0045972
Upper 95% Mea	9.2460352
Lower 95% Mean	9.2279648
N	440

Summary Statistics	
Mean	0.4989177
Std Dev	0.0655433
Std Err Mean	0.0031247
Upper 95% Mea	0.5050588
Lower 95% Mean	0.4927765
N	440

Quantiles			
100.0% maximum	7.67902		
99.5%	7.64894		
97.5%	6.89992		
90.0%	6.38679		
75.0%	quartile	5.86316	
50.0%	median	5.36804	
25.0%	quartile	4.86866	
10.0%		4.49768	
2.5%		4.14919	
0.5%		3.8851	
0.0%	minimum	3.5731	

Summary Statistics	
Mean	5.4030374
Std Dev	0.7223658
Std Err Mean	0.0344374
Upper 95% Mea	5.4707202
Lower 95% Mean	5.3353547
N	440

FIG. 2

Linear Density Data
95% of data is between:
Mean: 8.49 – 9.20
S.D.: 0.65 – 1.21
CV: 7.45 – 13.72
Mean(Lin.Den. (dtex))
Std Dev(Lin.Den. (dtex))
CV(Lin.Den. (dtex))
9.4
9.2
9
8.8
8.6
8.4
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
16
15
14
13
12
11
10
9
8
7
Quantiles
100.0% maximum 9.3248
99.5% 9.27802
97.5% 9.20319
90.0% 9.06092
75.0% quartile 8.9587
50.0% median 8.8334
25.0% quartile 8.7086
10.0% 8.60788
2.5% 8.48739
0.5% 8.39507
0.0% minimum 8.3136
Summary Statistics
Mean 8.8381364
Std Dev 0.1797675
Std Err Mean 0.0085701
Upper 95% Mea 8.8549799
Lower 95% Mean 8.821929
N 440
Quantiles
100.0% maximum 1.3987
99.5% 1.32074
97.5% 1.21212
90.0% 1.11669
75.0% quartile 1.02587
50.0% median 0.92911
25.0% quartile 0.8373
10.0% 0.75639
2.5% 0.64603
0.5% 0.59303
0.0% minimum 0.56247
Summary Statistics
Mean 0.9317258
Std Dev 0.1435042
Std Err Mean 0.0068413
Upper 95% Mea 0.9451715
Lower 95% Mean 0.91828
N 440
Quantiles
100.0% maximum 16.2798
99.5% 14.9098
97.5% 13.717
90.0% 12.5746
75.0% quartile 11.5915
50.0% median 10.4808
25.0% quartile 9.44448
10.0% 8.64335
2.5% 7.44798
0.5% 6.76573
0.0% minimum 6.6008
Summary Statistics
Mean 10.536312
Std Dev 1.5809329
Std Err Mean 0.075368
Upper 95% Mea 10.684439
Lower 95% Mean 10.388185
N 440

FIG. 3

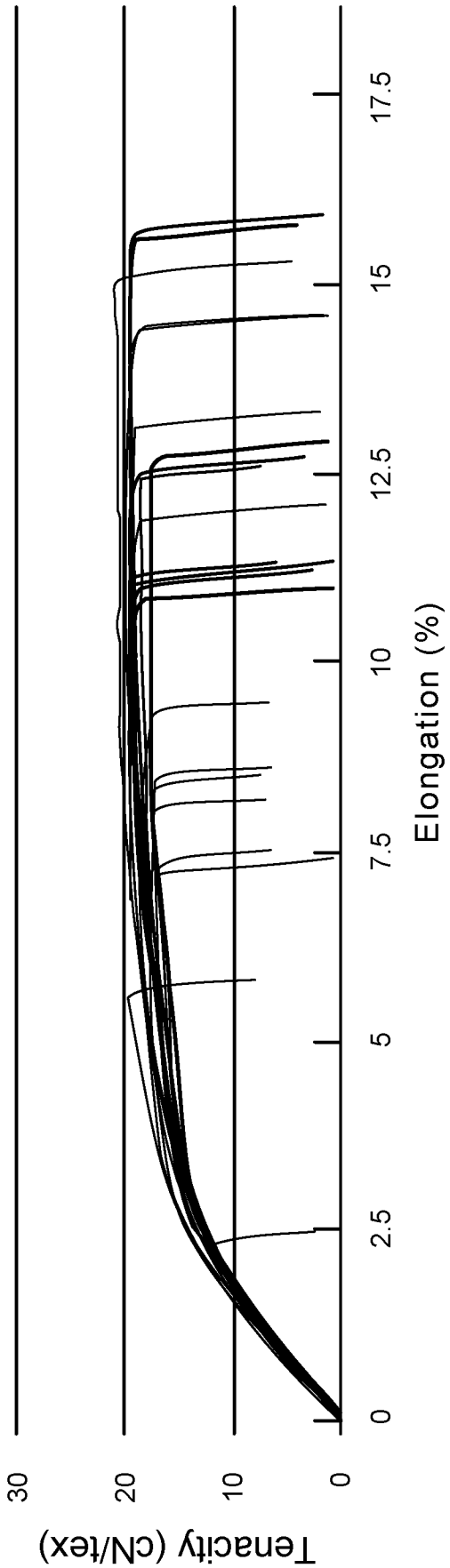


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US17/51668

A. CLASSIFICATION OF SUBJECT MATTER
IPC - C07K 14/435; D01D 5/00; D01F 4/00 (2017.01)
CPC - C07K 14/435, 14/43518; D01F 4/00; D02G 3/02, 3/042

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2016/0222174 A1 (BOLT THREADS, INC.) 4 August 2016; abstract	1-4
A	WO 2011/039345 A1 (VIB VZW, et al.) 7 April 2011; page 1, line 1; page 5, line 17	1-4
A	WO 2015/042164 A2 (REFACTORED MATERIALS, INC.) 26 March 2015; abstract; paragraph [0096]	1-4
A	(XIA, XX et al.) Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber. Proceedings of the National Academy of Sciences of the United States of America. 10 August 2010, Epub 26 July 2010, Vol. 107, No. 32; pages 14059-14063; abstract; DOI: 10.1073/pnas.1003366107	1-4
P, X	WO 2016/201369 A1 (BOLT THREADS, INC.) 15 December 2016; entire document	1-4
P, X	WO 2016/149414 A1 (BOLT THREADS, INC.) 22 September 2016; entire document	1-4

«

	Further documents are listed in the continuation of Box C.		See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 December 2017 (07.12.2017)	05 JAN 2018

Name and mailing address of the ISA/	Authorized officer
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US17/51668

Box No. I **Nucleotide and/or amino acid sequencc(s) (Continuation of item l.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed:
      ☐ in the form of an Annex C/ST.25 text file.
      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☒ furnished subsequent to the international filing date for the purposes of international search only:
      ☒ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☒ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US1 7/51668

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 5-29
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. Ill Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.