**MATERIALS AND METHODS**

**Sample collection**

Live larvae and pupae of *Parapsyche elsis* Milne, 1936 were collected from the South Fork of the Provo River (40.347519 N, 111.547645 W) between May and August 2018 and transported to either Brigham Young University or the University of Utah for further analysis. Silk capture nets were collected in the field by removing them from rocks and carefully placing them directly onto glass cover slips for SEM and elemental analysis. For the TEM analysis, *Hesperophylax* caddisworms were collected alive from Upper Red Butte Creek in Salt Lake City, UT (40.774202 N, 111.817752 W).

**40°46'27.1"N 111°49'03.9"W**

**Scanning Electron Microscopy/Energy Dispersive X-Ray Spectroscopy (SEM/EDS)**

All SEM and EDS analyses were conducted on an Apreo C Low Vac (Thermo Scientific) system at the Brigham Young University Microscopy Lab. For topographic imaging, silk nets on glass cover slips were Pt sputter coated. For non-quantitative elemental analysis, uncoated silk nets on glass coverslips were analyzed by EDS to compare concentrations of P, S, C, Si, Ca, Al, Fe, Mg, and Zn. Both line scans and 2D maps were performed with the TEAM EDS Software Suite (EDAX, Inc.) using an Octane Plus detector (EDAX, Inc.). For the line transect, a line scan was conducted at 0.5 μm resolution perpendicular to the long axis of a single fiber over a distance of 57.5 μm, which extended onto the glass cover slip on both sides of the fiber for background estimates. The resulting counts were then plotted using ggplot2 v3.0 in R v3.5.1[1] and the line plots were smoothed using locally estimated scatterplot smoothing (loess). For the 2D Map, EDS was performed on net junctions at 15 kV with a 200 μs dwell time over 55 frames.

**Transmission electron microscopy (TEM) of silk-precursors**

Silk gland dissection. Caddisworms were collected from Red Butte Creek (*Hesperophylax*) and the Provo River (*Parapsyche*) and kept in an aerated aquarium at 50 ºC. Caddisworms were anesthetized for 5 min in 30 mL water containing 15 drops ethyl ether, then pinned to a dissection tray ventral side up. An incision was created from the anus to just below the head, exposing the silk gland. The gland was pinched off with forceps at the anterior end, gently removed from the body without stretching, and rinsed in DI water. The anterior end of the gland was placed into an Eppendorf tube containing 200 L DI water. The gland was cut with scissors just below the forceps. The cut anterior end of the gland remained submerged and silk was allowed to flow out for 1 min before the gland was removed from the tube. The tube was closed and the sample mixed by gentle inversion 3-4 times.

Negative Staining. Silk gland samples (3.5 L) were applied to a carbon coated EM grid. After 30 s, a piece of damp filter paper was used to blot excessive sample from the grid, which was then placed sample side down onto a drop of DI water for 2 s. Excess water was wicked away using damp filter paper. Uranyl Acetate (1%, 3.5 L) was applied to the grid. After 30 s, the uranyl acetate solution was blotted from edge with damp filter paper and the grid was allowed to dry.

Imaging and Image Averaging. Samples were imaged using aTecnai F20 or aTecnai T12 TEM (FEI Inc.). Particles from the F20 images were averaged using *cis*TEM (cisTEM.org) using 120 fields containing over 1600 particles. Images from the F20 and F12 were used to measure particles using ImageJ.

**Elemental analysis by ICP-MS**

Samples of silk gland lumenal contents for ICP-MS were prepared as previously described.[2]

Briefly, freshly dissected silk glands were rapidly frozen in liquid N2 and lyophilized. The silk in the lumen of the silk gland was manually separated from the surrounding silk gland tissue with a pair of fine forceps.

Silk fiber nets were collected and kept intact in the field on glass coverslips. The silk samples were digested using a 5:1 mixture of 70% OPTIMA nitric acid and 30% ultrapure hydrogen peroxide overnight, then heated to dryness.  Dried samples were then resuspended in 2 mL 2% OPTIMA nitric acid for ICP-MS run. The average concentrations of two nets and two silk gland samples were reported. The difference between the two samples were less than 10% for all elements.

An Agilent 7900 ICP-MS was operated in helium (He) collision cell gas mode for all measurements.  In this work, Ca is measured using the isotope 44Ca.  To eliminate possible interference from 88Sr++, a calibration standard without Sr was used (Agilent Environmental Calibration Standard) and the run was done under half mass conditions applying the empirically determined interference equation: Mc(44) = [(M(44)\*1) – (M(43.5)\*10.9)]; where M(43.5) reflects the concentration of 87Sr++. An Agilent SPS4 autosampler was used to deliver the samples from a 96 position deep-well polypropylene microplate.  Agilent ICP-MS MassHunter 4.4 software allows auto-tune of all lens within the instrument.  Run parameters were as follows:  RF power 1600 W, sample depth 8.0 mm, carrier gas 1.05 L/min, sample flow rate = 0.2 mL/min, cell gas flow rate = 5 mL/min.

Calibration standards and samples were prepared in an acid matrix of 2% OPTIMA grade nitric acid. Agilent Environmental Calibration Standard was used to determine Ca concentration, Agilent Multi-Element Calibration Standard-4 was used to determine P concentration, and Agilent Multi-element Calibration Standard 2A was used to determine the concentration of the other elements.  All calibration curves had at least eight-points.  Agilent Germanium and Scandium Standard(s) were added online to standards, blanks and samples for use as internal standards to correct for potential sample matrix and/or nebulization effects.

**H-fibroin gene sequencing and assembly**

The H-fibroin gene consists of long repetitive regions and the first full length H-fibroin region was only recently assembled.[3] Previous efforts that relied solely on short read DNA sequencing were unsuccessful. Therefore, long-read, single-molecule Oxford Nanopore DNA sequencing was used to generate the raw data for H-fibroin gene sequence. First, a single larva was flash frozen in liquid N2 and high molecular weight DNA was extracted using an Agilent DNA extraction kit. After visualizing the DNA length using pulse field gel electrophoresis to ensure a high-quality DNA extraction, a sequencing library was prepared with the Oxford Nanopore LSK-109 library prep kit. The manufacturer’s instructions were followed with the following modifications: (1) following end repair and bead cleanup, DNA was eluted for 30 mins in a 37° C water bath, and (2) following adapter ligation, DNA was eluted for 30 min in a 37° C water bath. The library was then sequenced on two MinION FLO-MIN106 flow cells for 48 hr each using MinKnow v.18.05.5. Base calling was conducted using Albacore v.2.3.3.

The H-fibroin gene was assembled using Canu v.1.7 [4]. The conserved N and C termini were identified via BLAST [5] using existing *H. augstipennis* H-fibroin sequences on GenBank (accession numbers AB354591.1 and AB354592.1). Then the raw Nanopore reads were mapped to the gene sequence using Minimap2 [6] and the assembly/sequencing errors were corrected using Nanopolish v.0.10.2.[7] Open reading frames were identified in the H-fibroin gene using the Translate tool on the ExPASy Bioinformatics Resource portal. Three small sequencing errors were apparent from transitions of the H-fibroin ORF from one reading frame to another. The sequence was manually corrected using the highly repetitive nature of the H-fibroin.

**Tandem Mass Spectrometry**

Silk protein digestion. Silk fiber precursors were isolated from dissected silk glands by lacerating the silk gland and draining the contents into Eppendorf tubes. The isolate silk gland protein samples were digested by three different methods: i.) The sample was first heat denatured for 10 min at 95 ºC, then chilled on ice for 5 min. The proteins were then digested overnight at 38 ºC with Trypsin/LysC mixture in a 1:20 ratio of trypsin to sample volume. After the overnight digestion, the sample was acidified with 1% formic acid to a pH of 2-3. The peptide sample was then concentrated to 5 µL in a speedvac. ii.) The sample was first reduced with 5 mM dithiothreitol (DTT) for 45 min at 60˚C, then alkylated with 10 mM iodoacetamide (IAA) for 30 min at room temperature in the dark. Excess IAA was neutralized with 5 mM DTT and incubated for 45 min at 60˚C. The proteins were then digested overnight at 38 ºC with a Trypsin/LysC and Protease Max mixture (1:20:2 ratio of Trypsin/LysC: sample: 1% Protease Max). After the overnight digestion the sample was acidified with 1% formic acid to a pH of 2-3. The peptide sample was then concentrated to 5 µL in a speedvac. iii.) The proteins were digested overnight at 38 ºC with a Trypsin/LysC and Protease Max mixture (1:20:2 ratio of Trypsin/LysC: sample: 1% Protease Max). After the overnight digestion, the sample was acidified with 1% formic acid to a pH of 2-3. The peptide sample was then concentrated in a speedvac to 5 µL of solution.

LC/MS/MS Analysis. Reversed-phase nano-LC/MS/MS was performed on an Eksigent Ekspert nanoLC 425 system (SciEx) coupled to a Bruker MAXIS ETD II QToF mass spectrometer equipped with a nanoelectrospray source. Concentrated samples were diluted with a 1:1 ratio of sample:0.1% formic acid in water. Five µL of the samples were injected onto the liquid chromatograph. A gradient of reversed-phase buffers (Buffer A: 0.2% formic acid in water; Buffer B: 0.2% formic acid in acetonitrile) at a flow rate of 150 µL/min at 60 °C was set-up. The LC run lasted for 83 minutes with a starting concentration of 5% buffer B increasing to 55% over the initial 53 minutes and a further increase in concentration to 95% over 63 minutes. A 15 cm long/100 µm inner diameter nanocolumn was employed for chromatographic separation. The column was packed, in-house, with reverse-phase BEH C18 3.5 µm resin (Xbridge). MS/MS data was acquired using an auto-MS/MS method selecting the most abundant precursor ions for fragmentation. The mass-to-charge range was set to 350-1800.

Analysis of MS/MS Data. Mascot generic format (MGF) files were generated from the raw MS/MS data. Mascot (version 2.6) uses the MGF file for database searching and protein identification. For these samples, several custom caddisfly sequence databases were searched. The parameters used for the Mascot searches were: trypsin digest, two missed cleavages, carbamidomethylation of cysteine set as fixed modification, oxidation of methionine set as variable modifications, phosphorylation of S, T, and Y as variable modifications, methylation of R as variable modification, and the maximum allowed mass deviation was set at 11 ppm.

1. R. Core Team. 2019 *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org.

2. Ashton NN, Stewart RJ. 2018 Aquatic caddisworm silk is solidified by environmental metal ions during the natural fiber-spinning process. *FASEB J.* **33**, 572–583. (doi:10.1096/fj.201801029R)

3. Luo S, Tang M, Frandsen PB, Stewart RJ, Zhou X. 2018 The genome of an underwater architect, the caddisfly Stenopsyche tienmushanensis Hwang (Insecta: Trichoptera). *GigaScience* **7**. (doi:10.1093/gigascience/giy143)

4. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017 Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736. (doi:10.1101/gr.215087.116)

5. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990 Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. (doi:10.1016/S0022-2836(05)80360-2)

6. Li H. 2018 Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100. (doi:10.1093/bioinformatics/bty191)

7. Loman NJ, Quick J, Simpson JT. 2015 A complete bacterial genome assembled *de novo* using only nanopore sequencing data. *Nat. Methods* **12**, 733–735. (doi:10.1038/nmeth.3444)