

RSC Advances



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This *Accepted Manuscript* will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Structural Characterization of Nanofiber Silk Produced by Embiopterans (Webspinners)[†]

J. Bennett Addison,^a Tom Osborne Popp,^a Warner S. Weber,^a Janice S. Edgerly,^b Gregory P. Holland,^a and Jeffery L. Yarger^{*a}

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXX 20XX

First published on the web Xth XXXXXXXXXX 200X

DOI: 10.1039/b000000x

Embiopterans produce silken galleries and sheets using exceptionally fine silk fibers in which they live and breed. In this study, we use electron microscopy (EM), Fourier-transform infrared (FT-IR) spectroscopy, wide angle x-ray diffraction (WAXD) and solid-state nuclear magnetic resonance (ssNMR) techniques to elucidate the molecular level protein structure of webspinner (embiid) silks. Silks from two species *Antipaluria urichi* and *Aposthonia ceylonica* are studied in this work. Electron microscopy images show that the fibers are about 90-100 nm in diameter, making webspinner silks among the finest of all known animal silks. Structural studies reveal that the silk protein core is dominated by β -sheet structures, and that the protein core is coated with a hydrophobic alkane-rich surface coating. FTIR spectra of native embiid silk shows characteristic alkane CH₂ stretchings near 2800-2900 cm⁻¹, which decrease approximately 50% after washing the silk with 2:1 CHCl₃:MeOH. Furthermore, ¹³C ssNMR data shows a significant CH₂ resonance that is strongly affected by the presence of water, supporting the idea that the silk fibers are coated with a hydrocarbon-rich layer. Such a layer is likely used to protect the colonies from rain. FTIR data also suggests that embiid silks are dominated by β -sheet secondary structures similar to spider and silkworm silk fibers. NMR data confirms the presence of β -sheet nanostructures dominated by serine-rich repetitive regions. A deconvolution of the serine C β NMR resonance reveals that approximately 70% of all seryl residues exist in a β -sheet structure. This is consistent with WAXD results that suggest webspinner silks are 70% crystalline, which is the highest crystalline fraction reported for any animal silks. The work presented here provides a molecular level structural picture of silk fibers produced by webspinners.

1 Introduction

Embioptera (often called webspinners or embiids) produce silken galleries and sheets using exceptionally fine silk fibers in which they live and breed.^{1,2} Embiopterans are unusual insects because they are soft-bodied and flexible even as adults, and to a great extent, rely on silk instead of a tough exoskeleton for protection. Juvenile-form adult females, maternal care, and a colonial life that is defined by shared silk characterize this taxonomic order of insects. They spin silk produced in glands in the basal segment of their front feet (Figure 1D).¹ They step around the head and body as silk issues forth from multiple silk ejectors (for examples of spin step kinematics see Edgerly et al., 2012).³ Immatures, as well as adults, spin silk into domiciles that serve as retreats and egg chambers and into covered route-ways that lead to food. The two focal species in this study, *Antipaluria urichi* (Family Cloth-

didae) and *Aposthonia ceylonica* (Family Oligotomidae), are arboreal and tropical, living on tree bark where they graze on epiphytic algae and lichens within the relative safety of the silk (Figure 1A). Arboreal species share silk, affording them protection from predators, especially ants, which typically walk over silk apparently without recognizing that valuable prey lay beneath.

Webspinner silks are among the thinnest of all known silk-based biopolymers, but previous reports of fiber diameters are conflicting. Okada et al.⁴ found from SEM images that silks from the Australian webspinner *Aposthonia gurneyi* have fibers with a mean diameter of only 65 nm, while polarized optical microscopy data by Collin et al.^{5,6} found fiber diameters in the range of 500 nm for multiple embiopteran species. This discrepancy is curious, and is addressed in this work.

While the embiopteran taxonomy and behavior have been well studied,^{1-3,7-9} very little is known about the molecular-level structure of their silks. Significant insight was gained through cDNA sequencing of the primary silk protein sequence from multiple embiopteran species including *An. urichi* and *Ap. ceylonica*, which are the subjects of this study.^{6,10} Representations of the full (approximately 70 kDa) primary protein sequence for both species are seen in Figure

[†] Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

^a Department of Chemistry and Biochemistry, Arizona State University, Tempe, Arizona 85287-1604, USA. E-mail: jyarger@gmail.com

^b Department of Biology, Santa Clara University, Santa Clara, CA, USA.

1E. Like other silk-based biopolymers,^{11–13} webspinner silks are composed of highly repetitive protein sequences in which glycine, serine, and alanine are heavily represented.^{4,6,10} Although there exists very little experimental data characterizing webspinner silks, the obvious similarities between the embiopteran protein sequence and other well-characterized fibroin biopolymers allows us to make predictions. Silkworm and spider silks are also composed of highly repetitive proteins that contain runs of poly(GA) and poly(A).^{14,15} It is well known that these repeat motifs form water-inaccessible antiparallel β -sheet structures that are aligned along the fiber axis, and these rigid, nanocrystalline β -sheets are thought to be responsible for the impressive mechanical strengths of both spider and silkworm silks.^{16–18} When one looks at the cDNA sequences for many embiid silks,¹⁰ similarities are observed but alanine appears to have been mostly replaced by serine. Instead of runs of poly(GA), embiopteran primary protein sequences are dominated by runs of poly(GS) or GAGSGS repeats. Instead of runs of poly(A), some webspinner silks contain runs of poly(S). Serine has a higher propensity to adopt a β -sheet environment than alanine because of its ability to form side chain intra-sheet hydrogen bonds,^{19–21} therefore similar nanocrystalline β -sheet structures are fully expected from poly(GS), GAGSGS and poly(S) repeats found in embiopteran silks. Indeed, previous Fourier transform infrared spectroscopy studies on embiopteran silks have suggested that the fibers are dominated by β -sheet secondary structures within the protein core, presumably arising from said repetitive protein motifs.^{4–6,10}

These β -sheet structures within webspinner silks are potentially nanocrystalline and well aligned with respect to the fiber axis, similarly to other well-characterized silks. The percentage of the total fiber content that is nanocrystalline varies significantly among animal silks. Studies on spider dragline silk fibers show that the crystallinity content ranges from about 30% (*Nephila clavipes*) to 40% (*Latrodectus hesperus*),^{23,24} while fibers from the domesticated silkworm are typically higher, in the range of 40–60% crystallinity.^{25,26} The crystalline fractions arise mainly from repeated sequences, thus the primary protein sequence for webspinner silks should yield predictive power; protein secondary structure within spider dragline silk fibers has been shown to correlate quantitatively with silk primary protein sequences.^{24,27} With the exception of short C-terminal domains, the protein sequences of many embiopteran silks are virtually entirely composed of repetitive motifs, which likely adopt a β -sheet structure. For example, if all of the repeat domains within embiopteran silk from the species *An. urichi* adopt a β -sheet nanocrystalline structure, then the percent crystallinity may be upwards of 90%. If true, embiopteran silks are substantially higher in nanocrystalline content than other well characterized silk biopolymers. To the best of our knowledge, no attempt has been made to measure

the nanocrystallite size, orientation or crystalline fraction of any webspinner silks, which we address in this work through Wide Angle X-ray Diffraction (WAXD) techniques.

We are interested in better understanding the hierarchical structure of webspinner silks. Considering the many similarities in primary protein composition, it would not be surprising if the hierarchical structure of webspinner silk is similar to that of other well-characterized silks. The current skin-core model for spider dragline silks, for example, is a fibrous protein core that is encapsulated by a thin protective lipid-like coating.²⁸ As discussed, the protein core of spider dragline silk is composed of nanocrystalline β -sheet structures with dimensions of a few nanometers aligned along the fiber axis, separated by loosely-organized sheet, helical or randomly-oriented domains.^{22,29–31} This protein core is surrounded by a thin protective coating that is rich in long chain lipids and alkanes.^{32–34} Very little work has been performed characterizing this outer protective coating on spider silks, although Schulz et al.³⁴ conducted a thorough GC-MS analysis of the lipid and hydrocarbon content in silk from the orb weaver spider *Nephila clavipes*, revealing the presence of mostly long chain alkanes and methoxyalkanes of chain length between C₂₈ and C₃₄. The role of the lipid or alkane-rich surface coating is unclear, but it likely serves to both waterproof the silk fibers and maintain moisture within the silk core.³² Based on observations in the field and on laboratory colonies, embiopteran silken galleries are remarkably water-repellant, thus it is feasible that along with a β -sheet dominated protein core webspinner silk fibers possess a similar lipid or alkane-rich surface coating that serves a waterproofing purpose. Water-repellency would appear an especially adaptive trait for *An. urichi*, a tropical rainforest species from Trinidad and Tobago. Their colonies live on tree bark where they feed on epiphytic algae and lichens from within the protective covering of their silk. Rainfall is heavy, often torrential, and yet water appears to flow over the surface of the silk, leaving the insects dry beneath. *Ap. ceylonica* species are subtropical and tropical as well and are likely exposed to similar environmental conditions as the Trinidadian species. The availability of these two species in laboratory cultures and of published work on their silk provided us with an opportunity to further investigate the functional and structural aspects of embiopteran silk.

Interestingly, a few other silk-producing insects are known to contain a significant lipid content.^{35,36} It is often challenging to characterize a surface coating on spider silks with NMR techniques because this layer is thin relative to the overall fiber size and therefore hard to detect. The exceptionally fine fiber diameters of embiid silks provides us with a unique opportunity to observe and characterize any surface coating in addition to the protein core because of the higher surface area to volume ratio. In this present study, we use SEM, TEM, FT-IR, WAXD and solid-state NMR techniques to help elucidate both

molecular-level protein structure within the silk fiber core, and to interrogate the alkane-rich surface coating surrounding the silk fibers.

2 Materials and Methods

2.1 Silk collection

Cultures of approximately 50 *An. urichi* and *Ap. ceylonica* were established in separate plastic terrariums with dry oak leaves used as substrate. The insects were fed fresh romaine lettuce leaves every few days. The colony was misted with tap water every few days to maintain moisture and humidity within the domicile. The insects would spin silken sheets and tunnels, and the clean areas of these freshly spun networks were removed from the terrarium. Any visible debris (fragmented oak leaves, pebbles, dirt and waste) were removed from the silk using a pair of fine forceps under a dissecting microscope.

2.2 Fourier Transformed Infrared Spectroscopy

FT-IR analyses were performed using a Thermo Nicolet 6700 spectrometer equipped with an attenuated total reflectance (Smart Orbit ATR) accessory. For each sample, silk bundles were pressed onto a diamond window of the ATR attachment, and 64 scan averages were collected after 128 background scans.

2.3 Scanning Electron Microscopy

Cleaned webspinner silks were secured to conductive carbon tape, then were gold coated using a Denton vacuum sputter coater for 3 minutes at a deposition rate of 5 nm/min. The SEM was performed using a XL30 Environmental SEM-FEG built by FEI. The secondary electron (SE) detector was used for imaging. Images were collected under a vacuum pressure of less than 9×10^{-5} mbar and with a beam current of 5 kV. SEM images were processed using Gwyddion version 2.31 to measure fiber diameters from multiple images.³⁷

2.4 Transmission Electron Microscopy

Cleaned webspinner silk samples were immersed in 0.5% aqueous uranyl acetate and incubated 30 min at room temperature. An untreated control group was run in parallel and incubated in deionized water only. Both groups were washed with 3 consecutive changes of deionized water and allowed to air-dry overnight. Samples were then incubated in 3 consecutive changes of Spurr's epoxy resin over several hours, flat-embedded on Teflon-coated glass slides, and polymerized at 60°C for 24 hrs. Segments containing multiple fibers were cut from the thin resin layer and re-embedded in flat polymer

molds to obtain the desired cross-sectional orientation for microtomy. Sections were cut at 70 nm thickness with a diamond knife on a Leica Ultracut-R microtome. Images were generated on a Philips CM12 TEM operated at 80kV and captured with a Gatan model 791 CCD camera. TEM images were processed using Gwyddion version 2.31 to measure fiber diameters.³⁷

2.5 Wide Angle X-ray Diffraction

Embiid silk fibers were aligned parallel across a cardboard washer using a small amount of super glue at either end. Wide-angle x-ray diffraction measurements were collected at the Advanced Photon Source located at Argonne National Laboratory, Argonne IL, USA on the BioCars 14BM-C beamline which has x-ray wavelength of 0.9787 Å (12.668 keV) and a beam size of 130 x 340 μm (FWHM). The samples were mounted on a goniometer 300 mm from the ADSC Quantum-315 9-panel CCD array detector with the fibers offset a small angle from the beamstop to better view the meridian reflections. The exposure time was 60 seconds for each of ten images averaged and 5 background images were taken with the same parameters to remove air scattering. Images were then processed using Fit2D x-ray processing software and calibrated with cerium dioxide (CeO₂).

2.6 Peptide Preparation

A peptide model, GSGHGSSSSSSGDSGSGSGSGSGS-GSGA, was synthesized using Fmoc-chemistry to mimic web-spinner silk of the species *Ap. ceylonica*. A fully automated microwave-assisted peptide synthesizer (Liberty 1 by CEM) was used to prepare the peptide on the 0.1 mmol scale. The crude peptide was cleaved from the resin (preloaded Alanine-Wang resin from AAPPTeC) for 4 hours using a standard cleavage cocktail (95:2.5:2.5 TFA:TIS:H₂O). The crude peptide was precipitated out of TFA using cold diethyl ether, and after several ether washes, the crude peptide was dried, dissolved in 6M LiBr, and dialyzed against water at 4°C for 4 days. Any water-insoluble peptide was centrifuged out, and solubilized peptide was recovered by lyophilization. This lyophilized peptide represents *Ap. ceylonica* silks in a random-coil structure. Some of this peptide was dissolved in Formic acid and carefully pipetted onto a clean glass surface. The peptide crystallized into a β-sheet structure upon slow evaporation of Formic Acid. Peptides were then characterized using NMR spectroscopy.

2.7 Solid-State NMR.

¹³C solid-state NMR experiments were conducted on a 400 MHz Varian wide-bore instrument equipped with a 1.6 mm triple resonance MAS probe. Samples were spun at the magic

angle at 30 kHz. ^{13}C chemical shifts were referenced externally to TMS by setting the downfield adamantane resonance to 38.56 ppm. $^1\text{H} \rightarrow ^{13}\text{C}$ Cross Polarization under Magic Angle Spinning (CP-MAS) conditions were optimized using U- $[^{13}\text{C}, ^{15}\text{N}]$ glycine. For CP experiments, typical experimental conditions used were an initial 2 μs proton $\pi/2$ pulse, a 1 ms ramped ($\sim 15\%$) spin-lock pulse at a maximum of 150 kHz on the ^1H channel, and a square contact pulse set to the -1 spinning side band of the Hartmann-Hahn profile on the ^{13}C channel. A 100 kHz spectral width was used with 20 ms acquisition time, a 5 second relaxation delay, and 10240 scan averages. Direct Detection under Magic Angle Spinning (DD-MAS) experiments on hydrated embiid silk samples used a 2.5 μs $\pi/2$ pulse, 20 ms acquisition time, 100 kHz spectral width, a 1 second relaxation delay, and 40960 scan averages. All CP and DD-MAS experiments were collected using 150 kHz TPPM 38 proton decoupling during acquisition. To collect data on hydrated samples, silks were first soaked in DI water for at least 30 minutes then blotted with a Kim-Wipe to remove excess water.

3 Results and discussion

3.1 Fiber Diameters

Silk from the embiid species *Antipaluria urichi* and *Aposthonia ceylonica* were studied using SEM, TEM, FT-IR, WAXD and NMR spectroscopy to characterize the molecular-level protein structure as well as a hydrophobic surface coating rich in long-chain lipids and alkanes. Figure 1 shows both optical and SEM images of insects and silk produced from *An. urichi*. The insects produce silk out of their tarsal organs, or forelimbs, creating very thin sheets of silk protecting the colonies. An example of a silk in a natural, arboreal setting can be seen in Figure 1A. Fiber diameters from *An. urichi* were determined using SEM and TEM microscopy. Previous work by Collin et al. 5,6 reported fiber diameters in the range of 500 - 800 nm. However, the authors used polarized light microscopy techniques and thus could not resolve fibers below the optical resolution limit. Figure 1C shows how one could easily be fooled; it is likely that the authors were observing bundles of webspinner silks and were unable to resolve fine detail. Figure 2 shows a histogram of *An. urichi* fiber diameter measurements made from 68 isolated fibers over multiple SEM images and 82 fibers from TEM images. The fibers for SEM imaging had been coated with a layer of gold approximately 15 nm thick, thus 30 nm was subtracted from each edge to edge measurement. The average size was 93 ± 15 nm (one standard deviation), which is more consistent with the 65 nm fibers reported for a different webspinner species. 4 For TEM imaging, fiber bundles were stained by submerging in an aqueous solution containing 0.5% uranyl acetate for 30

minutes prior to resin embedment. The average fiber diameter over 82 measurements from 3 separate TEM images was 100 ± 15 nm. This is slightly larger than the 93 nm average result from SEM images, but we note that these fibers were soaked in an aqueous-based uranyl acetate stain prior to resin embedment and thus we are potentially observing a slight swelling of the fibers due to water absorption. This observation brings doubt into the validity of previous mechanical testing results on embiid silks. Webspinner silks show elasticity similar to spider silks (15-40% extensibility), but silk strengths were reported at only about 150 MPa. 5,6 If correct, this is many times weaker than spider dragline fibers and silkworm silk. As a silk used primarily for structural and protective purposes and not for absorbing impact, it would be surprising if webspinner fibers possess similar gigapascal-level strengths as spider dragline fibers. Nevertheless, embiid silks galleries must be strong enough to both deter predators, which often walk on top of the silk.

To the best of our knowledge, the only available mechanical data obtained on silk produced by webspinner insects is unreliable due to improper fiber diameter measurements. Therefore in an attempt to better estimate embiid fiber tensile properties, we collected tensile stress-strain curves on silk bundles prepared for *An. urichi*. Samples were prepared by carefully brushing an E-shaped cardboard card across the tarsus of adult female insects of *An. urichi*. Fibers were superglued to each of the three anchor points on the E. Stress-strain curves were obtained by stretching one side of the E-shaped card at a rate of 1% per second, while the other unstretched side was analyzed using SEM to approximately obtain the number of fibers present. Additional experimental details and results are included as supplemental material. Results suggest that webspinner silks are significantly stronger than previously thought; we observed an average of 500 MPa mean ultimate stress and about 30% extensibility over 14 measurements. Due to the small fiber diameters and extreme difficulty in obtaining consistent samples, this result should only be interpreted as a rough estimate.

3.2 Infrared Spectroscopy: Surface coating and protein core

An additional key feature of embiid silken galleries is the ability to prevent water from penetrating into the silk networks. While composed primarily of protein, their silk is exceptionally hydrophobic; silken sheets protecting colonies do not appear to allow water to penetrate into the domiciles. If the surface of the silks were composed of glycine, alanine and serine-rich protein, one might expect water to easily penetrate a thin layer of silk. The hydrophobic nature of the silk suggests some form of lipid or alkane-rich surface coating similar to the water repelling layer found on spider silks. 28,32,33 Our

results provide evidence for such a surface coating. FT-IR data and NMR data both show the presence of alkanes even after washing the silk with the detergent sodium dodecyl sulfate (SDS), indicating that the surface layer must be significantly longer than C_{12} as to not be removed by SDS. Indeed, initial GC-MS analysis of a $CHCl_3$:MeOH extraction from *An. urichi* silk shows the presence of straight and branched-chain alkanes of at least length C_{34} , and likely even longer (supporting information). FT-IR spectra seen in Figure 3 on native web-spinner silks show characteristic alkane CH_2 symmetric and asymmetric stretch absorbances at 2850 and 2920 cm^{-1} , respectively.³⁹ We attribute this to alkane CH_2 absorbance from the hydrocarbon-rich surface coating on the silk. Previous FT-IR studies on embiopteran silks also show a large CH_2 absorbance, but it was not assigned.⁴ The native silk samples were then treated with 2:1 mixture of $CHCl_3$:MeOH overnight in attempt to remove any surface lipid or alkane-rich layer, and FT-IR analysis was then repeated on the treated silk samples. The amide I and amide III bands did not change suggesting that the protein structure remained unaffected (supporting information), but the alkane CH_2 stretchings dropped approximately 50% in intensity after extraction (Figure 3B). This result clearly shows the presence of a lipid or alkane-rich surface coating. It also becomes clear that the surface coating is not covalently attached to the protein core, so perhaps this surface layer is co-secreted along with the protein fiber.

The protein core structure was also analyzed using infrared spectroscopy. FT-IR is an extremely common technique used to study protein secondary structure of silk-based biopolymers; some examples include silks produced from silkworm,^{40–43} spiders,^{42,44} caddisfly larvae,⁴⁵ silverfish,⁴⁶ lacewing,³⁶ and web-spinners.^{4,5} The amide I, II and amide III absorbances are commonly used to infer protein secondary structure. There seems to be a general consensus in the literature regarding the amide I band, but significant debate in assigning secondary structures to components of the amide III band. For example, the amide III β -sheet and random coil absorbances were assigned at 1263 and 1230 cm^{-1} respectively for silkworm silk from *Bombyx mori*,⁴⁰ but are assigned to 1222 and 1242 cm^{-1} for silkworm silk from *An. pernyi*. It is possible that the inversion of peak location is a result of different repetitive motifs; *Bombyx mori* silk is dominated by poly(GA) and GAGAGS units while *An. pernyi* contains runs of poly(A) similar to spider silks. In support, careful secondary structure assignments to the amide III band for poly(A)-rich spider silks agree well with *An. pernyi* but not with *Bombyx mori*.⁴² Thus the primary protein sequence has a major impact on FT-IR absorption peak positions, as does fiber strain and overall secondary structure.⁴⁷ Considering significant variation in peak position, width, and shape, we think that a quantitative deconvolution of FT-IR absorbance bands will not give convincing results. Qualitative and especially

comparative analysis of IR absorbance profiles, however, can be extremely powerful. Figure 4 shows that the FT-IR absorbance profile of *An. urichi* is virtually identical to that of *Ap. ceylonica*. The strong amide I absorbance maximum at 1627 wavenumbers is identical to reports from Okada et al.,⁴ on web-spinner silk from the congener *Ap. gurneyi*, which is similarly dominated by poly(GS) and poly(S) motifs. For comparison, the amide I band for β -sheet-rich silkworm silk is found at 1615 cm^{-1} (supporting information). The shift towards higher frequency absorbance for web-spinner silk relative to silkworm silk could be attributed to a larger random coil component, as is often done in the literature through peak deconvolution. While this is a possibility, we think it is more likely that the shift in frequency is a result of differing primary protein sequences rather than a lower β -sheet fraction, especially when considering the high expected crystalline fraction for web-spinner silks. Amide I absorbances from random-coil structures are typically found near 1650 cm^{-1} while β -sheet structures show absorbances at lower frequencies (1620 cm^{-1}).⁴⁸ Additionally, the shoulder absorbance at 1695 cm^{-1} is regularly assigned to β -sheet protein structures,^{40,48} thus it is quite clear that like silkworm silks, web-spinner silk fibers are dominated by β -sheet structures, and are remarkably similar across species.

3.3 Wide Angle X-ray Diffraction

Characterization of web-spinner silk from *An. urichi* by wide-angle x-ray diffraction (WAXD) indicates an amorphous fraction with diffuse scattering and a crystalline fraction comprised of nanocrystalline β -sheet structures. The diffraction pattern observed for axially-aligned web-spinner silk from *An. urichi* shown in Figure 5A closely resembles our previously published spider silk patterns, confirming the presence of regular nanocrystalline β -sheet structures within the silk protein core.^{23,49} Careful analysis of the WAXD profile gives crystallite size, orientation of the nanocrystallite with respect to the fiber axis, and the overall percent crystallinity of the fiber. The 12 highest intensity unique reflections were identified and assigned to an orthorhombic unit cell of dimensions $a = 6.6\text{ \AA}$, $b = 9.6\text{ \AA}$, $c = 19.2\text{ \AA}$. Each reflection was fit as an individual component in d spacing using an iterative Gaussian fitting protocol until it converged. The radial broadening of the crystalline reflections, as demonstrated by the full width at half max (FWHM) in 2θ space (Figure 5C, 5D), was used to calculate the average crystallite size along the a , b , and c axes using Scherrers formula. This resulted in crystallite dimensions of 3, 4, and 5 nm in the a , b , and c axes, respectively.²⁵ Azimuthal broadening of the crystalline reflections is indicative of variance in alignment of the crystalline fraction with respect to the fiber axis. Hermans orientation factor of the crystalline fraction, f_c , was calculated to be 0.93, from the FWHM of ra-

dial integration at 4.3 \AA^{-1} where $f_c = (3(\cos^2\phi)-1)/2$ and ϕ is the angle between the c axis of the nanocrystallite and the fiber axis. Difficulty manually aligning the silk fibers during sample preparation leads to fiber misalignment, which means this is a low estimate of the actual nanocrystallite orientation within the fiber. The overall crystalline fraction can be estimated as a fraction of the area represented by radial integration 15° on either side of the equator for the equatorial (200) and (120) crystalline peaks relative to the area represented by the integrated intensity of the full diffraction pattern resulting in 69% crystallinity. This high crystallinity content is likely due to the high fraction of repetitive motifs and significant serine content, both of which correlate to increased β -sheet content in the silk protein core.

3.4 Solid-State NMR

To get a more complete understanding of both the silk protein core and the encapsulating surface coating, we used solid-state NMR techniques on both native and CHCl_3 :MeOH-washed silks in their dry and hydrated states. Solid-state NMR is a very powerful tool for characterizing biopolymers; it has been widely utilized to elucidate molecular-level environments of repeat motifs such as poly(GA) and poly(A) in silkworm and spider silks, respectively. For embiid silks from both species discussed in this work, glycine and serine are the most dominant amino acids represented at approximately 45% and 36%, respectively.¹⁰ The ^{13}C chemical shift of glycine $C\alpha$ is not very sensitive to secondary structure, however both the alanine and serine $C\alpha$ and especially $C\beta$ chemical shifts are influenced dramatically by local conformation.^{50,51} ^{13}C chemical shifts have been used extensively to identify secondary structures within silk-based biopolymers, and can be applied here to elucidate molecular level environments within embiid silks. Cleaned, native silk from *Ap. ceylonica* and *An. urichi* were studied with ^{13}C solid-state NMR spectroscopy. ^{13}C resonances for the three most abundant amino acids are assigned, and secondary structures are indicated (Figures 6, 7). We are limited in the amount of information we can extract from ^{13}C data on naturally-abundant samples, but we can still draw many strong conclusions. For one, serine exists predominantly in β -sheet structures based on the ^{13}C chemical shifts observed for serine $C\beta$ (Figure 6). As indicated in Figure 6, the serine $C\beta$ resonances in the CP-MAS spectra lie at 64.5 ppm, a shift consistent with serine in a β -sheet environment. Serine from GAGSGS and poly(S) repeat motifs are most likely responsible for this observation. We also can see a minor shoulder at 62 ppm that is assigned to serine containing regions in a randomly-oriented environment.

Further confirmation of the secondary structural assignments were made by comparing $^1\text{H} \rightarrow ^{13}\text{C}$ CP-MAS NMR data of peptides mimicking webspinner repetitive motifs that

were either trapped in a random-coil confirmation or crystallized into a β -sheet structure from formic acid (Figure 8). To mimic webspinner silk in a random-coil structure, a 30 AA peptide representing the protein sequence of *Ap. ceylonica* silk (illustrated in Figure 1E) was synthesized using solid-phase peptide synthesis, dissolved in 6M LiBr, dialyzed against DI water for 4 days, and then lyophilized. The peptide was then dissolved in formic acid and slowly dried, inducing a β -sheet structure. This method is commonly used to capture silkworm model peptides in both random-coil and β -sheet conformations.^{52,53} $^1\text{H} \rightarrow ^{13}\text{C}$ CP-MAS data was collected on both peptides (8A). Using DMFIT software⁵⁴, the spectra for random-coil and β -sheet model peptides were independently fit to extract precise ^{13}C chemical shifts, peak widths, and % Gaussian vs. % Lorentzian line shapes. This information was then used to deconvolute NMR spectra for native silks to approximate β -sheet content; only peak amplitudes were allowed to vary when fitting native webspinner silk data. Deconvolution of NMR data obtained for *Ap. ceylonica* (8B) and *An. urichi* (data not shown) silks indicate that serine $C\beta$, the residue most isolated and most sensitive to secondary structure, is approximately 70% β -sheet and 30% random-coil for both species. This estimation is consistent with both our FT-IR and our XRD data in that FT-IR absorbance profiles for the two species suggest that the silk protein structures are virtually identical and are dominated by β -sheet structures, and that analysis of XRD results reveal that *An. urichi* silk is 70% crystalline. The alanine $C\beta$ resonance shows a similar story in both silks, although it is more clear for *An. urichi* silk fibers where alanine is better represented (6C, 6D). Alanine, found almost exclusively in GAGSGS repeat motifs from *An. urichi* silk, clearly exists predominantly in β -sheet structures.

The contrast between CP and DD-MAS spectra in Figure 7 is very useful to highlight rigid versus mobile regions of hydrated silks.⁵⁵⁻⁵⁸ Domains of the silk that are affected by water will typically have shortened ^{13}C T_1 and longer T_2 relaxation times due to increased molecular motion, thus are easily seen as sharp peaks in the direct spectra. The DD-MAS experiment uses a fast (1 second) recycle delay so that ^{13}C resonances from rigid regions of the silk with long T_1 relaxation properties become saturated and therefore are not observed. While the majority of serine adopts a β -sheet structure within the silks, a new sharp resonance emerges in the direct spectra from both species at 62.2 ppm. This peak corresponds to serine $C\beta$ in a random coil or loose helical environment. Serine is found almost exclusively in poly(GS), poly(GAGSGS) or poly(S) repetitive motifs in both silks, therefore not all seryl residues from these repeats exist in a β -sheet structure. A native silk gallery will naturally repel water due to its hydrophobic coating, but in this case we encouraged hydration of the fibers by completely submerging the silk samples in water. Perhaps when the silk becomes forcibly hydrated,

poly(GS) and GAGSGS repeats from *Ap. ceylonica* and *An. urichi* silks are stable in both conformations, or can exchange between sheet-like and randomly-oriented structures on a slow timescale. Alanine resonances are weak from *Ap. ceylonica* silks, but we also observe a sharp peak emerge at 17.5 ppm for alanine C β from *An. urichi* silk in the direct spectra. Similar to the random coil serine C β peak evident by the sharp peak at 62.2 ppm, a portion of alanine also appears to adopt a random coil environment. Alanine from *An. urichi* silks comes from GAGSGS repeat domains, further supporting the idea that such repeats are found in both sheet-like and random structures when hydrated. For comparison, *Bombyx mori* silkworm silk is dominated by similar GAGAGS repeat domains, however a parallel experiment conducted on these fibers does not result in an emerging sharp helix or random-coil resonance upon fiber wetting. It appears then that when water penetrates into and solvates webspinner silk fibers, a higher fraction of silk protein becomes mobile as compared to similar regions of silkworm silk. Webspinner silk protein is much smaller than that of silkworm (70 kDa vs. 350 kDa). Perhaps for both silkworm and webspinner silks then, the dominant protein repeat units are embedded primarily in β -sheet nanostructures, but C- and N-terminal ends are less restricted. The shorter length of the webspinner protein and therefore higher percentage of protein near the C- and N-terminal domains may account for the apparent increase in random-coil protein content after fiber wetting.

Figure 7 also yields information on the surface coating surrounding the silk fibers. For both *Ap. ceylonica* and *An. urichi* silks, one notices that the alkane CH₂ resonance is significantly pronounced in the direct spectra when the silk is wet. The alkane CH₂ resonance from both silks is both sharp and strong in the direct spectra, presumably because of fast molecular motion induced by the presence of water on the silk surface. Additionally, both *An. urichi* and *Ap. ceylonica* silks show a new sharp resonance in the direct spectra at 14 ppm, which we assign to terminal or branched alkane CH₃ groups on the silk surface. Similar to FT-IR results in Figure 3, NMR data shown in Figure 7 on *Ap. ceylonica* silk fibers reveal the severe reduction of alkane CH₂ and CH₃ signals after washing the fibers with a 2:1 mixture of CHCl₃:MeOH (red). We conclude that the surface coating surrounding the silk protein core is composed of long-chain lipids or alkanes that are non-covalently attached to the protein.

4 Conclusions

An in depth analysis has been performed on webspinner silks from *An. urichi* and *Ap. ceylonica*, and our findings, summarized in Figure 9, allow us to improve the current understanding of the molecular-level and hierarchical structure of embiopteran silk fibers. Our best electron microscopy results

show that webspinner silks are 90-100 nm in diameter, making them among the thinnest known silk-based biopolymers. Both FT-IR and NMR results confirm what could be predicted by the repetitive primary protein sequences; the protein core is dominated by nanocrystalline β -sheet structures arising from glycine- and serine-rich repeat motifs. Wide angle x-ray diffraction results show that for *An. urichi* silk, and presumably other embiopteran silks, these nanocrystallites are highly ordered, are well aligned with respect to the fiber axis (low estimate: $f_c = 0.93$), and the crystallites are approximately 3 to 5 nm in dimension. Additionally, x-ray data show that webspinner silks possess an extremely high crystalline fraction at 69%. This is consistent with solid-state NMR results, which indicate that 70% of serine residues reside in nanocrystalline β -sheet structures. The silk is naturally water-repellent, but when the silk is water saturated, some regions of the protein, possibly the N- and C-termini of the relatively small (70 kDa) protein, become mobile and randomly-oriented. NMR and FT-IR data also provide evidence for a long-chain lipid or alkane-rich surface coating on webspinner silk fibers. Characteristic alkane infrared absorbances and ¹³C resonances decrease 50% in intensity after treating the silk with 2:1 CHCl₃:MeOH, suggesting that this surface coating is non-covalently adhered to the protein core. The hydrophobic nature of silken galleries appears to serve to protect the insects from being dislodged from their arboreal dwellings during heavy rain. It also protects the interior of their domiciles and their bodies from extreme wetting, which has multiple negative consequences including increased risk of suffocation and of fungal attack. During severe dry periods, the silk might also serve to limit desiccation of the insects sheltering inside. The basic hierarchical structure of webspinner silk fibers is illustrated in this study: the thin fiber core is composed of glycine and serine-rich repetitive protein motifs that primarily adopt a β -sheet nanostructure, and the protein core surrounded by a thin protective hydrophobic lipid or alkane-rich shell.

5 Acknowledgements

We gratefully acknowledge the School of Life Sciences EM Laboratory at Arizona State University for TEM imaging, and the LeRoy Eyring Center for Solid State Science at Arizona State University for SEM imaging. We also thank Dr. Brian Cherry for help with NMR and Robert Henning for his help during the x-ray experiments. Use of the Argonne National Laboratory (ANL) Advanced Photon Source (APS) was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract no. DE-AC02-06CH11357 and DOE-EPSCoR under contract number DE-SC0004791. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under grant number RR007707. The De-

partment of Defense (DOD) Air Force Office of Scientific Research (AFOSR) under award no. FA9550-14-1-0014 and the National Science Foundation (NSF), Division of Materials Research under award no. DMR-1264801 supported this work. NSF grant DEB-0515865 awarded to JSE at SCU supported collecting trips into the field and maintenance of cultures at SCU. We also thank Khaaliq Dejan for assistance in harvesting and cleaning silk. He was supported by the Faculty Student Research Assistant Program at SCU.

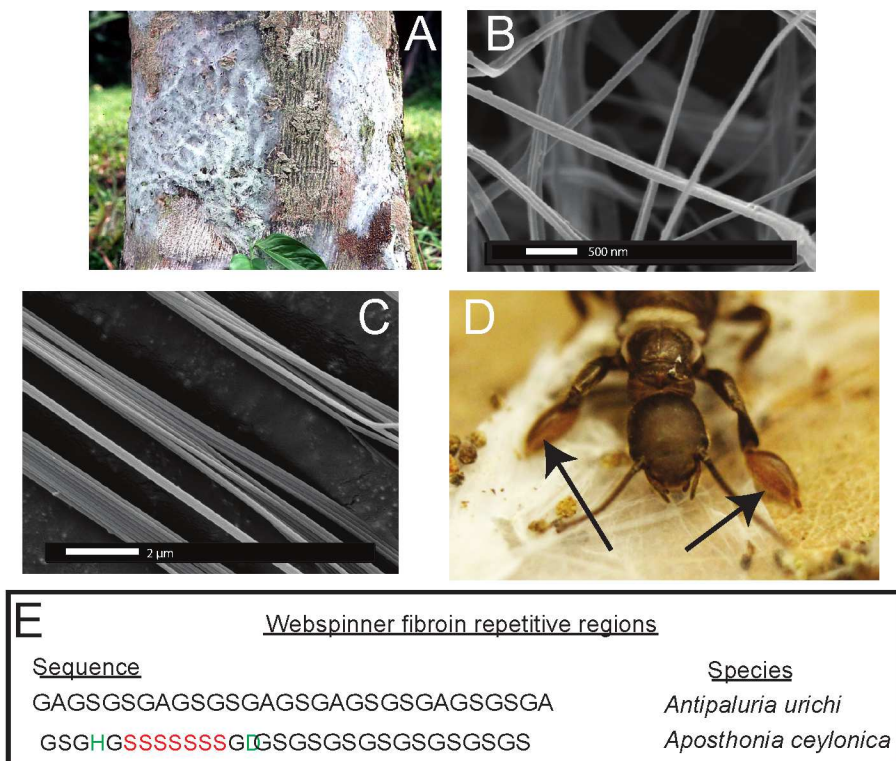


Fig. 1 (A - D) Optical and SEM images of arboreal embidoptera insect silk from the species *Antiplauria urichi*. The insects spin silken sheets and tunnels that protect and harbor the colonies (A) using exceptionally fine (~90 nm) fibers (B, C) from their tarsus (forearms), indicated with arrows in (D). Adult insects are approximately 1.6 cm in length. A representation of the repetitive protein regions from both species is shown in (E). Sequences from *Antiplauria urichi* and *Aposthonia ceylonica* silks were obtained from GenBank [FJ361212] and [EU170437], respectively

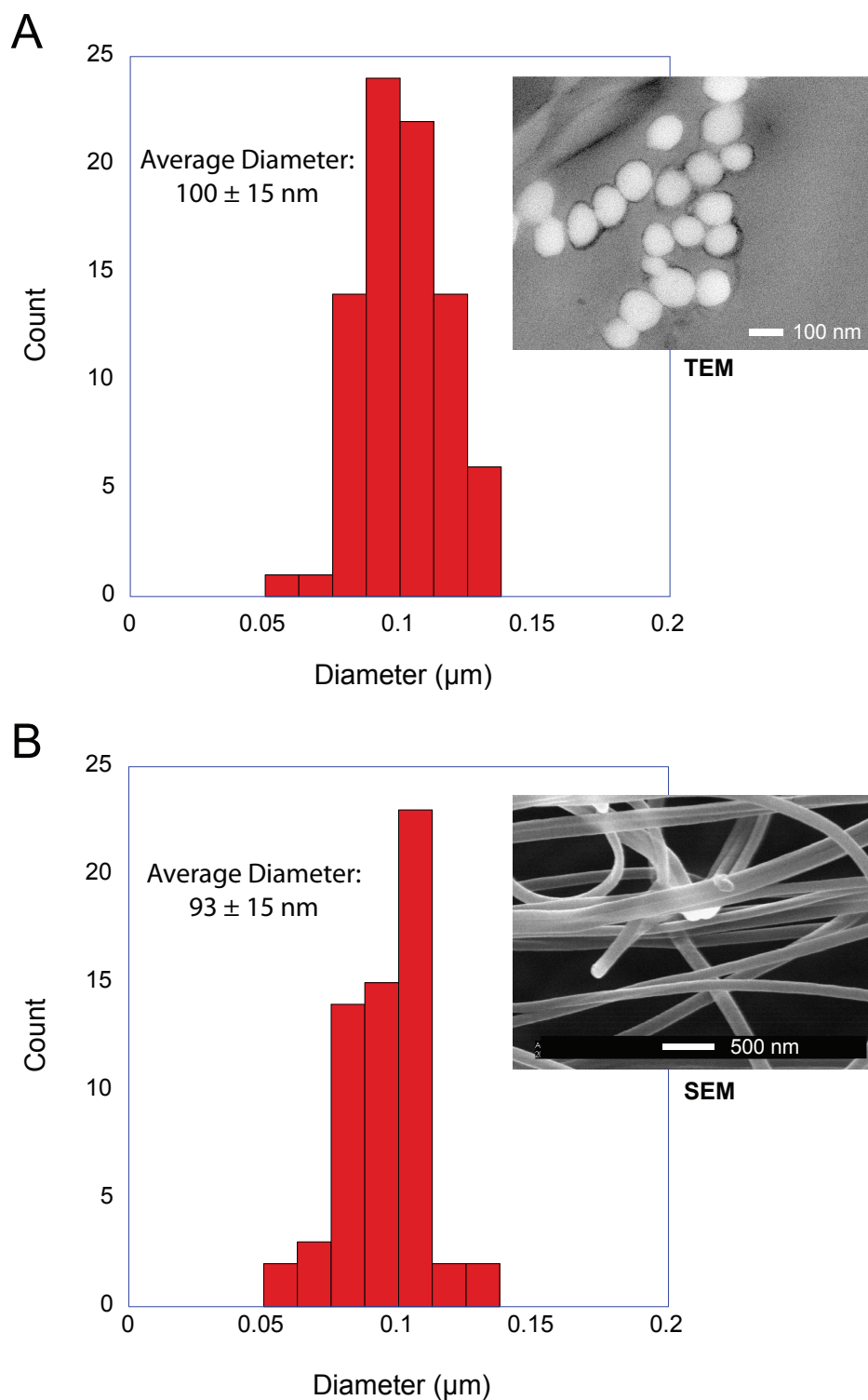


Fig. 2 Fiber diameter distribution for silk obtained from adult female *Antipaluria urichi* using electron microscopy. 82 and 68 separate measurements were combined from multiple TEM (A) and SEM (B) images, respectively. The results indicate that *An. urichi* silk fibers are approximately 90 - 100 nm in diameter.

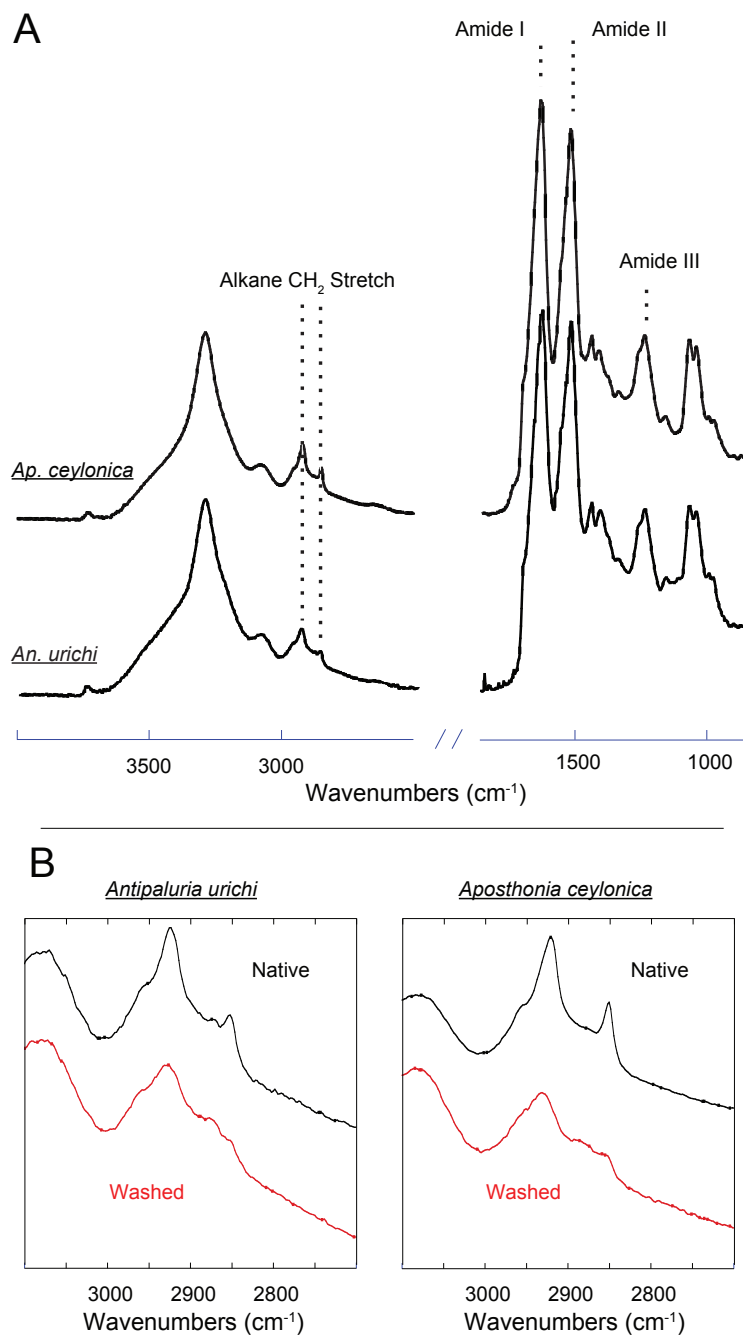


Fig. 3 A: FT-IR spectra of native *Antipaluria urichi* and *Aposthonia ceylonica* silk bundles. The strong amide I and II absorbances at 1625 and 1514 cm⁻¹, respectively, suggest that the protein core is predominantly β -sheet. **B:** Native webspinner silks (black) are compared to silks that were washed with 2:1 CHCl₃:MeOH for 24 hours (red). Amide I, II and III bands remain unchanged (data not shown), but the alkane CH₂ symmetric and asymmetric stretches at 2850 and 2920 wavenumbers, respectively, drop approximately 50% in intensity after treatment. This is associated with the removal of a lipid or alkane-rich surface coating on the silk fibers.

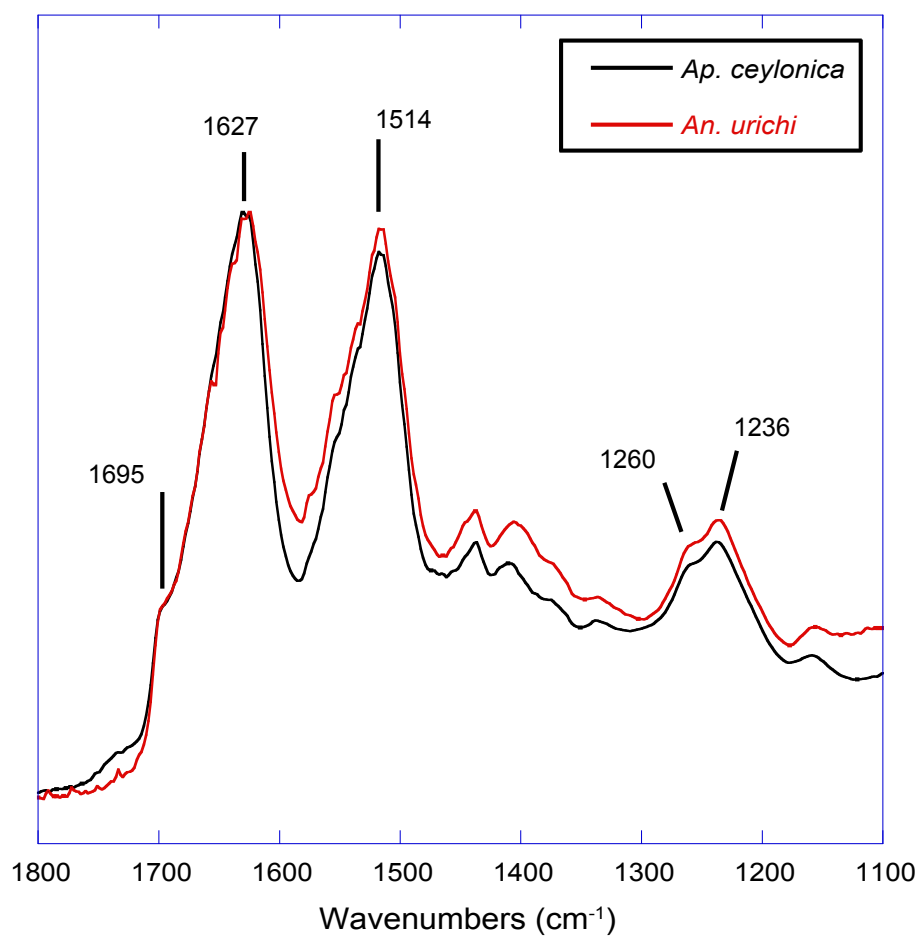


Fig. 4 FT-IR absorbance profiles comparing web spinner silks from the species *Antipaluria urichi* (red) and *Aposthonia ceylonica* (black). Spectra are virtually identical, emphasizing their remarkable similarity in structure.

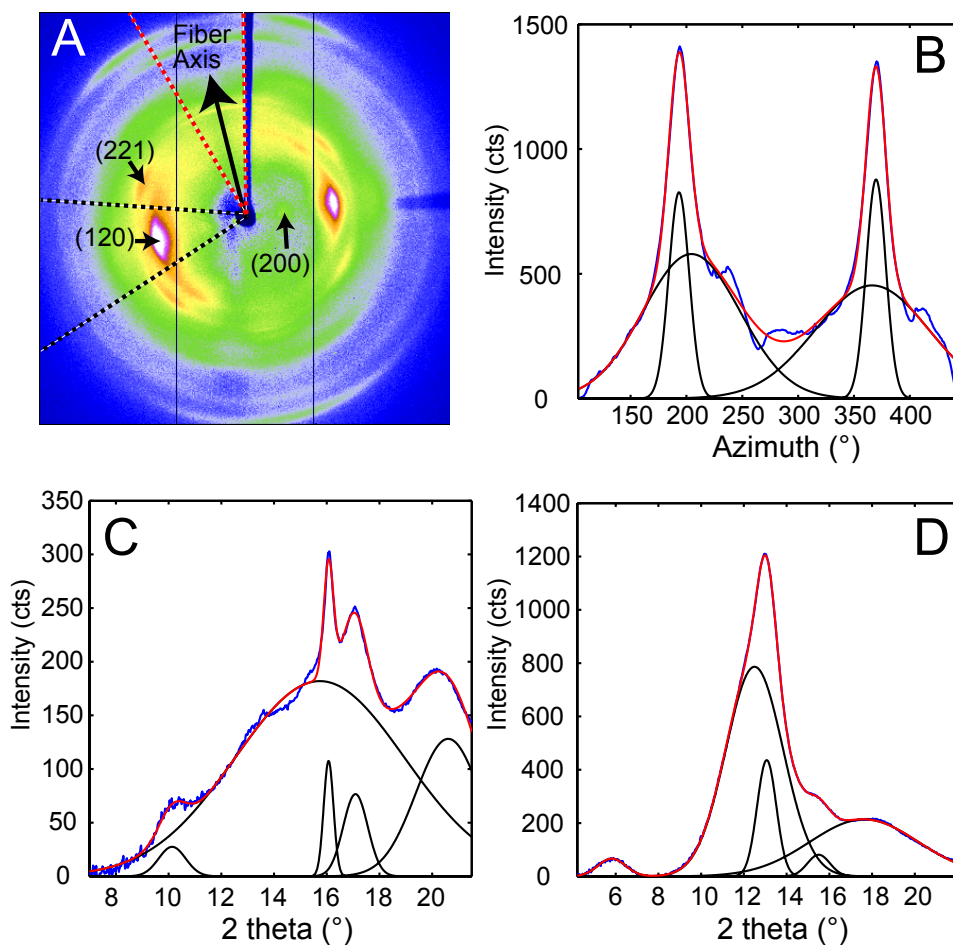


Fig. 5 Summary of Wide Angle X-ray Diffraction (WAXD) data. (A) 2D WAXD pattern of axially aligned webspinner silk from *Antipaluria urichi*. The fiber axis relative to the beamstop shadow (vertical) is indicated with the arrow. Equatorial reflections (120) and (200) are associated with nanocrystalline β -sheets, and the (221) reflection is associated with the amorphous component. (B) 1D azimuthal intensity profile of narrow ring around the (120) reflection (4.3 \AA^{-1}) starting at the beamstop and continuing counter-clockwise. (C) 1D radial intensity profile of the meridian wedge, indicated with red dashed lines in (A). (D) 1D radial intensity profile of the equatorial wedge, indicated with black dashed lines in (A).

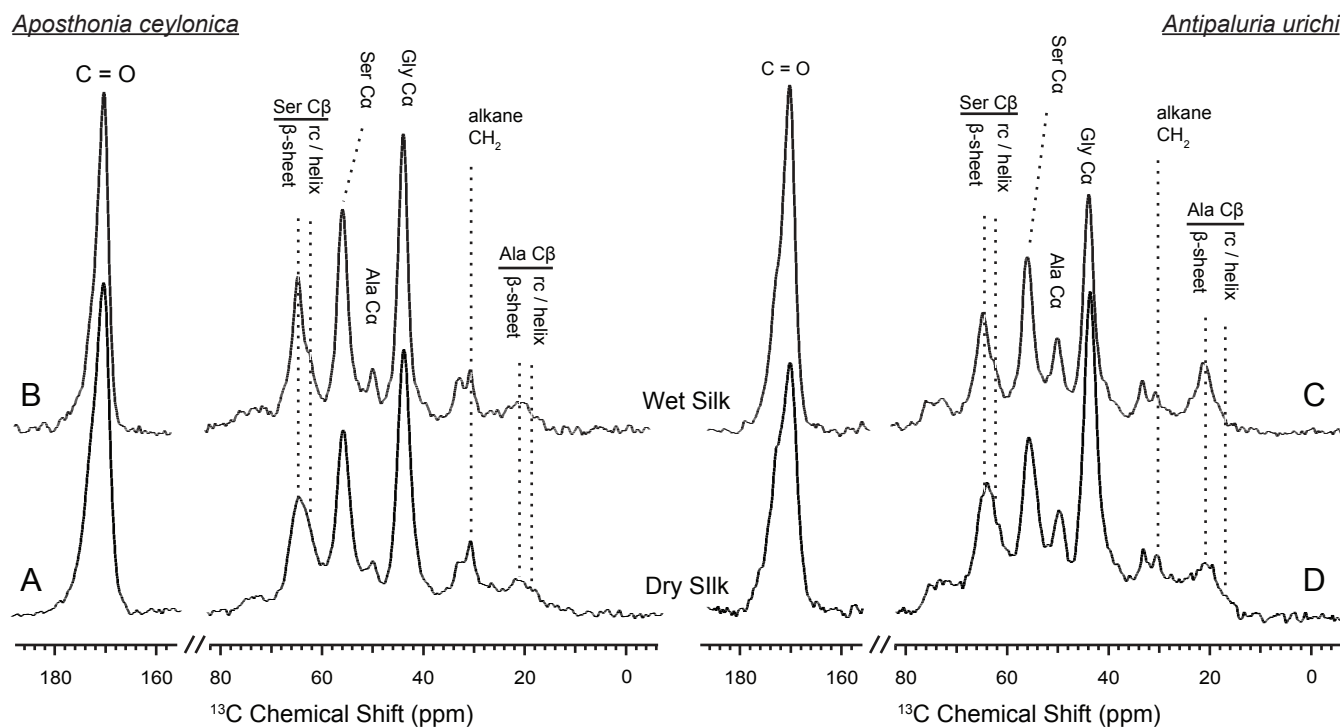


Fig. 6 $^1\text{H} - ^{13}\text{C}$ CP-MAS NMR spectra on wet (top) and dry (bottom) embiid silks from *Aposthonia ceylonica* (left) and *Antipaluria urichi* (right). The three most common amino acids have been assigned, and secondary structures are indicated for serine and alanine $\text{C}\beta$ resonances. An alkane CH_2 peak is clearly present in both silks. Silk from both species appears structurally very similar, although *An. urichi* contains significantly more alanine than *Ap. ceylonica*.

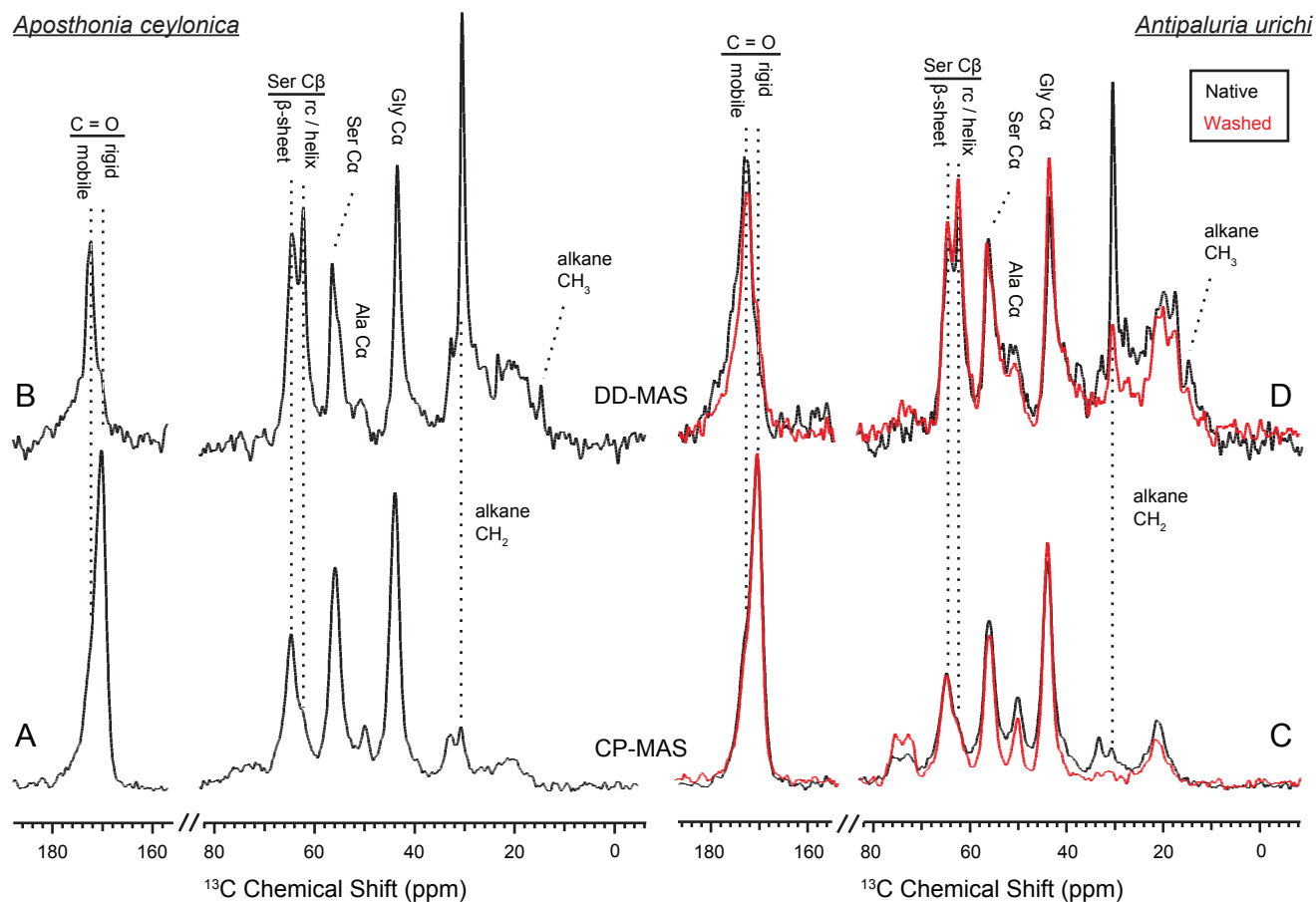


Fig. 7 ^1H - ^{13}C CP-MAS (bottom) and DD-MAS (top) NMR on wet embiid silks from *Aposthonia ceylonica* (left) and *Antipaluria urichi* (right). Data from native silks is seen in black, while experiments performed on *An. urichi* silk that was extracted with 2:1 CHCl_3 :MeOH is overlaid in red. Peaks observed in the direct spectra are significantly plasticized by water, while those observed in the CP spectra are from more rigid regions. The CH_2 alkane resonance is clearly affected by water, suggesting its presence on the fiber surface. A new CH_3 peak is seen in the direct spectra at 14 ppm, which we assign to terminal or branched alkane methyl groups from the surface coating.

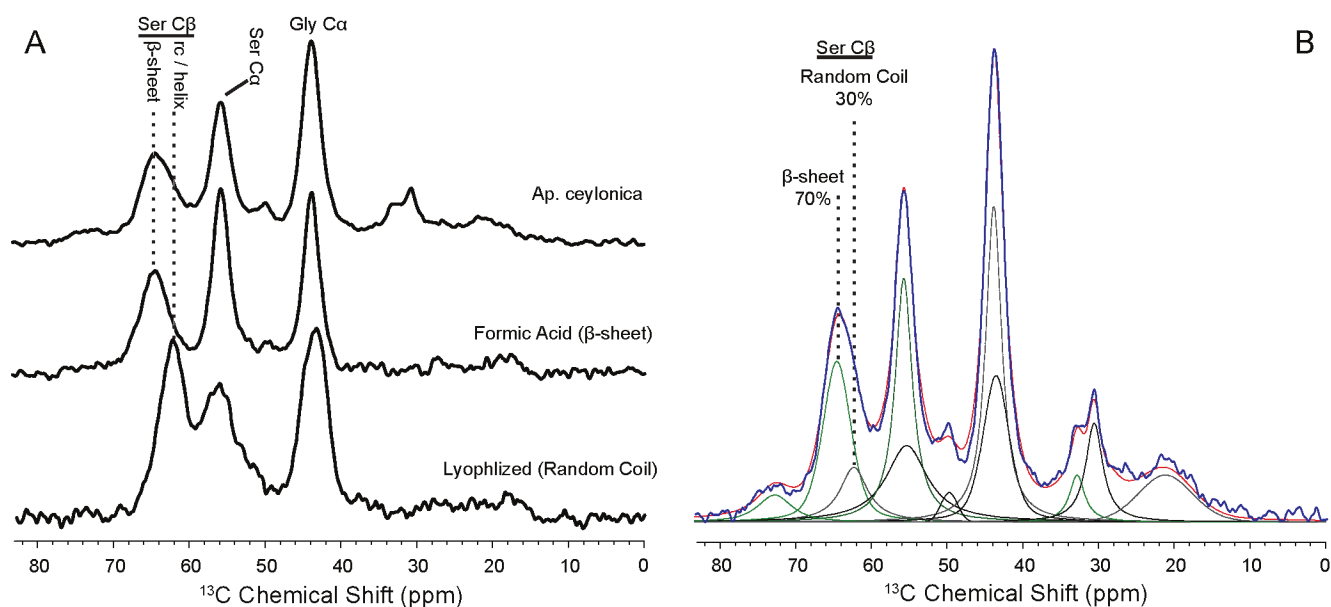


Fig. 8 (A) ^1H - ^{13}C CP-MAS NMR data comparing native webspinner silk (*Aposthonia ceylonica*, top) to representative peptide models mimicking β -sheet (middle) or random-coil (bottom) conformations. (B) Using peak locations, widths and lineshapes obtained from data on peptide mimics in (A), native webspinner silk data was fit such that only peak amplitude was allowed to vary. The data is shown in blue, and the resulting fit in red. Results suggest that approximately 70% of seryl residues adopt a β -sheet structure, which is in strong agreement with WAXD results.

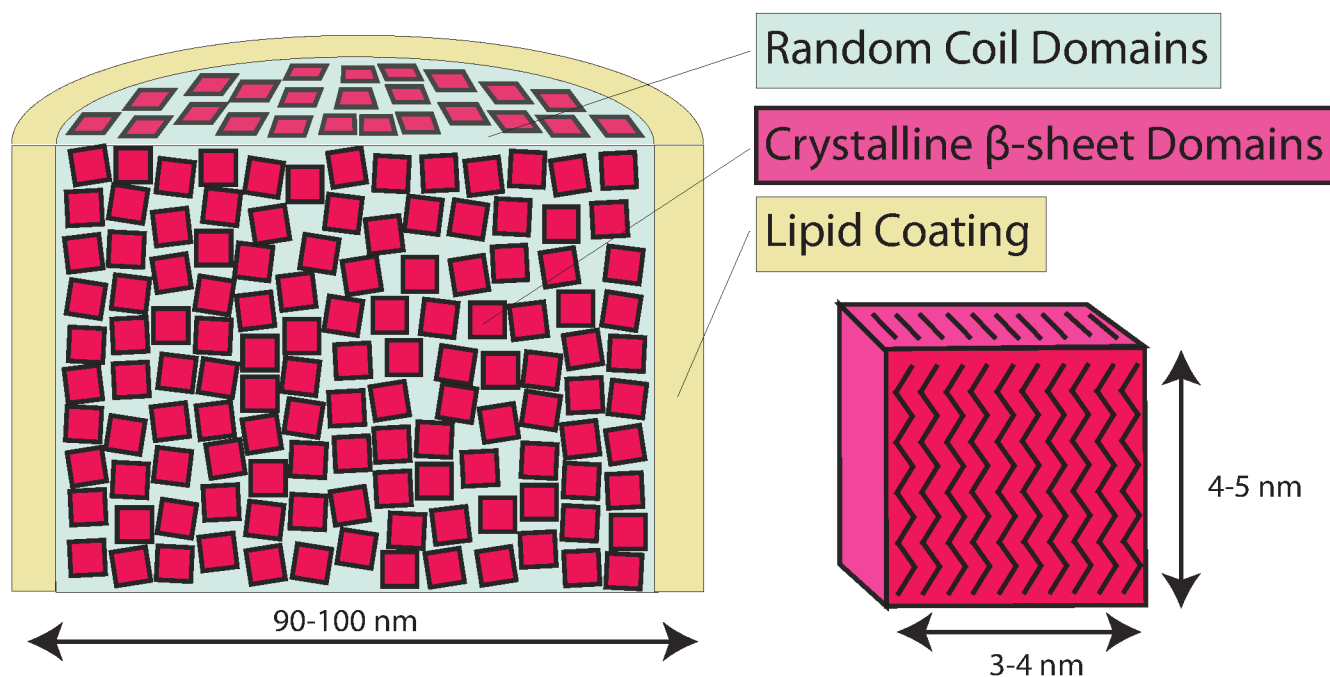


Fig. 9 Conceptual model of webspinner silk based on experimental results. β -sheet nanocrystallites (red) are a few nanometers in dimension, are well-aligned with respect to the fiber axis, and they make up about 70% of the silk protein core. Separating these nanocrystalline regions are random or loosely-oriented domains (blue). The silk protein core is surrounded by a thin hydrophobic lipid or alkane-rich layer (yellow).

References

- 1 J. S. Edgerly, J. A. Davilla and N. Schoenfeld, *Journal of Insect Behavior*, 2002, **15**, 219–242.
- 2 J. S. Edgerly, *The Evolution of Social Behavior in Insects and Arachnids*, Cambridge University Press, 1997, pp. 14–25.
- 3 J. S. Edgerly, S. Büsse and T. Hörschemeyer, *Zoologischer Anzeiger*, 2012, 297–306.
- 4 S. Okada, S. Weisman, H. E. Trueman, S. T. Mudie, V. S. Haritos and T. D. Sutherland, *International Journal of Biological Macromolecules*, 2008, **43**, 271–275.
- 5 M. A. Collin, E. Camama, B. O. Swanson, J. S. Edgerly and C. Y. Hayashi, *Biomacromolecules*, 2009, **10**, 2268–2274.
- 6 M. A. Collin, J. E. Garb, J. S. Edgerly and C. Y. Hayashi, *Insect biochemistry and ...*, 2009, **39**, 79–82.
- 7 J. S. Edgerly, A. TADIMALLA and E. P. DAHLHOFF, *Functional Ecology*, 2005, **19**, 255–262.
- 8 E. S. Ross, *Encyclopedia of Entomology*, Springer Netherlands, 2008, pp. 4169–4172.
- 9 K. B. Miller, C. Y. Hayashi, M. F. Whiting, G. J. Svenson and J. S. Edgerly, *Systematic Entomology*, 2012, **37**, 550–570.
- 10 M. A. Collin, J. S. Edgerly and C. Y. Hayashi, *Zoology (Jena, Germany)*, 2011, **114**, 239–246.
- 11 R. V. Lewis, *Chemical reviews*, 2006, **106**, 3762–3774.
- 12 G. H. Altman, F. Diaz, C. Jakuba, T. Calabro, R. L. Horan, J. Chen, H. Lu, J. Richmond and D. L. Kaplan, *Biomaterials*, 2003, **24**, 401–416.
- 13 T. D. Sutherland, J. H. Young, S. Weisman, C. Y. Hayashi and D. J. Merritt, *Annual review of entomology*, 2010, **55**, 171–188.
- 14 C. Z. Zhou, F. Confalonieri, N. Medina, Y. Zivanovic, C. Esnault, T. Yang, M. Jacquet, J. Janin, M. Duguet, R. Perasso and Z. G. Li, *Nucleic Acids Research*, 2000, **28**, 2413–2419.
- 15 M. Xu and R. V. Lewis, *Proceedings of the National Academy of Sciences of the United States of America*, 1990, **87**, 7120–7124.
- 16 S. Keten, Z. Xu, B. Ihle and M. J. Buehler, *Nature Materials*, 2010, **9**, 359–367.
- 17 S. Keten and M. J. Buehler, *Applied physics letters*, 2010, **96**, 153701.
- 18 C. Y. Hayashi, N. H. Shipley and R. V. Lewis, *International Journal of Biological Macromolecules*, 1999, **24**, 271–275.
- 19 J. Johansson, C. Nerelius, H. Willander and J. Presto, *Biochemical and biophysical research communications*, 2010, **402**, 515–518.
- 20 R. W. Williams, A. Chang, D. Juretić and S. Loughran, *Biochimica et Biophysica Acta*, 1987, **916**, 200–204.
- 21 D. L. Minor, Jr and P. S. Kim, *Nature*, 1994, **367**, 660–663.
- 22 D. Grubb and L. Jelinski, *Macromolecules*, 1997, **30**, 2860–2867.
- 23 S. Sampath, T. Isdebski, J. E. Jenkins, J. V. Ayon, R. W. Henning, J. P. R. O. Orgel, O. Antipoa and J. L. Yarger, *Soft Matter*, 2012, **8**, 6713–6722.
- 24 J. E. Jenkins, S. Sampath, E. Butler, J. Kim, R. W. Henning, G. P. Holland and J. L. Yarger, *Biomacromolecules*, 2013, **14**, 3472–3483.
- 25 K. Liang, Y. Gong, J. Fu, S. Yan, Y. Tan, R. Du, X. Xing, G. Mo, Z. Chen, Q. Cai, D. Sun and Z. Wu, *International Journal of Biological Macromolecules*, 2013, **57**, 99–104.
- 26 T. Asakura, J. Yao, T. Yamane, K. Umemura and A. S. Ulrich, *J. Am. Chem. Soc.*, 2002, **124**, 8794–8795.
- 27 J. E. Jenkins, M. S. Creager, R. V. Lewis, G. P. Holland and J. L. Yarger, *Biomacromolecules*, 2009, **11**, 192–200.
- 28 A. Sponner, W. Vater, S. Monajembashi, E. Unger, F. Grosse and K. Weisshart, *PLoS ONE*, 2007, **2**, e998.
- 29 G. P. Holland, R. V. Lewis and J. L. Yarger, *J. Am. Chem. Soc.*, 2004, **126**, 5867–5872.
- 30 Z. Yang, D. Grubb and L. Jelinski, *Macromolecules*, 1997, **30**, 8254–8261.
- 31 C. Riekel, C. Bränden, C. Craig, C. Ferrero, F. Heidelbach and M. Müller, *International Journal of Biological Macromolecules*, 1999, **24**, 179–186.
- 32 F. Vollrath, T. Holtet, H. C. Thogersen and S. Frische, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 1996, **263**, 147–151.
- 33 S. Frische, Maunsbach, AB and F. Vollrath, *Journal of Microscopy*, 1998, **189**, 64–70.
- 34 S. Schulz, *Lipids*, 2001, **36**, 637–647.
- 35 E. Victoriano, D. O. Pinheiro and E. A. Gregório, *Neotropical entomology*, 2007, **36**, 707–711.
- 36 S. Weisman, H. E. Trueman, S. T. Mudie, J. S. Church, T. D. Sutherland and V. S. Haritos, *Biomacromolecules*, 2008, **9**, 3065–3069.
- 37 D. Nečas and P. Klapetek, *Central European Journal of Physics*, 2011, **10**, 181–188.
- 38 A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi and R. G. Griffin, *The Journal of Chemical Physics*, 1995, **103**, 6951.
- 39 R. N. A. H. Lewis and R. N. McElhane, *Methods in molecular biology (Clifton, N.J.)*, 2007, **400**, 207–226.
- 40 M. Boulet-Audet, T. Lefèvre, T. Buffeteau and M. Pézolet, *Applied Spectroscopy*, 2008, **62**, 956–962.
- 41 J. Shao, J. Zheng, J. Liu and C. M. Carr, *Journal of applied polymer science*, 2005, **96**, 1999–2004.
- 42 S. Ling, Z. Qi, D. P. Knight, Z. Shao and X. Chen, *Biomacromolecules*, 2011, **12**, 3344–3349.
- 43 S. Ling, Z. Qi, D. P. Knight, Y. Huang, L. Huang, H. Zhou, Z. Shao and X. Chen, *Biomacromolecules*, 2013, **14**, 1885–1892.
- 44 P. Papadopoulos, R. Ene, I. Weidner and F. Kremer, *Macromolecular Rapid Communications*, 2009, **30**, 851–857.
- 45 N. N. Ashton, D. R. Roe, R. B. Weiss, T. E. Cheatham, III and R. J. Stewart, *Biomacromolecules*, 2013, **14**, 3668–3681.
- 46 A. A. Walker, J. S. Church, A. L. Woodhead and T. D. Sutherland, *Insect biochemistry and molecular biology*, 2013, **43**, 572–579.
- 47 P. Papadopoulos, J. Sölter and F. Kremer, *The European physical journal. E, Soft matter*, 2007, **24**, 193–199.
- 48 V. SYu and N. N. Kalnin, *Biopolymers*, 1990, **30**, 1259–1271.
- 49 J. O. Warwicker, *Journal of Molecular Biology*, 1960, **2**, 350–362.
- 50 H. Saito, *Magnetic Resonance in Chemistry*, 1986, **24**, 835–852.
- 51 A. Shoji, T. Ozaki, H. Saito, R. Tabeta and I. Ando, *Macromolecules*, 1984, **17**, 1472–1479.
- 52 T. Asakura, K. Suita, T. Kameda, S. Afonin and A. S. Ulrich, *Magnetic Resonance in Chemistry*, 2004, **42**, 258–266.
- 53 T. Asakura, R. Sugino, T. Okumura and Y. Nakazawa, *Protein science : a publication of the Protein Society*, 2002, **11**, 1873–1877.
- 54 D. Massiot, F. Fayon, M. Capron, I. King, S. Le Calvé, B. Alonso, J. O. Durand, B. Bujoli, Z. Gan and G. Hoatson, *Magnetic Resonance in Chemistry*, 2002, **40**, 70–76.
- 55 M. S. Creager, J. E. Jenkins, L. A. Thagard-Yeaman, A. E. Brooks, J. A. Jones, R. V. Lewis, G. P. Holland and J. L. Yarger, *Biomacromolecules*, 2010, **11**, 2039–2043.
- 56 G. P. Holland, J. E. Jenkins, M. S. Creager, R. V. Lewis and J. L. Yarger, *Biomacromolecules*, 2008, **9**, 651–657.
- 57 Z. Yang, O. Liivak, A. Seidel, G. LaVerde, D. Zax and L. Jelinski, *J. Am. Chem. Soc.*, 2000, **122**, 9019–9025.
- 58 J. B. Addison, N. N. Ashton, W. S. Weber, R. J. Stewart, G. P. Holland and J. L. Yarger, *Biomacromolecules*, 2013, **14**, 1140–1148.