HHS Public Access

Author manuscript

*ACS Biomater Sci Eng*. Author manuscript; available in PMC 2018 August 14.

Published in final edited form as:

*ACS Biomater Sci Eng*. 2017 August 14; 3(8): 1542–1556. doi:10.1021/acsbiomaterials.7b00292.

# Predicting Silk Fiber Mechanical Properties through Multiscale Simulation and Protein Design

**Nae-Gyune Rim**[†,‡], **Erin G. Roberts**[†,#], **Davoud Ebrahimi**[§], **Nina Dinjaski**[§,‖], **Matthew M. Jacobsen**[‡], **Zaira Martín-Moldes**[‖], **Markus J. Buehler**[§], **David L. Kaplan**[‖], and **Joyce Y. Wong**[*,‡,#]

[‡]Department of Biomedical Engineering, Boston University, 44 Cummington Mall, Boston, Massachusetts 02215, United States

[#]Division of Materials Science and Engineering, Boston University, 44 Cummington Mall, Boston, Massachusetts 02215, United States

[§]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

[‖]Department of Biomedical Engineering, Tufts University, 4 Colby Street, Medford, Massachusetts 02155, United States

## Abstract

Silk is a promising material for biomedical applications, and much research is focused on how application-specific, mechanical properties of silk can be designed synthetically through proper amino acid sequences and processing parameters. This protocol describes an iterative process between research disciplines that combines simulation, genetic synthesis, and fiber analysis to better design silk fibers with specific mechanical properties. Computational methods are used to assess the protein polymer structure as it forms an interconnected fiber network through shearing

[*]**Corresponding Author:** jywong@bu.edu., Tel.: +1-617-353-2374. Fax: +1-617-353-6766.

[†]**Author Contributions**

N.R. and E.G.R. contributed equally to this work.

**Supporting Information**

The Supporting Information is available free of charge on the ACS Publications website at DOI:10.1021/acsbiomaterials.7b00292.

Compressed file that includes the following files used for simulation. Open the zip file and choose the appropriate file to run computational simulation as outlined in this protocol (ZIP)

pair_soft_modified.cpp (cpp file), equipment setup Generate_Configuration.m (m file), step 1 of simulation procedure

equil_shear_stretch.in (in file), step 3 of simulation procedure

analize_full_process (tar file), step 6 of simulation procedure

connectivity_Analysis.m (m file), step 13 of simulation procedure

Network_Conductance.m (m file), step 13 of simulation procedure

Node_bridge_diagram_multiplicity.m (m file), step 13 of simulation procedure

**ORCID**

Erin G. Roberts: 0000-0001-8985-8799

Davoud Ebrahimi: 0000-0001-9898-7023

Nina Dinjaski: 0000-0002-4765-2730

Matthew M. Jacobsen: 0000-0002-0368-3439

Zaira Martín-Moldes: 0000-0002-2932-8064

Markus J. Buehler: 0000-0002-4173-9659

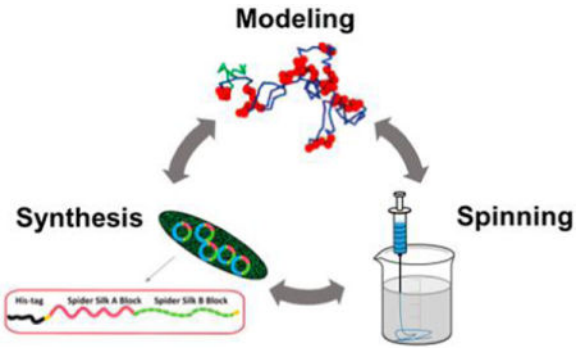
David L. Kaplan: 0000-0002-9245-7774

**Notes**

The authors declare no competing financial interest.

and how this process affects fiber mechanical properties. Model outcomes are validated experimentally with the genetic design of protein polymers that match the simulation structures, fiber fabrication from these polymers, and mechanical testing of these fibers. Through iterative feedback between computation, genetic synthesis, and fiber mechanical testing, this protocol will enable a priori prediction capability of recombinant material mechanical properties via insights from the resulting molecular architecture of the fiber network based entirely on the initial protein monomer composition. This style of protocol may be applied to other fields where a research team seeks to design a biomaterial with biomedical application-specific properties. This protocol highlights when and how the three research groups (simulation, synthesis, and engineering) should be interacting to arrive at the most effective method for predictive design of their material.

## Graphical abstract




### Keywords

recombinant silk; computational modeling; genetic synthesis; spinning

## INTRODUCTION

### Power of Predictive Biomaterial Design

The predictive design of biomaterials provides an instrumental toolset and a wealth of opportunities for efficient determination of fabrication parameters necessary for materials with desired functional properties. Specific to the tissue engineering and drug delivery fields, materials can be better designed for mechanical performance or in vivo functionality with limited amounts of experimental iteration as the need for traditional trial-and-error methodology is greatly reduced through computational simulation; our research teams have documented many such application specific, simulation supported experimentation studies in the literature.[1–4]

In this protocol, we will describe a collaborative material design process which utilizes iterative feedback between three research components. The basis of this process is matching coarse-grained modeling to overall physical behavior of a correlated peptide sequence. The strength of this hierarchical view of protein fiber assembly is that it relies on universally applicable thermodynamic phenomena to describe fiber assembly, namely entropic reorientation of hydrophilic chains and enthalpic bonding of hydrophobic chains.[5]

Therefore, even though this protocol focuses on work done with recombinant silks, the predictive capabilities of this protocol are not limited to specific sequences nor experimental conditions, and may be tailored to the specific peptide designs of the research team. Consequently, the methods described herein are relevant to collaborative research groups that have teams who are strong in both computation and experimentation disciplines.

## Motivation for Predictive Design of Silk Biomaterials

Silk is a promising material for biomedical applications because of its exceptional mechanical properties, versatile processing in water into different material formats, and biocompatibility.[6–9] Applications for silk materials include tissue regeneration, sutures, and gene/drug delivery.[8,10–13] Much of today's silk biomaterial research is focused on understanding how the desirable mechanical properties of naturally produced silk can be replicated synthetically through proper amino acid sequences and processing parameters such as shear flow. Researchers seek to recreate what is found in nature and to utilize this understanding to then tailor synthetic silk properties to meet property requirements specific to different biomedical applications.[14–19] The opportunity to tailor the properties of silk toward specific biomedical applications offers an important step forward in matching material design to functional need. The versatile features of silk chemistry and genetic manipulation in combination with multiscale simulation of these properties allows scientists to determine how to design and process silk block copolymer sequences and structures that possess specific mechanical properties. Computational methods can be utilized to assess protein–polymer interactions at different length scales, and outcomes from these models can then be validated experimentally with controllable and precise genetic design of protein polymers and processing parameters for fiber fabrication. Once models and experimental results are in reasonable agreement, the simulations can then be used to inform future protein synthesis and silk fiber fabrication with targeted strength and toughness that satisfy specific mechanical performance requirements. It is getting to this computational and experimental agreement that requires continued feedback between the three entities, a crucial component that has been highlighted within this protocol and distinguishes it from others.

## Study Summary and Overview of Experimental Design

As discussed above, understanding how the desirable mechanical properties of naturally produced silk can be reproduced and tailored to a specific application is a prominent challenge in the field of silk research; as well as a topic of focus in polymer-materials research. This protocol describes an iterative design approach to solve this challenge by combining results and inputs from three different research disciplines: computational simulation, silk block copolymer synthesis, and biopolymer processing using fiber spinning. The integration of these three disciplines permits the elucidation of structure–function relationships for silk biomaterials, from the molecular level designs to the macroscopic fiber performance. The specific example presented in this protocol describes synthetic silk protein sequences containing three major building blocks: a hydrophobic domain, a hydrophilic domain, and a hexahistidine fusion tag.[20] The research focuses on how the lengths and the ratio of the hydrophobic/hydrophilic domains, the overall chain length of the polymer, and shear forces during fiber processing affect the polymer network conductivity and fiber

mechanical properties. Figure 1 depicts the cross-talk between the three major disciplines involved in this process. In the following sections, we will discuss the general functions and experiments that each of the three research sectors (identified by the hexagons) performs. It is important to note that information flows in both directions between the sectors; examples of information exchanged is shown in the arrows. Although protocols specific to each entity are prominent in the literature,[20–22] this protocol demonstrates how these groups should be interacting to arrive at methods for effective predictive design.

## Simulation Work

Dissipative particle dynamics (DPD) is the simulation method used in this protocol to study polymer network analysis and to predict mechanical properties. DPD is a stochastic simulation technique used to simulate the dynamic and rheological properties of fluids.[23] The method is carried out using Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS), a classical molecular dynamics package, which can be used to model atoms or simulate particles at the atomic, meso, or continuum scale.[24] Silk fibers are modeled as polymeric soft matter; this approach allows for more structural flexibility for the protein building blocks of silk and is more suitable for modeling the large structural evolution during the shear flow spinning process. In this approach, silk peptides and proteins are described as multiblock copolymer chains composed of hydrophobic “a” and hydrophilic “b” beads, with each bead representing 3 amino acids in the “A” and “B” domains of the copolymer. The initial random structure of the polymer is established in the model and then shear flow is applied to this structure. Polymer network analysis is then performed by comparing node-bridge diagrams before and after shear flow to study changes in network connectivity. Simulated tensile and effective stress/strain measurements before and after this process are also taken. Experimental work that has been conducted to determine the amino acids and secondary structure in naturally occurring silk sequences informs the computational model in terms of which molecular entities and mechanisms to use in the simulation structures. For the simulated spinning process, experimental process parameters such as concentration of spinning solution and spin speed are fed into the model as well.

## Experimental Work: Genetic Block Copolymer Synthesis and Expression

The same polymers that are modeled are also synthesized experimentally. Block copolymers composed of ‘A’ (amino acid sequence = SGAGAAAAAGGAGT) and ‘B’ (amino acid sequence = SGRGGLGGQGGYT) domains are cloned. The cloned sequences are chosen based on consensus sequences in spider silk and synthesized as short single-stranded oligonucleotides (up to 100 bp) by commercial oligonucleotide synthesis or used directly as polymerase chain reaction products from cDNA libraries. Large repetitive sequences can be constructed using the following cloning strategies: concatemerization, step-by-step directional approach and recursive ligation.[25] Concatemerization is used for construction of genetic libraries; this approach cannot be used for preparation of genes of specific sizes.[26] Step-by step and recursive directional ligation allows for facile modularity, where control over the size of the genetic cassettes is required. Moreover, recursive directional ligation eliminates the restriction sites at the junctions between monomeric genetic cassettes without interrupting key gene sequences with additional base pairs[27] and is the strategy that has been used here. Designed sequences are transferred into *Escherichia coli* and transformed cells

are cultivated using a fermenter for high cell-density cultivation. After cultivation, cells are then harvested, lysed, and the supernatant is run through a separation column; the silk fractions are collected, dialyzed, and then lyophilized. SDS-PAGE is used to confirm purity and mass spectrometry is used to identify protein content. SEM is used for further morphological characterization, FTIR for secondary structure analysis, and AFM for force measurements.

### Experimental Work: Bioinspired Spinning and Mechanical Fiber Testing

Once the synthetic silks are fabricated, the recombinant silk proteins are dissolved in a suitable solvent, such as hexafluoroisopropanol (HFIP) and extruded through a thin needle/spinneret into an coagulation bath to form solid fibers.[28] The hydrodynamic flow focusing and shear mimic aspects of the spinning of native silk fibers. Fiber segments are mounted onto frames for imaging and fiber diameter measurements. A uniaxial tensile tester or Instron is used to displace the ends of the fiber at a constant rate while measuring the force due to pulling. Fibers are pulled until failure and modulus, toughness, and ultimate tensile data can be calculated from these measurements.

### Interactions between Three Research Sectors

As discussed above, the main point emphasized in this protocol is the necessary cross-talk and feedback between the three research entities as each one informs the others of process inputs and characterization outputs. It is the integration of this data that allows for the eventual predictive design of silk biomaterials. Within the procedure section of this protocol, in addition to flagging critical steps and caution areas, collaboration points have also been called out. These are areas where an individual sector should be checking in with one of the other teams as noted.

## MATERIALS

### Simulation

#### Equipment

- LAMMPS: Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) is a molecular dynamics program from Sandia National Laboratories. The LAMMPS software and its installation manual can be found at http://lammps.sandia.gov/. In this paper, we assume software is installed under the Linux environment. We used the LAMMPS, November 17th, 2016, version.
- OCTAVE: Octave is a free open source programming language primarily intended for numerical computations. It is compatible with MATLAB and can be downloaded from (http://www.gnu.org/software/octave/download.html). In this paper, we assume software is installed under the Linux environment.

#### Equipment Setup

- Installation: Compile LAMMPS in the Linux system with a modified “pair_soft.cpp” to represent the additional hydrogen bonding between hydrophobic (A) blocks.

- Download “pair_soft_modified.cpp”. This file and all other referenced in this protocol are available in the Supporting Information.
- Overwrite “pair_soft.cpp” in the lammps/src directory and compile LAMMPS.
- Install octave packages per the procedure detailed in the provided link (see Equipment section).

## Genetic Block Copolymer Synthesis and Expression

### Reagents

### Chemicals and Solvents

- Kanamycin sulfate powder (Sigma-Aldrich, St. Louis, MO, USA)
- 0.22 $\mu$m Durapore PVDF filters (EMD Millipore, Billerica, MA, USA)
- Lysogeny broth (LB) powder (Fisher Scientific, Pittsburgh, PA, USA)
- Glycerol (Sigma-Aldrich, St. Louis, MO, USA)
- Hyper Broth (Athena Environmental Sciences, Inc., Baltimore, MD, USA)
- Glucose Nutrient Mix (Athena Environmental Sciences, Inc., Baltimore, MD, USA)
- Agarose powder (Sigma-Aldrich, St. Louis, MO, USA)
- Loading dye electrophoresis (Fisher Scientific, Pittsburgh, PA, USA)
- Ethidium bromide (EtBr) (Fisher Scientific, Pittsburgh, PA, USA)
- IPTG (Aldrich-Sigma Chemical Co. Ltd., cat. no. I6758; see Reagent Setup)
- Ni-NTA His-bind resin (Novagen, EMD Millipore, Billerica, MA, USA)
- NuPAGE Novex 12% Bis-Tris Protein Gels, 1.0 mm, 10-well (Thermo Fisher Scientific, Waltham, MA, USA)
- NuPage sample reducing agent (Thermo Fisher Scientific, Waltham, MA, USA)
- NuPAGE LDS Sample Buffer (4X) (Thermo Fisher Scientific, Waltham, MA, USA)
- SimplyBlue Safe Stain (Thermo Fisher Scientific, Waltham, MA, USA)
- Urea (Sigma-Aldrich, St. Louis, MO, USA)
- Sodium phosphate monobasic (Sigma-Aldrich, St. Louis, MO, USA)
- Tris hydrochloride (Sigma-Aldrich, St. Louis, MO, USA)
- Empty columns (Bio-Rad, Hercules, CA, USA)

### Vectors and Bacterial Strains

- pET30a(+) (Novagen, San Diego, CA, USA)
- pUC57(Kn)-A, pUC57(Kn)-B (GenScript, Piscataway, NJ, USA)

- *E. coli* DH5α chemically competent cell (New England Biolabs, Ipswich, MA, USA)
- *E. coli* BL(DE3) chemically competent cells (New England Biolabs, Ipswich, MA, USA)

### Enzymes

- T4 DNA Ligase (New England Biolabs, Ipswich, MA, USA)
- Restriction endonucleases: Nco, XhoI, NheI, SpeI (New England Biolabs, Ipswich, MA, USA)
- Antarctic Phosphatase (New England Biolabs, Ipswich, MA, USA)

### Molecular Weight Markers

- One kb DNA Ladder (New England Biolabs, Ipswich, MA, USA)
- Novex Sharp Pre-Stained Protein Standard (Thermo Fisher Scientific, Waltham, MA, USA)

### Kits

- DNA isolation: QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA, USA)
- Gel extraction: QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA)

### Buffers and Solutions

- NEB endonuclease restriction enzyme buffer set (New England Biolabs, Ipswich, MA, USA)
- DNA Ligase Buffer (New England Biolabs, Ipswich, MA, USA)
- TAE (Tris-acetate-EDTA) buffer (Fisher Scientific, Pittsburgh, PA, USA)
- 10X Antarctic Phosphatase Reaction Buffer (New England Biolabs, Ipswich, MA, USA)
- NuPAGE MES SDS Running Buffer (20X) (ThermoFisher Scientific, Waltham, MA, USA)
- Denaturing buffer (8 M urea, 0.1 M NaH2PO4, 0.01 M Tris-HCl, pH 8.0)

### Reagent Setup

- Antibiotics stock solution–kanamycin (Kn) (50 mg/mL): dissolve 0.5 g of kanamycin sulfate powder (Sigma-Aldrich, St. Louis, MO, USA) in a 15 mL Falcon tube containing 10 mL of deionized water. Filter with sterilized 0.22 μm Durapore PVDF filters (EMD Millipore, Merck KGaA, Darmstadt, Germany) in biohood to sterilize antibiotics solution. Store at −20 °C.
- Culture plates for colony selection, LB (Kn) plates: dissolve 7.5 g of Lysogeny broth (LB) powder (Fisher Scientific, Pittsburgh, PA, USA) in a 500 mL flask containing 300 mL of deionized water. Add 4.5 g of agar powder into the LB

solution. Stir well on a stir plate, and autoclave at 121 °C 15 psi for 20 min. Add 300 μL of kanamycin stock solution at about 50 °C in biohood. Mix well on a stir plate. Pour LB/agar/Kn solution into sterilized Petri dishes, wait until solidified, and put the lid back on for each plate in biohood. Store at 4 °C, in sterilized plastic bags.

- Culture media, LB (Kn) media: dissolve 25 g of LB powder in a 4 L flask containing 1 L of deionized water. Stir well on a stir plate, and autoclave at 121 °C 15 psi for 20 min. Add 1 mL of kanamycin stock solution at 37 °C in biohood. Mix well on a stir plate. Store at room temperature.
- Prepare 1×TAE (Tris-acetate-EDTA) buffer from 50×TAE buffer stock solution (Fisher Scientific, Pittsburgh, PA, USA).
- To make 40 mL of a 1% agarose gel: dissolve 0.4 g of agarose powder (Sigma-Aldrich, St. Louis, MO, USA) in 40 mL of 1×TAE buffer by heating in a microwave 30s twice. Let agarose solution cool down for 5 min and add ethidium bromide (EtBr) (Fisher Scientific, Pittsburgh, PA, USA) to a final concentration of approximately 0.2–0.5 μg/mL (usually about 2–3 μL of lab stock solution per 100 mL of gel). Seal the open ends of the plastic tray with the rubber stoppers and pour the 1% agarose gel solution into the sealed plastic gel tray. Position the 13-well comb at one end of the tray so that it is 0.5–1.0 mm above the bottom of the tray and not touching the rubber stopper; allow the gel to harden.
- DNA isolation: QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA, USA).
- Gel extraction: QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA).
- IPTG 100 mM stock: Dissolve 0.238 g in 10 mL of deionized water. Filter sterilize using a 0.22 mm pore-size filter. Aliquot and store at −20 °C for up to 2 years. Use at concentration of 1 mM.

## Equipment

- 20 mL syringes (BD Medical, Franklin Lakes, NJ, USA)
- 0.22 mm syringe filters (Nalgene, Thermo Fisher Scientific, Waltham, MA, USA)
- Sterile 100 mm × 25 mm Petri dishes (Midwest Scientific Inc., Minneapolis, MN, USA)
- Laminar airflow biohood (1300 Series Class II, Type A2 Biological Safety Cabinet Packages, Thermo Fisher Scientific, Waltham, MA, USA)
- 2 mL microcentrifuge tubes (Eppendorf, Hauppauge, NY, USA)
- Surgical blades (Fisher Scientific, Pittsburgh, PA, USA)
- UV–visible light spectrophotometer (GENESYS 10S UV–vis spectrophotometer, Thermo Fisher Scientific, Waltham, MA, USA)

- Sorvall refrigerated centrifuge (Sorvall RC 6 Plus centrifuge, Thermo Fisher Scientific, Waltham, MA, USA)
- 2 L centrifuge bottles with screw caps (Thermo Fisher Scientific, Waltham, MA, USA)
- 50 mL conical tubes (Falcon, Fisher Scientific, Pittsburgh, PA, USA)
- Slide-A 2000 Da molecular weight cutoff dialysis cassette (Slide-A-Lyzer Dialysis Cassettes, 2K MWCO, Thermo Fisher Scientific, Waltham, MA, USA)
- Four L glass Erlenmeyer flasks (Thermo Fisher Scientific, Waltham, MA, USA)
- Ten L beakers (Thermo Fisher Scientific, Waltham, MA, USA)
- Analytical balance (Mettler Toledo, Columbus, OH, USA)
- Lyophilizer (FreeZone 2.5 Liter Benchtop Freeze-Dry Systems, Labconco, Kansas City, MO, USA)
- Vortex mixer (Fisher Scientific Analog Vortex Mixer, Thermo Fisher Scientific, Waltham, MA, USA)
- NanoDrop 200 Spectrophotometers (Thermo Fisher Scientific, Waltham, MA, USA)
- Gel electrophoresis equipment (Mupid-2plus System, Takara, Jusatsu, Shiga, Japan)
- Protein gel electrophoresis chamber system (XCell *SureLock* Mini-Cell, Thermo Fisher Scientific, Waltham, MA, USA)
- Incubator (Precision Incubator, 11.2 cu ft, 2-Door, Mechanical; 120 V, Thermo Fisher Scientific, Waltham, MA, USA)
- pH meter (Fisher Scientific accumet AB15 Basic and BioBasic pH/mV/°C Meters, Thermo Fisher Scientific, Waltham, MA, USA)
- Shaking incubator set at 37 °C/250 r.p.m. (New Brunswick G 25 Shaker Incubator, GMI, Ramsey, MN, USA)
- Centrifuge (Centrifuge 5804 R, Eppendorf, Hauppauge, NY, USA)
- Rotating Tube Mixer (Benchmark Scientific, Edison, NJ, USA)
- Thermoblock (THERMOBLOCK TD 200 P1, FALC INSTRUMENTS S.R.L., Treviglio, BM, Italy)
- Thermocycler (MJ Research PTC-200 Thermal Cycler, GMI, Ramsey, MN, USA)
- Autoclave (Consolidated Sterilizer Systems, Boston, MA, USA)
- Gel imager (G:BOX, Syngene, Frederick, MD, USA)
- Shaking water bath (SHEL LAB, Sheldon Manufacturing, Inc., Cornelius, OR, USA)

## Bioinspired Spinning Process

### Reagents

### Protein Solvent Options

- Deionized water
- Phosphate-buffered solution (PBS) (Sigma-Aldrich, St. Louis, MO, USA)
- Hexafluoroisopropanol (HFIP) (Sigma-Aldrich, St. Louis, MO, USA)
- 9.8 M lithium bromide (LiBr) (Sigma-Aldrich, St. Louis, MO, USA)
- Formic acid (Sigma-Aldrich, St. Louis, MO, USA)

***Coagulation Bath Solvent Options:*** A solution that further concentrates the fiber stream through dehydration of the spinning dope as it exits the spinning needle. Different solvents lead to different crystallization times, again providing design options to the investigator. Solvents that are miscible with water can be further tuned in their composition to affect the crystallization time as well.

- Methanol (Sigma-Aldrich, St. Louis, MO, USA)
- 2-propanol (Sigma-Aldrich, St. Louis, MO, USA)
- Acetone (Sigma-Aldrich, St. Louis, MO, USA)

### Equipment

- MiniSpin microcentrifuge (Eppendorf, Hauppauge, NY, USA)
- 1.5 mL microcentrifuge tubes (Fisher Scientific, Pittsburgh, PA, USA)
- 600 mL glass beakers (Fisher Scientific, Pittsburgh, PA, USA)
- 20 mL disposable scintillation glass vials with lids (Kimble, Fisher Scientific, Pittsburgh, PA, USA)
- Spinning apparatus (see Equipment Setup)
- Gastight syringe (250 uL, Hamilton Company, Reno, NV, USA)
- Syringe pump (Fusion 100, Chemyx, Stafford, TX, USA)
- Soldering iron (RadioShack, Fort Worth, TX, USA)
- Copier transparency Film (3M, St. Paul, MN, USA)
- Microscope (Axiovert S-100, Carl Zeiss AG, Oberkochen, Germany)
- Instron microtester or similar tensile testing apparatus (Instron, Norwood, MA)

### Equipment Setup

- Mount the syringe on the Fusion 100 syringe pump according to the manufacturer’s instructions, and set the pump speed rate at 15 uL/min. Set up the syringe pump vertically. For fiber extrusion, prepare the desired coagulation solvent in beakers to use as coagulation bath.

- Cut transparency film into the appropriate frame size to mount the fibers.

## PROCEDURE

### Simulation: Timing 170 CPU Hour

**DPD Setup Using LAMMPS Package—**Note: The actual simulation time depends on many factors, including setting of the system (for instance, the size of the system, length of the simulation, time step), calculation algorithm in LAMMPS, RAM, software and hardware configurations of the computer used to run the simulation. The estimation of required time is for simulating $H(AB)_2$ system with specified system setup in the attached supplementary files. We used the computational resources provided by the National Science Foundation through the Extreme Science and Engineering Discovery Environment (XSEDE) on the bridges server (https://portal.xsede.org/psc-bridges).

**1** Generate initial configurations and LAMMPS input file (randomly distributed peptide chains in a water box) by downloading “Generate_Configuration.m”.

**a.** Setup elongated rectangular box with a square cross-section and periodic boundary condition. The box length in the direction of flow ($x$ direction in our example) should be chosen to be longer than a single extended chain to avoid artifacts from boundary conditions.

**b.** Define polymer beads based on the coarse-graining level of three amino acid residues per bead (Figure 2 and Table 1).

**Collaboration Point!** Informed by Block Co-Polymer Synthesis Team–provides information on the silk molecular sequence such that the correct simulation properties are assigned to the polymer beads representing the amino acid residues used for synthesis.

**c.** Fix concentration at 20% volume fraction of coarse-grained polymer beads; the rest is filled with polymer beads (each representing 3 amino acids).

**Collaboration Point!** Informed by Spinning Team–provides concentration of silk solution used to spin fibers such that simulation concentration matches processing conditions.

**d.** For starting structure, grow peptide chains randomly with a fixed distance $a$ (equilibrium bond distance in chain) between beads in simulation box up to 20% of box volume.

**e.** Fill in box with solvent beads until a number density of 3 is reached.

**Critical Step!** Key variables one can change:

repeat_motif_spider: repeating units or motif, e.g., ’A1B1’

num_hydrophobic_spider: how many ‘a’ beads in A block

num_hydrophilic_spider: how many ‘b’ beads in B block

num_histidine_spider: how many ‘b’ beads in H block

num_repeat_spider: how many repeats of the defined motif

**Caution!** System characteristics such as number of repeating units, ratio of hydrophobic to hydrophilic block, and length of hydrophobic and hydrophilic blocks should be defined by adjusting above parameters.

**Caution!** In the simulation, we have assumed water as a solvent and the parameters used for interaction of hydrophobic and hydrophilic beads correspond to system solvated in water. With this assumption, as a first approximation, we could capture the effect of hydrophobic to hydrophilic ratio, length and concentration of the peptides, and shear flow in aggregation of silk block copolymer. This assumption needs to be revisited (by introducing new bead types and/or changing interaction parameters of solvent beads) to study the effect of other solvents.

**2** Execute “Generate_Configuration.m”. By installing Octave, one can execute the MATLAB code in mode, e.g., Octave “Generate_Configuration.m, m”. This will generate a LAMMPS data file with bead coordinates and bonding information, as well as “psf” files with or without water beads for connectivity mapping (for visualization in VMD and network analysis).

**3** To prepare LAMMPS input file, download LAMMPS input file “equil_shear_stretch.in”. The following steps are included inside the LAMMPS input file:

**a.** Name of the data file generated in the previous step should be modified (inside “#file name” section)

**b.** Simulate spinning process (NVE ensemble with a fixed time step of delta $t = 0.03\tau$ where $\tau = 0.75$ ns is the characteristic time scale).

**Collaboration Point!** Informed by Spinning Team–provides flow rate used to spin fibers such that simulation shear matches processing conditions.

**c.** Achieve steady state using equilibration simulation where root-mean-square-deviation (RMSD) is tracked for total number of equilibration steps (variable tot_equil_step = 210,000 time steps in the example). In this example, this step will generate 14 configurations as they are saved every 15 000 steps (variable dumpfreq = 15 000).

**Caution!** If RMSD shows the system has not reached to steady state, increase number of equilibration steps (tot_equil_step).

**d.** Apply shear rate of $\dot{\gamma}_{xy} = \dot{0}.01\tau_{xy}^{-1} = 0.01\tau^{-1}$ along $x$ axis for total number of shear steps (variable tot_shear_step = 420 000 time steps in this example). This is applied using the Lees boundary condition (Figure 3). In this example, this step will generate 28 configurations as they are saved every 15 000 steps (variable dumpfreq = 15 000).

**e.** Confirm equilibrations of self-assembled aggregates without and with shear flow by plateau in time evolutions of size aggregates (number of beads per node) and number of bridges before shear and during shear flow, respectively. Use additional time steps without shear to equilibrate system after shear flow. (Variable tot_equilafter_step = 150 000 time step in this example.) In this example, this step will generate 10 configurations as they are saved every 15 000 steps (variable dumpfreq = 15 000).

**Caution!** If aggregate size and/or number of bridges before and after applying shear has not reached to a plateau, increase number of equilibration steps (step c, tot_equil_step) and/or number of shear steps (step d, tot_shear_step).

**f.** Perform tensile simulations in x direction by stretching simulation box at a constant engineering strain rate of $7.5 \times 10{-}6\,\tau^{-1}$, whereas the $y$ and $z$ directions are adjusted according to their aspect ratio to maintain a constant volume (variable tot_stretch_-step = 100 000 time step in this example). In this example, this step will generate 10 configurations as they are saved every 10 000 steps (variable dump-freq2 = 10 000).

**Collaboration Point!** Informed by Spinning Team–stretch rate used on tensile tester such that simulation sampling matches mechanical testing.

**4** Run LAMMPS code. This step will generate the following “dcd” trajectory and stress data files:

equil_11111.dcd

shear_11111.dcd

shear_relaxed_11111.dcd

stretch_11111.dcd

all_stress_11111.txt

## Data Analysis of Polymer Network

**5** Copy output files from Step 4 into the same directory.

**6** Download “analize_full_process.tar” into that same directory.

**7** Unzip file: “tar-xvf analize_full_process.tar” (multiple C++, MATLAB, and bash script files that automatically analyze full simulation steps).

**8** Copy the psf file without water in Step 2 (the file with “water_only.psf” extension) as “ref.psf.”

**9** Using VMD (we have used 1.9.1 version), “ref.psf,” and any “dcd” file (step 4), save one frame as “ref.pdb” (this is only for the number counting so any single frame is fine).

**10** Check these files: “network” (needs to be recompiled from “network.cpp”), “catdcd,” “*.m,” “ref.pdb,” and “ref.psf” before running “anal_*.sh.”

**11** Run the bash script: “./anal_full_process_single_-seed.sh.” For single frame analysis, download and use “analize_single_network.tar”. You might need to modify the number of steps in the loops depending on the number of steps in the LAMMPS input file (see number 3). In the first three lines (starting with “catdcd”) from the configurations saved as “dcd” file, every other one is saved into “pdb” format for further analysis. For instance, from the 14 configurations generated in the equilibrium part of the example, 7 configurations will be saved in the “pdb” format and analyzed inside the first loop.

**Caution!** This step is memory intensive. We have run it on XSEDE servers. If you get segmentation fault due to insufficient memory, you need to use a computer with larger RAM.

**12** Analyze the network:

- **a.** Use the depth-first search algorithm to find the cluster distribution efficiently.
- **b.** Each isolated cluster is then defined as a node whose coordinate is defined as the center of mass of “a” beads in the cluster.
- **c.** Determine the number of connections between two nodes. It is equal to the number of polymers through which the two nodes are linked. Node connectivity was used to generate the node-bridge diagrams.

**13** Plot results of polymer network analysis: “tar-cvf data.tar */boundar*.txt */*cluster*.txt */*connectivity*.”

- **a.** Copy the “data.tar” file into a Windows machine.
- **b.** Copy “Connectivity_Analysis.m” and “draw_ss_-curve.m” after unzip.
- **c.** “Connectivity_Analysis.m” plots the time evolution of network property, and “draw_ss_curve.m” plots the stress–strain curve.
- **d.** To plot the node-bridge diagram, copy “node_bridge_diagram_multiplicity.m” file into the directory “shear_evolve_*” (this will give equilibrium structure) and the directory “stretch_evolve_*” (this will give after-shear structure) and then run in MATLAB.
- **e.** To plot the network conductance (as a measure of the network connectivity), download “Network_-Conductance.m” and run it to get the time evolution of network conductance.

  **Collaboration Point!** Confirm with Block Co-Polymer Synthesis Team–simulation properties match those seen by experimental characterization of protein.

**14** See Figure 4A–C for sample results.[21]

### Data Analysis of Mechanical Properties

**15** Plot results of mechanical properties from step 3(e) above. Effective stress is calculated using the following formula: $\sigma_{\text{eff}} = \sigma_x - \nu(\sigma_y + \sigma_z)$.

Note: Stimulated system right after shear flow is prestretched. A compressive simulation is performed (negative strain rate of $-5 \times 10^{-6}\ \tau^{-1}$) to relax the system and determine stress-free configuration with a reduced box dimension as the new reference for subsequent tensile simulation and calculations.

**Collaboration Point!** Confirm with Spinning Team–compare effective stress–strain results from simulation mechanical properties and stress–strain results from fiber mechanical testing.

**16** See Figure 4D for sample results.

## Genetic Block Copolymer Synthesis and Expression: Timing 19–24 Days

### Generate Vector pET30L

**1** Cloning cassette linker: the linker sequence is designed as *Xho*I and *Nco*I inserted to be incorporated into pET30a-(+) (Figure 5A). The linker is used to generate modified pET30a(+) vector, pET30L that carries *Nhe*I and *Spe*I restriction sites.

**Critical Step!** Use appropriate software (e.g., Bioedit) to design sequences. General recommendations for sequence design are absence of dimerization capability, absence of significant hairpin formation (usually > 3 bp) and low specific binding at the 3′ end. Commercially obtain the oligonucleotide linker sequence. Prepare the oligonucleotide linker sequences as recommended by the manufacturer. Both forward and reverse linker oligonucleotides should be designed.

**2** To generate the cloning cassette linker, anneal two synthetic nucleotides, forward and reverse (1:1 v/v), by decreasing the temperature of a 20 pmol/μL oligonucleotide solution from 95 to 20 °C at a gradient of 0.1 °C/s. Mismatched double strands are denatured at 70 °C followed by another temperature decrease to 20 °C. Repeat this cycle three times.

**3** Digest a commercially available pET30a(+) vector with restriction enzymes *Nco*I and *Xho*I to insert an adaptor sequence.

**4** Prepare the reaction (50 μL total) by adding 1 μg of plasmid DNA, 5 μL of 10X NEBuffer, 1 μL of each restriction enzyme (*Nco*I and *Xho*I) and add ultrapure water to 50 μL. Incubate 1 h at 37 °C.

**Critical Step!** All restriction enzymes should be maintained on ice while in use. Temperature change might cause enzyme degradation and activity loss. The enzymes should be added last to the reaction mix to allow for uniform distribution.

5 Ligate double stranded adaptor sequence (Linker), carrying *Nhe*I and *Spe*I restriction sites, into *Xho*I and *Nco*I sites of a pET30a(+) vector to generate pET30L vector. Note that both restriction sites are preserved after ligation. The resulting cloning vector is referred to as pET30L.

6 Perform ligation reaction (20 μL total) by adding 2 μL of 10X T4 DNA Ligase Buffer, 0.02 pmol of vector DNA (pET30a(+)) and 0.06 pmol of insert DNA (double stranded linker DNA) and 1 μL of T4 DNA Ligase.

7 Add nuclease-free water to a total of 20 μL. Gently mix the reaction by pipetting up and down and microfuge briefly.

8 Incubate the reaction at 16 °C overnight or at room temperature for 10 min.

9 Heat inactivate at 65 °C for 10 min.

10 Chill the ligation mix on ice and transform 1–5 μL of the reaction into 50 μL high efficiency *E. coli* DH5α chemically competent cell following the protocol recommended by the manufacturer. Select the positive clones on LB plates supplemented with kanamycin (50 μg/mL).

**Critical Step!** Prepare the 50 °C water bath in advance to allow the autoclaved LB agar to cool down, but not start solidifying. After autoclaving leave the bottle in the bath until the temperature of the media reaches approximately 50 °C. Once the bottle can be held in your hands, add 1 mL of 50 mg/mL Kn stock under sterile conditions. Do not autoclave kanamycin as it is thermosensitive. Kanamycin should be filter-sterilized before use (pore size 0.22 μm). In the case of LB broth, incubation of the media in a water bath is not necessary, as the antibiotic can be added when the media cools down. When pouring the plates, take care to avoid bubbles. If bubbles form, use pipetman to draw it back into the pipet while the agar is still liquid.

11 Inoculate selected clones in LB (Kn) medium.

12 Culture overnight at 37 °C at 250 rpm.

13 Isolate the plasmids using QIAprep Spin Miniprep Kit. Check the concentration of isolated plasmid by nanodrop and on agarose gel.

**Critical Step!** The concentration of DNA can be estimated on agarose gel when compared to the standards used for DNA quantification (e.g., New England Biolabs DNA ladder). In addition, agarose gel provides information on possible contamination and/or degradation of the DNA. Nanodrop allows for determination of DNA concentration by reading absorbance at 260 nm. An absorbance ratio of 260 and 280 nm gives an estimate of the purity of the solution. Pure DNA and RNA solutions have $OD_{260}/OD_{280}$ values of 1.8 and 2.0, respectively. Contamination of a sample with protein or phenol will result in a value significantly less than the above numbers, typically $OD_{260}/OD_{280}$ of around 1.6.

**14** Perform the restriction reaction using *Nco*I and *Xho*I enzymes. Perform the restriction reaction using *Nhe*I and *Spe*I endonucleases to confirm the successful ligation of the linker sequences.

**15** Prepare the reaction (50 μL total) by adding 1 μg of plasmid DNA, 5 μL of 10X NEBuffer, 1 μL of each restriction enzyme (*Nhe*I and *Spe*I) and add ultrapure water to 50 μL.

**16** Incubate 1 h at 37 °C. Run 1% agarose gel to confirm the successful ligation of pET30a(+) and linker sequence.

**Critical Step!** As the insertion of the linker does not change the antibiotic resistance it is difficult to select positive clones where the ligation was successful, as both positive and negative clones grow on LB (Kn) plates. Therefore, posterior to plasmid isolation, digestion should be performed and positive clones sent for sequencing. In the case of pET30L, standard T7, and T7 term primers can be used for sequencing.

**17** Prepare LB (Kn) precultures of selected clones carrying pET30L, incubate overnight at 37 °C at 250 rpm. Prepare 10% glycerol stock and freeze at −80 °C.

**Critical Step!** Due to the high density of 100% glycerol it is difficult to pipet, therefore 80% glycerol solution should be prepared and autoclaved. This will allow accurate pipetting. In addition, it is recommended to cut the pipet tip to allow for easier pipetting. The sterility of all solutions and equipment should be maintained. Glycerol is toxic to bacteria at high concentrations so once added to bacterial culture, the culture should be immediately frozen. The glycerol stock of all strains should be made and stored at −80 °C in the case there is the need to return back to any of the previous cloning steps.

**18** Sequencing the plasmids with correct size from step 13 to prevent frameshifting mutation.

### Generate Vector pET30L-A

**19** Design the genes encoding recombinant silk-like proteins based on the sequence of the MaSp I protein of *Nephila clavipes* (Accession Numbers: P19837 and ACF19411.1, see 1).

**20** Design two individual blocks, a hydrophobic poly alanine rich block (A) and a hydrophilic glycine-rich block (B) (see example Figure 5B). The A block consists of one poly(GA) repeat (GAGAAAAAGGAG) responsible for β-sheet formation, whereas the B block is composed of four GGX repeats, separated by the GSQGSGR sequence.

**21** Optimize the codons for the production in E. coli and commercially obtain the sequence as the *Nhe*I/*Spe*I inset in pUC57 (Kn). Follow the manufacturer's instructions to dissolve the plasmids.

**22** Transform high efficiency *E. coli* DH5α chemically competent cell following the protocol recommended by the manufacturer. Select the positive clones on LB plates supplemented with kanamycin (50 μg/mL).

**23** Inoculate selected clones in LB (Kn) medium. Culture overnight at 37 °C at 250 rpm.

**24** Isolate the plasmids using QIAprep Spin Miniprep Kit.

**25** Check the concentration of isolated plasmid by nanodrop (see 13).

**26** Perform the restriction reaction using *Nhe*I and *Spe*I endonucleases to isolate silk-like monomers. Prepare the reaction (50 μL total) by adding 1 μg of plasmid DNA, 5 μL of 10X NEBuffer, 1 μL of each restriction enzyme (*Nhe*I and *Spe*I), and ultrapure water to 50 μL. Incubate 1 h at 37 °C.

**27** Stop the restriction reaction by incubating the mix at 80 °C for 20 min.

**28** Prepare the samples for gel electrophoresis by adding 1/6 volume of 6× DNA Loading Dye to the sample and 1% agarose gel (reagent setup).

**Caution!** EtBr is a known mutagen. Wear a lab coat, eye protection, and gloves when working with this chemical. Do not add EtBr while the gel is warm.

**29** Run the gel at 100 V until the dye line is approximately 75–80% of the way down the gel to visualize released monomer.

**30** Visualize the sample using Gel Imager.

**31** Perform the gel extraction using QIAquick Gel Extraction Kit following the procedure recommended by manufacturer.

**32** Analyze the concentration of DNA fragment by nanodrop (see 13), and store at −20 °C.

**33** Prepare LB (Kn) precultures of selected clones carrying pUC57-A, incubate overnight at 37 °C at 250 rpm.

**34** Prepare 10% glycerol stock and freeze at −80 °C.

**35** Digest pET30L with restriction enzyme *Nhe*I and *Spe*I (see 4). Run 0.9% agarose gel and gel extract the vector.

**36** Determine the plasmid concentration by nanodrop (see 13).

**37** Ligate pET30L and monomer sequence corresponding to A block (see 6), both previously digested with *Nhe*I and *Spe*I. The resulting cloning vector is referred to as pET30L-A.

**38** Transform high efficiency *E. coli* DH5α chemically competent cell (see 10–11) and select the cells on LB (Kn) plates.

**39** Isolate the pET30L-A plasmid (see 14).

**Critical Step!** Send selected clones for sequencing to confirm the incorporation of sequence A.

**40** Store successful *E. coli* clones at −80 °C (see 17).

## Generate Vector pET30L-AB

**41** Use the successful pET30L-A clones to perform digestion with *Spe*I.

**42** Stop the reaction by incubation at 80 °C for 15 min.

**43** To avoid the self-ligation of the pET30L-A vector digested with *Spe*I, remove 5′ phosphate from the DNA using Antarctic Phosphatase. To set up the reaction, add 1/10 volume of 10× Antarctic Phosphatase Reaction Buffer to 1–5 μg of pET30-A cut with *Spe*I.

**44** Add 1 μL of Antarctic Phosphatase (5 units) and mix.

**45** Incubate for 15 min at 37 °C.

**46** Heat inactivate at 80 °C for 2 min.

**47** Perform the ligation of pET30-A cut with *Spe*I and sequence corresponding to B block cut with *Spe*I and *Nhe*I.

**48** Transform high efficiency *E. coli* DH5α chemically competent cell and select the cells on LB (Kn) plates. Step-by-step directional ligation: ligate B block into the vector followed by stepwise ligation of the A block. By using a step-by-step directional ligation approach, direct control over the assembly of monomeric genes into complex sequences can be achieved. (Ex) AB, $A_3B$, $AB_3$, $(AB)_n$

**49** Isolate the plasmid and send selected clones for sequencing to confirm the incorporation of sequence B. Store successful clones at −20 °C.

**Collaboration Point!** Inform Simulation Team of final protein sequences. In future, create protein sequences that the Simulation Team predicts will have the desired properties.

## Gene Expression and Protein Production

**50** For the expression and protein production, transform the successful clones into *E. coli* BL(DE3).

**51** Inoculate 50 mL LB (Kn) medium with *E. coli* BL(DE3) pET30L-AB to prepare preculture.

**52** Detailed culture condition: plate and incubate in Hyper Broth medium supplemented with kanamycin and glucose. Grow cells at 37 °C and 250 rpm.

**53** Induce gene expression with 1 mM IPTG (isopropyl β-D-1-thiogalactoside) when the optical density, $OD_{600}$ is between 1.2 and 1.5.

## Protein Extraction and Purification

**54** Harvest cells 4 h after induction using centrifuge at 10 000 g.

**55** Perform protein purification under denaturing conditions on Ni-NTA resin.

**56** Resuspend the harvested cells in denaturing buffer (see Reagent Setup).

**57** Lyse the cells by stirring for 30 min at RT.

**58** Centrifuge cell lysate at 9000 rpm, 4 °C for 25 min to separate insoluble cell fragments and soluble proteins.

**59** Mix the cell lysate with Ni-NTA resin in 4 to 1 ratio and leave overnight under constant agitation.

**60** Load the lysate–resin mixture carefully into an empty column with the bottom cap attached.

**61** Remove the bottom cap and collect flow-through.

**62** Wash four times with denaturing buffer (pH 8.0), twice with denaturing buffer (pH 6.7), and two times with denaturing buffer (pH 5.3). To minimize carbamylation of silk peptides by urea, all buffers are prepared fresh daily.

**63** Elute the proteins using the denaturing buffer (pH 4.5).

**64** Analyze collected fractions by SDS-PAGE on NuPage 4%–10% bis-tris gels (follow the recommendations of manufacturer for sample preparation) and use Novex Sharp pre-stained protein standard.

**65** Run the gel for 40 min at 200 V and stain the gel afterward with Simply Blue (follow the instructions recommended by manufacturer).

**66** Transfer the proteins into Slide-A-Lyzer Cassette with MWCO 2000 Da.

**Critical step!** Load the syringe without the needle, once the syringe is loaded, put the needle and fill the cassette. It is important to avoid the use of needle to minimize shear.

**67** Dialyze the proteins against water for 3 days at 4 °C with steering on steering plate.

**68** Change the water every 3 h first day and two times per day next 2 days.

**69** Lyophilize the dialyzed proteins. Store proteins at room temperature.

**70** See Figure 7 for sample results.

## Bioinspired Spinning Process: Timing 1–24 h

### Vertical Wet Spinning

**1** Measure and record the weight of a single 1.5 mL micro-centrifuge tube. Then add an appropriate amount of lyophilized protein followed by the corresponding amount of the protein solvent into the tube to generate a solution of concentration 15 wt % or greater. Measure and record the weight after each addition to have a record of the concentration.

**Caution!** Several protein solvents have important safety considerations. Every chemical should be handled with appropriate chemical care, following safety guidelines. Both HFIP and formic acid are classified as high health hazards and

corrosive. LiBr solutions are strongly exothermic upon mixing and also considered a health hazard. Be sure to wear appropriate protective apparel and operate in a properly ventilated environment.

**Critical Step!** The more caustic solvents listed are generally highly polar, denaturing, and/or chaotropic solutions that enable the highest protein concentrations to be achieved. Solubility up to functional spinning concentrations (15% w/v) is achievable in pure water, so it is at the discretion of the investigator to decide the best solvent to use based on the concentration demands of the setup.

**2** Mix the solvent and lyophilized protein by gentle pipetting and vortexing. Be sure to allow the solvent sufficient time to interact with the protein. At a minimum, spinning concentrations can be obtained in a manner of 1–5 min after mixing. Full dissolution of the protein may require additional time, up to 24 h at room temperature.

**Critical Step!** If gelation is occurring, mixing and dissolution can be done at 4 °C. It is important that the solvent is equilibrated to that temperature prior to mixing.

**3** Centrifuge the solution in the MiniSpin at 12 kRPM for 2 min to separate any undissolved protein from the solution. Remove the supernatant via pipetting. This is the solution used for spinning.

**Critical step!** The protein sample must be completely solubilized before spinning. The presence of insoluble aggregates in the solution may clog the needle of the spinning apparatus.

**4** If there is an undissolved protein pellet remaining after removing the supernatant, place the tube in an 80 °C oven for 20 min to remove the remaining water. Remove the tube from the oven and leave in room temperature for 20 min to equilibrate. Then, measure and record the weight of the tube with the remaining protein fraction to determine the true concentration of the spinning dope with the following equation.

$$\text{spinning concentration } (\%\text{w/v}) = \frac{(m_{\text{initial protien}} - m_{\text{remaining protien}})}{V_{\text{solvent}}} 100$$

**5** Load a minimum of 25 μL of spinning solution into a 250 ul Hamilton Gastight syringe (Hamilton Company, NV) with a needle gauge no lower than 22 in order to generate sufficient shearing to initiate fiber assembly.

**Critical Step!** Higher gauge needles with lower inner diameters produce higher degrees of shear, but also reduce the diameter of the extruding fibers. Fibers can be generated consistently using up to 33 gauge needles, but the flow rates may need to be adjusted (5 μL/min instead of 15 μL/min) to compensate for the overall reduction in the volume of the fibers, which may not be able to overcome the inertia of the solvent when injected at a high rate.

6 Extrude the silk solution using the syringe pump system as shown in Figure 6A, which is in a vertical setup. The speed of the pump can be managed between 5 and 15 μL/min as discussed above. The collection bath is the reservoir for the chosen coagulation solvent.

**Critical Step!** The selection of the coagulation bath depends on many factors, but most notable is the coagulation time. The advantage of a slower coagulation is that the spinning solution has time to clear the needle tip before becoming an insoluble fiber, which reduces the chance of clogging. However, the time the solution spends in the bath would necessarily increase, in the case of using a weaker coagulant. Of the solvents mentioned, their order of fastest to slowest coagulation of silk are methanol > acetone > isopropanol.

**Critical step!** Make sure to start the pump before the needle contacts the coagulation bath; otherwise, coagulation of the protein may start within the needle, resulting in clogging.

**Collaboration Point!** Work with Block Co-Polymer Synthesis team on protein structure that allows for spinable fiber. Certain sequences do not always form uniform fibers.

**Collaboration Point!** Inform Simulation Team of fluid flow and needle diameter during spinning so that computational model incorporates accurate shear properties.

7 Collect the generated fibers from the bath by transferring them into scintillation vials filled with the coagulation solvent for storage until needed for further observation and analysis.

**Critical step!** Be careful when transferring the fibers. Try to avoid inadvertent drying or tangling when they are transiently removed from the solvent.

## Fiber Mechanical Testing

8 Transparent film frames are used to handle the individual fibers for further observation (microscopic image, tensile testing) (see example in Figure 6B).

9 Using forceps and scissors, cut the sample fibers into approximately 2 cm long pieces and mount onto the transparent film frames where double-sided tape is already attached on each side to hold the fiber in place.

10 Frame mounted fibers are easily manipulated for bright field microscope imaging without damage by imaging through a Petri dish containing several framed fibers. The diameter is determined using ImageJ assigning each pixel a physical length based on objective magnification and resolution of the microscope camera. Five measurements of the diameter are taken across the image and averaged to get a value for mechanical property calculations.

11 The frames can then be mounted onto clips attached to the Instron Microtester. Once secure, the sides of the frame must be carefully burned away using a soldering iron before performing stretching.

**Critical step!** Scissors can be used the cut the frame, but we found that the torque introduced to the frame from cutting the plastic often led to failures of the fiber prior to testing. The soldering iron does not introduce any bending moments on the fiber; however, the heat generated by the soldering iron can cause problems if it is placed too close to the fiber.

**12** Under dry testing conditions, the fibers should be stretched at a slow strain rate (0.01/s) until the fiber breaks. If the fibers are wetted or submerged, then greater rates (up to 0.05/s) may be applied without causing failures strictly due to the strain rate. Record the force as a function of either time or strain (both are interchangeable if the strain rate is fixed).

**13** Once collected, the force data can be normalized to stress values using the diameter measurements from beforehand. For fibers, a cylindrical volume is assumed, which leads to a circular cross-sectional area ( $A=\pi\left(\frac{d}{2}\right)^2$, where A is cross-sectional area and d is diameter).

**14** See Figure 7 for sample results.

## ANTICIPATED RESULTS

### Simulation

In Figure 4A, quantitative network analysis is used to visualize the results after equilibration as seen by the sample output frames in the top row. Hydrophobic, hydrophilic, and histidine tags are colored in red, blue, and green, respectively. Water beads are not shown. The bottom row shows results of the network analysis via a node-bridge diagram on these same frames. It can be seen here that only the $H(AB)_2$ sequence has network connectivity. This means that a 1:1 ratio of the hydrophobic and hydrophilic blocks results in a connected polymer network while a lower hydrophobic to hydrophilic ratio does not. A “node” is defined as the center of mass of an individual crystal (A block) aggregate. A “bridge” indicates a physical connection via amorphous phase (B and H blocks) between the individual crystal aggregates. In Figure 4B, quantitative network analysis is also used to visualize the effect of applying shear to the $H(AB)_{12}$ construct. In the top row, sample output frames before (left) and after (right) applying shear from simulations are shown. Water beads are not shown. In the bottom row, the results of the network analysis, node-bridge diagram, are shown for the same frames; it can be seen that applying shear results in the alignment of the connections with the direction of the shear. The thickness of the black lines is linearly proportional to the number of connections.

Figure 4C shows the time evolution of quantitative network structure. During both equilibration and shear flow, the median sizes of nodes do not vary a lot among all sequences (a), but longer copolymers have significantly more bridges (b). During shear flow, copolymers are stretched under shear flow and the cross-links are continuously broken and reformed until the steady state is reached for the simulated polymer properties in a–c. Although the network connectivity is not increased (in fact, it drops by a little due to shearing-induced perturbation) during shear flow (b), the network conductance for the long

sequences ($n$ = 8 and 12) is greatly enhanced because of the shearing-induced polymer alignment (c). The discontinuity between shear flow and tensile stretching periods in a–c is due to the stress relaxation simulation (results not reported here) in between the two periods. Interestingly, the shorter sequences ($n$ = 2 and 4) form weak networks that are not able to maintain structural integrity due to the shearing-induced perturbation and, therefore, their network conductance drops to zero (c). The distribution histograms of the numbers of connected nodes for $H(AB)_4$ and $H(AB)_{12}$ after shear flow are shown in (d). The $H(AB)_{12}$ sequence has node connectivity distribution peaks at higher values (~5) than $H(AB)_4$ (~2).

Figure 4D shows the simulated fiber stress–strain curves for mechanical tensile tests of polymer networks before and after shear flow. Before the shear flow, all networks are weak without stiffening at high tensile strains (low maximum stress), and it follows the trend that longer sequences are stronger with higher Young's modulus. After the shear flow, the network Young's modulus remains the same except that the Young's modulus for the two shorter sequences ($n$ = 2 and 4) drops to zero, and the longer sequences are stronger with much higher maximum stress and stiffening behavior at high tensile strains. Note that due to the coarse-grained nature of the DPD model, the absolute stress values are approximate and we will focus on the relative values and comparisons among different protein sequences in a qualitative way. This factor, in addition to the presence of solvents in simulations also leads to simulated ultimate network strains much higher than that of common silk fibers (10–30%).[29] Therefore, simulation predicted quantitative Young's modulus values are not to be compared with the later experimental tensile testing results directly. Section (c) shows the schematic of the $H(AB)_{12}$ network deformation process under various tensile strains. The red clusters (stiff cross-linked $\beta$ Sheet crystals) slightly deform during stretching and maintain their structural integrity, whereas the blue bridges (soft amorphous regions) deform significantly or even break to release the applied stress.

## Experimental: Synthesis and Fiber Processing

Switching from simulation to experiment, sample results for biosynthesis and processing can be seen in Figure 7. Figure 7A features an SDS-PAGE gel showing purification steps of $H(AB)_{12}$ (left) and $H(AB)_{12}$ protein after dialysis (right) and Figure 7B shows SEM images of the separated $H(AB)_{12}$ fibers after rehydration, demonstrating a relatively uniform fiber diameter. In (C), experimental FTIR-derived percentages of the protein secondary-structure contents for the three peptides sequences are plotted, and (D) features AFM nanoindentation measurements, the Young's modulus maps (left) and the corresponding Young's modulus distributions (right) of $H(AB)_{12}$ fiber. The $H(AB)_{12}$ sequence forms spherical aggregates with diameters of 15–20 nm (right). The AFM height plots provide quantitative values of the aggregate sizes measured along the red line in the AFM images. Finally, (E) is the mechanical data for fibers pulled using the Instron. A representative stress–strain curve from mechanical tensile testing on various silk fiber samples confirms high fiber robustness and Young's modulus.

## SUMMARY

The ability to synthesize silk fibers with mechanical properties tailored for specific biomedical applications without experimental iteration would be an amazing realization of the value of predictive material design. This protocol outlines one approach for using computer simulation along with experimental protein synthesis and fiber spinning to establish a method for predictive design of silk fibers.

The computer model first predicts that a 1 to 1 ratio of hydrophobic to hydrophilic blocks is needed within the protein structure in order to have a connected polymer network. Furthermore, maintaining this ratio, the model predicts that longer copolymer chains have more bridges and under shear flow, have higher network conductance when compared to shorter chains. Before the shear flow, all networks are relatively weak during tensile testing; however, longer sequences are stronger with higher Young's modulus. After the shear flow, the network Young's modulus remains the same for the longer sequences but a much higher maximum stress and corresponding higher tensile strain are reached; for the shorter chains that lack network conductance, the modulus goes to zero.

Experimental protein synthesis and fiber spinning are then used to create copolymers and fibers representative of these simulated materials. Mechanical properties determined by AFM and tensile testing are used to either confirm or disprove that the simulation predictions are accurate. In general, experimental tests and trends agree with results predicted by simulation. However, at this point, the absolute stress values predicted by the DPD model are approximate and only relative values among different protein sequences can be compared to when determining whether or not the simulation and experiments agree.

Although the model system does not allow for full predictive abilities at this time, it does provide a better understanding of how copolymer sequence, chain length, and shear affect fiber mechanical strength, which allows for more rational design. This protocol also provides insight into which steps in the process the three teams can continue to iterate around in order to improve the correlation between the properties resulting in the simulation and in the actual biomaterial. Methods are presented that are relevant to biomaterials other than just silk fibers.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors thank NIH U01 EB014976-05 (M.J.B., D.L.K., J.Y.W.) for support of this work. N.R. is a Howard Hughes Medical Institute International Student Research fellow. We also thank SITEP team members W. Huang and J. Yeo for reviewing and providing their feedback on our manuscript

## References

1. Dinjaski N, Ebrahimi D, Ling S, Shah S, Buehler MJ, Kaplan DL. Integrated Modeling and Experimental Approaches to Control Silica Modification of Design Silk-Based Biomaterials. ACS Biomater Sci Eng. 2016; doi: 10.1021/acsbiomaterials.6b00236

2. Dinjaski N, Ebrahimi D, Qin Z, Giordano JEM, Ling S, Buehler MJ, Kaplan DL. Predicting rates of in vivo degradation of recombinant spider silk proteins. J Tissue Eng Regener Med. 2017; doi: 10.1002/term.2380

3. Huang W, Ebrahimi D, Dinjaski N, Tarakanova A, Buehler MJ, Wong JY, Kaplan DL. Synergistic Integration of Experimental and Simulation Approaches for the de Novo Design of Silk-Based Materials. Acc Chem Res. 2017; 50(4):866–876. [PubMed: 28191922]

4. Gronau G, Krishnaji ST, Kinahan ME, Giesa T, Wong JY, Kaplan DL, Buehler MJ. A review of combined experimental and computational procedures for assessing biopolymer structure-process-property relationships. Biomaterials. 2012; 33(33):8240–55. [PubMed: 22938765]

5. Jacobsen MM, Tokareva OS, Ebrahimi D, Huang W, Ling S, Dinjaski N, Li D, Simon M, Staii C, Buehler MJ, Kaplan DL, Wong JY. Effect of terminal modification on the molecular assembly and mechanical properties of protein-based block copolymers. Macromol Biosci. 2017:1700095.

6. Jin HJ, Kaplan DL. Mechanism of silk processing in insects and spiders. Nature. 2003; 424(6952): 1057–1061. [PubMed: 12944968]

7. Schafer A, Vehoff T, Glisovic A, Salditt T. Spider silk softening by water uptake: an AFM study. Eur Biophys J. 2008; 37(2):197–204. [PubMed: 17851658]

8. Altman GH, Diaz F, Jakuba C, Calabro T, Horan RL, Chen JS, Lu H, Richmond J, Kaplan DL. Silk-based biomaterials. Biomaterials. 2003; 24(3):401–416. [PubMed: 12423595]

9. Jelinski LW. Establishing the relationship between structure and mechanical function in silks. Curr Opin Solid State Mater Sci. 1998; 3(3):237–245.

10. Chao PHG, Yodmuang S, Wang XQ, Sun L, Kaplan DL, Vunjak-Novakovic G. Silk hydrogel for cartilage tissue engineering. J Biomed Mater Res, Part B. 2010; 95b(1):84–90.

11. MacIntosh AC, Kearns VR, Crawford A, Hatton PV. Skeletal tissue engineering using silk biomaterials. J Tissue Eng Regener Med. 2008; 2(2–3):71–80.

12. Numata K, Hamasaki J, Subramanian B, Kaplan DL. Gene delivery mediated by recombinant silk proteins containing cationic and cell binding motifs. J Controlled Release. 2010; 146(1):136–143.

13. Lammel A, Schwab M, Hofer M, Winter G, Scheibel T. Recombinant spider silk particles as drug delivery vehicles. Biomaterials. 2011; 32(8):2233–2240. [PubMed: 21186052]

14. Holland C, Terry AE, Porter D, Vollrath F. Natural and unnatural silks. Polymer. 2007; 48(12): 3388–3392.

15. Jin HJ, Fridrikh S, Rutledge GC, Kaplan D. Electrospinning bombyx mori silk with poly(ethylene oxide). Biomacromolecules. 2002; 3:1233–1239. [PubMed: 12425660]

16. Wang M, Jin HJ, Kaplan DL, Rutledge GC. Mechanical properties of electrospun silk fibers. Macromolecules. 2004; 37(18):6856–6864.

17. Guerette PA, Ginzinger DG, Weber BHF, Gosline JM. Silk properties determined by gland-specific expression of a spider fibroin gene family. Science. 1996; 272(5258):112–115. [PubMed: 8600519]

18. Lazaris A, Arcidiacono S, Huang Y, Zhou JF, Duguay F, Chretien N, Welsh EA, Soares JW, Karatzas CN. Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. Science. 2002; 295(5554):472–476. [PubMed: 11799236]

19. Rammensee S, Slotta U, Scheibel T, Bausch AR. Assembly mechanism of recombinant spider silk proteins. Proc Natl Acad Sci USA. 2008; 105(18):6590–6595. [PubMed: 18445655]

20. Lin SC, Ryu S, Tokareva O, Gronau G, Jacobsen MM, Huang WW, Rizzo DJ, Li D, Staii C, Pugno NM, Wong JY, Kaplan DL, Buehler MJ. Predictive modelling-based design and experiments for synthesis and spinning of bioinspired silk fibres. Nat Commun. 2015; 6:6892. [PubMed: 26017575]

21. Rabotyagova OS, Cebe P, Kaplan DL. Self-assembly of genetically engineered spider silk block copolymers. Biomacromolecules. 2009; 10(2):229–36. [PubMed: 19128057]

22. Koeppel A, Holland C. Progress and Trends in Artificial Review Silk Spinning: A Systematic Review. ACS Biomater Sci Eng. 2017; 3(3):226–237.

23. Groot RD, Warren PB. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. J Chem Phys. 1997; 107(11):4423.

24. Chen S, Phan-Thien N, Fan X-J, Khoo BC. Dissipative particle dynamics simulation of polymer drops in a periodic shear. J Non-Newtonian Fluid Mech. 2004; 118:17.

25. Tokareva O, Michalczechen-Lacerda VA, Rech EL, Kaplan DL. Recombinant DNA production of spider silk proteins. Microb Biotechnol. 2013; 6(6):651–63. [PubMed: 24119078]

26. Meyer DE, Chilkoti A. Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. Biomacromolecules. 2002; 3(2):357–67. [PubMed: 11888323]

27. Higashiya S, Topilina NI, Ngo SC, Zagorevskii D, Welch JT. Design and preparation of beta-sheet forming repetitive and block-copolymerized polypeptides. Biomacromolecules. 2007; 8(5):1487–97. [PubMed: 17388563]

28. Teule F, Cooper AR, Furin WA, Bittencourt D, Rech EL, Brooks A, Lewis RV. A protocol for the production of recombinant spider silk-like proteins for artificial fiber spinning. Nat Protoc. 2009; 4(3):341–55. [PubMed: 19229199]

29. Agnarsson I, Kuntner M, Blackledge T. A Bioprospecting finds the toughest biological material: extraordinary silk from a giant riverine orb spider. PLoS One. 2010; 5(9):e11234. [PubMed: 20856804]

30. Tokareva OS, Glettig DL, Abbott RD, Kaplan DL. Multifunctional spider silk polymers for gene delivery to human mesenchymal stem cells. J Biomed Mater Res, Part B. 2015; 103(7):1390–401.

31. Huang W, Krishnaji S, Hu X, Kaplan D, Cebe P. Heat Capacity of Spider Silk-like Block Copolymers. Macromolecules. 2011; 44(13):5299–5309. [PubMed: 23869111]

32. Omenetto FG, Kaplan DL. New opportunities for an ancient material. Science. 2010; 329(5991): 528–31. [PubMed: 20671180]


**Figure 1.**
Depiction of the interactions between the three sectors of the silk collaboration, SITEP (Silk Integrative Theory Experiment Project).

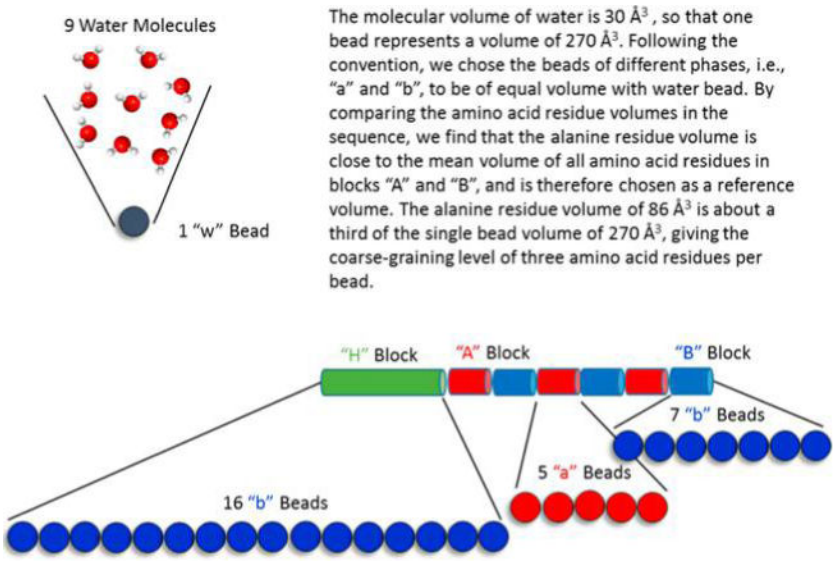



**Figure 2.**

Diagram of coarse-grained beads used in simulation. The figure depicts a schematic of coarse graining scheme used to simulate the formation of micelles. Three types of coarse grained beads are used: hydrophobic ("a"), hydrophilic ("b"), and water ("w") beads. For definition of each block, see Table 1.

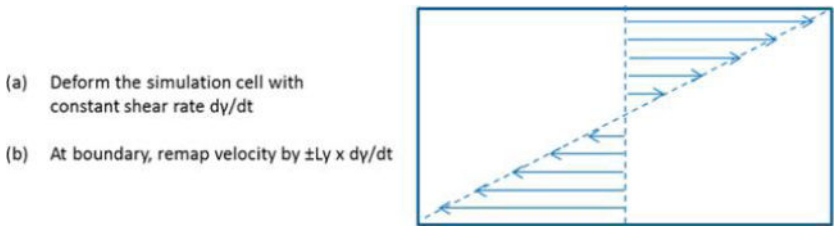



**Figure 3.**
Lees Edwards boundary condition for applying shear in the simulation.

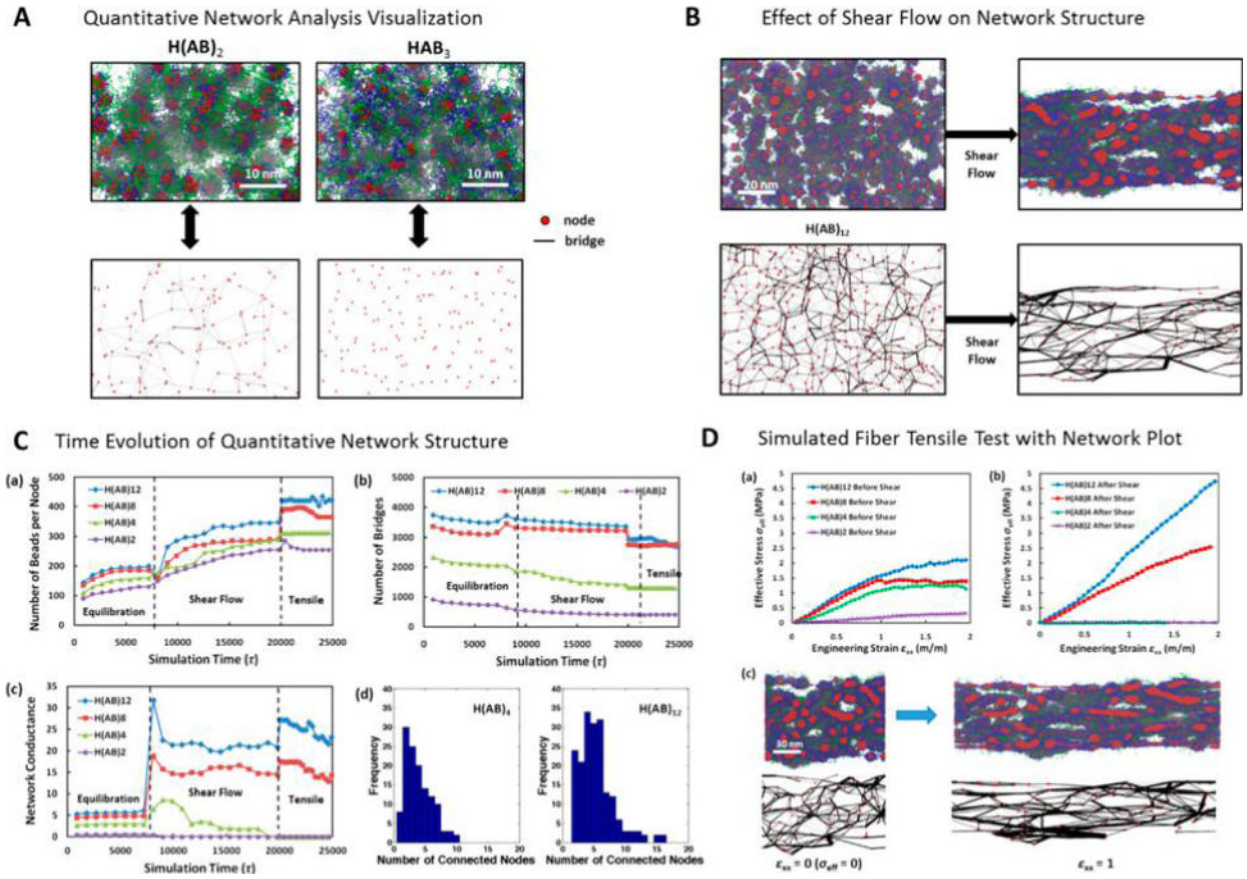



**Figure 4.**

Sample results for simulation. (A) Quantitative network analysis to visualize the results after equilibration. (B) Quantitative network analysis to visualize effect of applying shear on $H(AB)_{12}$ construct. Top row: Hydrophobic, hydrophilic, and histidine tag are colored in red, blue, and green. Bottom row: "Node" indicates center of mass of individual crystal (A block) aggregates. "Bridge" indicates physical connection via amorphous phase (B and H blocks) between individual crystal aggregates. (C) Time evolution of quantitative network structure. (a) Median number of 'a' beads per node as a measure of the aggregate size. (b) Total number of bridges as a measure of the connectivity. (c) Polymer network conductance as a measure of the degree of protein alignment. (d) Distribution histograms of the numbers of connected nodes for $H(AB)_4$ and $H(AB)_{12}$ after shear flow. (D) Simulated fiber tensile test. Stress–strain curves for mechanical tensile tests of polymer networks (a) before and (b) after shear flow. (c) Schematic of the $H(AB)_{12}$ network deformation process under various tensile strains. Reproduced with permission from ref[20]. Copyright 2015 Nature Publishing Group.

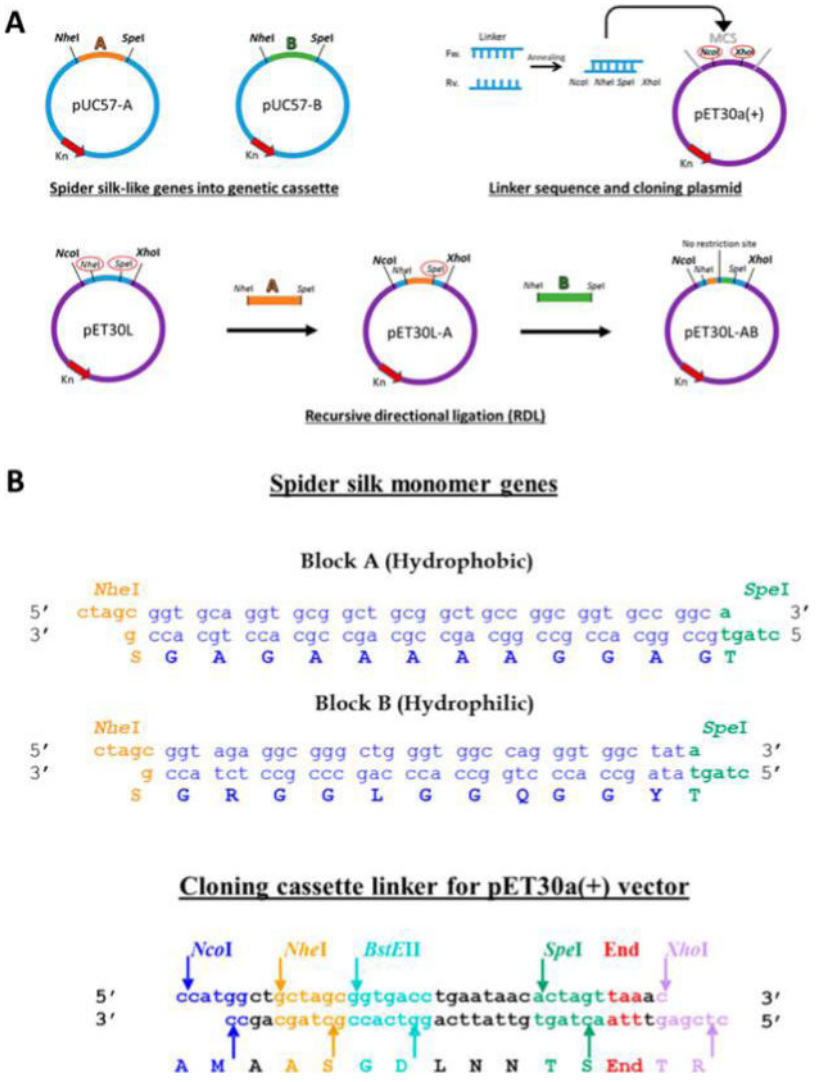



**Figure 5.**

Synthesis summary. (A) Schematic representation of the recombinant DNA strategy used to make spider silk block copolymers. *Nco*I, *Nhe*I, *Spe*I, *Xho*I are restriction enzymes. Dark blue indicates pET30a(+) plasmid, light blue indicates cloning linker, orange indicates silk hydrophobic block A, and green indicates silk hydrophilic block B. Reproduced with permission from ref[30]. Copyright 2015 John Wiley & Sons. (B) Sequence of hydrophobic poly alanine-rich block A and a hydrophilic glycine-rich block B and cloning cassette linker used as an example for this paper. Reproduced with permission from ref[21]. Copyright 2009 American Chemical Society.

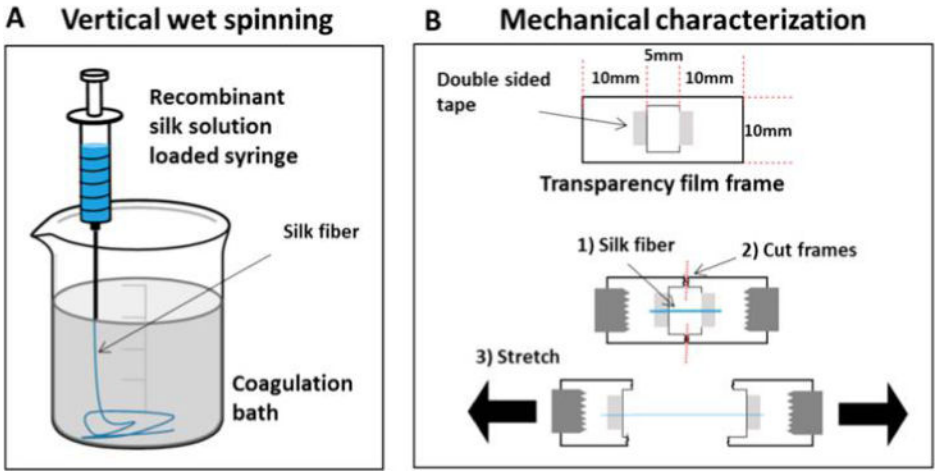



**Figure 6.**

(A) Vertical wet spinning setup to fabricate recombinant silk block copolymer fiber. (B) Transparency film was used to make frames with certain dimensions as shown in the figure to hold single-fiber samples for mechanical characterization. Double-sided tape was used to easily fix the fiber on the frame to make it easier to transfer samples for diameter measurement via microscope and mechanical testing. Once the frame (with fiber sample) was loaded onto the Instron, the frame was carefully cut using scissors or soldering iron and then tensile testing was performed.

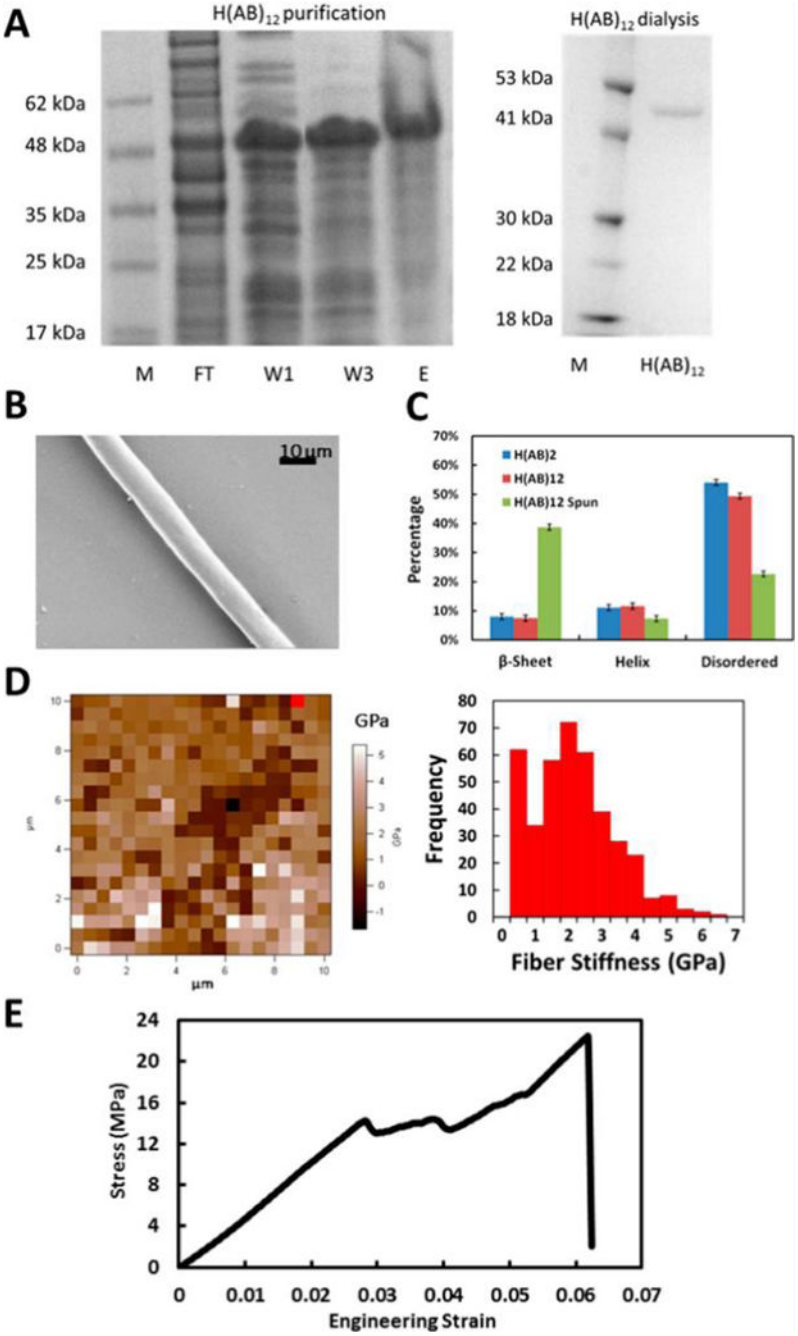



**Figure 7.**

Sample results for biosynthesis and processing. (A) SDS-PAGE gel showing purifications steps of $H(AB)_{12}$ (left) and $H(AB)_{12}$ protein after dialysis (right). (B) SEM images of the separated $H(AB)_{12}$ fiber after rehydration. (C) FTIR-derived percentages of the protein secondary-structure contents. (D) AFM nanoindentation: The Young's modulus maps (left) and the corresponding Young's modulus distributions (right) of $H(AB)_{12}$ fiber. (E) Instron: Stress–strain curve from mechanical tensile testing on various silk fiber sample. (B–E) Reproduced with permission from ref[20]. Copyright 2015 Nature Publishing Group.

**Table 1**

Amino Acid Sequences of the Investigated Silklike Protein Blocks

| block | silklike protein amino acid sequence |
|---|---|
| “A” (hydrophobic), MW = 1247 Da | G A G A A A A A G G A G T S |
| “B” (hydrophilic), MW = 1960 Da | Q G G Y G G L G S Q G S G R G G L G G Q T S |
| “H” (hydrophilic), MW = 5239 Da | M H H H H H H S S G L V P R G S G M K E T A A A K F E R Q H M DS P D L G T D D D D K A M A A S |