Article

Structural Diversity of Native Major Ampullate, Minor Ampullate, Cylindriform, and Flagelliform Silk Proteins in Solution

Imke Greving,* Ann E. Terry,* Chris Holland, Maxime Boulet-Audet, Isabelle Grillo, Fritz Vollrath, and Cedric Dicko*

Read Online

| ACCESS | Metrics & More | Article Recommendations | Supporting Information |

ABSTRACT: The foundations of silk spinning, the structure, storage, and activation of silk proteins, remain highly debated. By combining solution small-angle neutron and X-ray scattering (SANS and SAXS) alongside circular dichroism (CD), we reveal a shape anisotropy of the four principal native spider silk feedstocks from *Nephila edulis*. We show that these proteins behave in solution like elongated semiflexible polymers with locally rigid sections. We demonstrated that minor ampullate and cylindriform proteins adopt a monomeric conformation, while major ampullate and flagelliform proteins have a preference for dimerization. From an evolutionary perspective, we propose that such dimerization arose to help the processing of disordered silk proteins. Collectively, our results provide insights into the molecular-scale processing of silk, uncovering a degree of evolutionary convergence in protein structures and chemistry that supports the macroscale micellar/pseudo liquid crystalline spinning mechanisms proposed by the community.
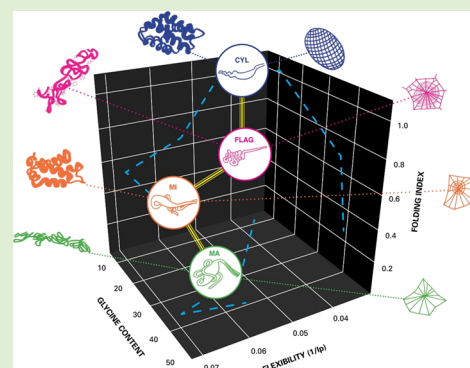
INTRODUCTION

Biological materials are typically grown while silks are spun.[1] The process of silk spinning has been shown to have a significant overlap with and can be taken as a model for industrial polymer processing whereby a liquid feedstock undergoes solidification into a fiber as a result of pultrusion.[1−3] Silk fibers start as aqueous protein melts[2,4,5] originating from bespoke glands, ducts, and spigots with each type of silk having a specific protein sequence.[6,7] Not surprisingly, there has been considerable debate about whether the secret of a spider's unique ability to tune silk fiber properties is the primary amino acid sequence[4,5] or the spinning process.[6,7] Most likely, it is a combination supported by two recent studies showing that major ampullate silk proteins are packed, and their reactivity modulated, in micellar to granular subunits.[8,9] Indeed, the amphipaticity of silk has, for long, been cited as the reason for molecular un- and re-folding.[10,11]

One barrier to a better understanding of the full process has been the strong bias toward studying in spiders the major ampullate dragline silk and in insects the cocoon silk of the mulberry worm *Bombyx mori*. The focus on these two silks precluded any generalization of both the chemical composition of the dope and mechanics of the spinning process. Looking at independently evolved lineages of silk-producing organisms, Walker et al.[12] suggested a convergence toward the occurrence of liquid-crystal intermediates (mesophases) to reduce the viscosity of the silk dope and assist in the formation of the supramolecular structure.[13,14] However, the exact nature of the liquid-crystal-forming units (mesogens) in the silk dope and the relationship between liquid crystallinity, protein structures, and interaction followed by the processing of the dope are yet to be fully elucidated.[15] A starting point for this will be a deeper understanding of the solution behavior of the four most studied spider silk proteins from a structural and colloidal perspective.

Here, we seek to uncover the correlations between the size, shape, and structure of four distinct spider silk proteins before they are spun into a fiber. Previously, a range of structural characterization techniques have been deployed to uncover this relationship; however, sample availability, preparation, and lack of complementary techniques have hampered a consistent evaluation of the observations.[16,17] We update the state-of-the-art by combining small-angle scattering (SAS) and circular dichroism (CD) with a semiflexible theoretical model to examine native silk proteins to allow us to draw salient relationships between folding and chain flexibility and the implication for spinning.
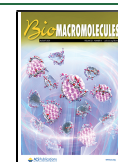
■ MATERIALS AND METHODS

Spiders and Protein Extractions. *Nephila edulis* (Tetragnathidae) golden-orb spiders were raised in a greenhouse under controlled humidity and temperature conditions. The major (MA) and minor (MI) ampullate, flagelliform (FLAG), and cylindriform (CYL) glands were retrieved by dissection of mature female spiders and transferred into milliQ water at 6 °C. The glands were gently peeled to remove the epithelium (except for FLAG, where the epithelium was separated by gravity after overnight dissolution), and the contents of the lumen were gently blotted and placed in Eppendorf tubes. Additional milliQ water was added to fill the tubes to exclude air, and the proteins were left to dissolve for 17 h at 6 °C. MA, FLAG, and CYL completely dissolved, and the resulting solutions were visually transparent and homogeneous; however, MI did not dissolve completely, and the obtained solutions showed a tendency to flocculate. For MI, only the dissolved portion of the material in the Eppendorf tube was used as the stock solution for further measurements resulting in a much smaller sample volume. Two sets of stock solutions for each type of silk were prepared on different occasions for the experimental measurements. Before use, if a sample showed visual aggregation, it was not used for measurements or data analysis.

The protein concentrations of the stock solutions were calculated from dry weight estimates following the drying of an aliquot of stock solution in a vacuum oven for 2 h at 80 °C. The error in the protein concentration by this method was estimated to be ±0.1 mg/mL. The MA stock solutions had concentrations of 18.1 and 20.0 mg/mL, the MI had concentrations of 1.4 and 2.7 mg/mL, and the CYL had concentrations of 3.0 and 1.5 mg/mL. Due to the low volume of one FLAG stock solution, the protein concentration was estimated from the SAXS intensity of another FLAG stock solution of 0.3 mg/mL, scaled to overlap in the higher-q region. From the stock solutions of each protein, a dilution series was prepared by weight to the stated dilution ratio (sample-1 = stock solution; sample-2 = dilution 1:1; sample-4 = dilution 1:4; sample-8 = 1:8; sample-16 = 1:16). Note that the pH for all silk solutions was not buffered, resulting in a measured pH of approximately 6.5 for all the solutions.

Circular Dichroism. Circular dichroism (CD) spectra were measured for all four silk stock solutions using a Jasco 810 spectrometer at the EMBL (Grenoble) before the SAS experiments, about 17 h after initial dissection of the spiders. The samples were measured in quartz cells (Hellma) of 0.01 mm path length for MA, which had the highest protein concentration, and 0.1 mm path length for MI, FLAG, and CYL. Three consecutive spectra were recorded and averaged using a 1 nm resolution step and a 200 nm/min scan rate. The data were reduced using the CDTool software.[18]

Small-Angle Scattering. Small-angle X-ray scattering (SAXS) was recorded at the Bio-SAXS beamline ID14−3 at the ESRF (Grenoble). Each sample of the dilution series was gently pipetted into separate wells of a standard 96 well plate, sealed with Cristal tape. The samples were transferred to the measuring point, a quartz 1 mm-diameter capillary, using a modified HPLC pump system.[19] Between each sample, the capillary, syringe, and connecting tubing were flushed clean with 12 M urea solution before rinsing with milliQ water and drying. The SAXS was recorded for the water solution between each silk sample to confirm that the capillary was thoroughly clean before the next measurement and that no residue of silk was left following the cleaning step. The SAXS data were collected using a Pilatus 1 M X-ray detector at a fixed sample detector distance of 2.43 m and incident radiation wavelength of 0.93 Å, giving an accessible q range of 0.05−5.8 nm^{-1}.[19] A test was made to estimate when radiation damage of the silk proteins might be significant. At data collection exposure times greater than 5 s, evident changes were observed in the SAXS pattern.

For this reason, exposure times were kept to a maximum of 1 s, and multiple exposures (up to 50 frames per sample) were taken while continuously flowing the sample through the capillary using a minimum flow rate of 0.5 μL/s to additionally avoid radiation damage. At these flow rates, we did not observe any orientation in the 2D scattering patterns and concluded that there was no significant effect of shear or extension. The data were reduced using the ID14−3 data reduction pipeline[19] with bovine serum albumin serving as a calibration standard for M_w. The obtained three-column Ascii files for up to 50 frames per sample were averaged using PRIMUS[20] after the exclusion of anomalous data due to empty capillary or bubbles in the sample.

Time-of-flight small-angle neutron scattering (SANS) measurements were conducted at SANS2d, ISIS Facility, U.K. On SANS2d, the collimation can be optimized for the sample-to-detector distance, which in these experiments was 4 m. By selecting 1.75 to 16.5 Å wavelength neutrons, the q range is 0.004 to 1.5 Å^{-1}, recorded by a single 0.96 m2 Ordela 21000 N detector. Complementary SANS measurements were performed at the D11 beamline at the Institut Laue Langevin (ILL). The energy was set to 6 Å and the sample was scanned at 2, 10, 28 m sample detector distances. The samples were prepared in the same manner as for the SAXS measurements with degassed milliQ water as the solvent. As the main spider silk glands are duplicated in a spider, both glands were used to prepare a single stock solution.

Further dilution was also with milliQ water. 400 μL of each sample was prepared in this way and sealed in quartz standard Hellma cells of 1 mm path length. The cells were placed in a thermostated and automated sample changer at 20 °C. Small-angle neutron scattering data were collected for 15 min from each sample, and the process was repeated until sufficient signal-to-noise ratios were achieved, with typical collection times ranging from 30 min to 2 h depending on the protein concentration. During the SAXS measurements, the sample was static unlike the SAXS samples. In this manner, instabilities or changes in the sample could be monitored. The neutron scattering contrast came from the inherent scattering length density difference of protein to water with no additional contrast variation using deuterium exchange. All data reduction and background subtraction were made using Mantid (http://dx.doi.org/10.5286/software/mantid). The scattering from a partially deuterated polymer standard allowed data to be normalized to absolute intensities.

Data Analysis. Both SAXS and SANS data were initially interpreted by applying the Guinier approximation using PRIMUS.[20] From a plot of ($\ln I(q)$ vs q^2), the slope can be used to determine the radius of gyration (R_g). The projected y-intercept yields the intensity at $q = 0$ nm^{-1}, $I(0)$, which can be used to calculate the molecular weight, M_w, of the scattering entity.

The $p(r)$ function or distance distribution function describes the distances between points within an object. In a protein ensemble, this proves to be useful for visualizing conformational changes as small changes in the relative positions of a few residues can be resolved in the shape of the $p(r)$ distribution. The $p(r)$ function obtained by two methods were compared: by the indirect Fourier transformation method (using GNOM[21]) and by Bayesian statistics (BayesApp[22]). In principle, the advantage of the Bayesian statistics over GNOM is that the maximum dimension of the scattering object (D_{max}) is estimated from the program with no user constraints. In contrast, in GNOM, the standard settings were chosen for $p(r)$, and D_{max} was estimated visually from the fit[23] using a q_{max} of 0.4 nm^{-1}. The cross-sectional $p(r)c$ was obtained using Bayesian statistics (BayesApp[22]) on curves truncated at q^*. In our results, both approaches gave similar results, and the values from GNOM were chosen as it is the more conventional approach.

The slope of the plot of ($\ln I(q)$ vs $\ln q$) provides information about the local interface and fractal dimension of the scattering entity. At high q, the Porod region of the scattering curve, a slope of −2 is representative of a Gaussian chain in a dilute solution, whereas a slope of −1 signifies rigid rods. A slope of −4 represents an interface or surface, which is smooth, and between −3 and −4 is an interface that is seen as rough at that length scale.[24]

It is also possible to fit the form factor $p(q)$ of a scattering entity in dilute solution using a mathematical expression. The silk proteins were found to be similar to a worm-like chain as they could be fitted using the flexible cylinder model of the SASView program (method 3 in ref 25). Fitting of the cross section was also performed with SASVIEW (http://www.sasview.org).

■ RESULTS AND DISCUSSION

Molecular Weight and Tertiary Structure. SAS analysis provides three critical pieces of information of relevance for understanding the prespun silk dope: (i) the molecular weight of the scattering entities, (ii) their shape, and (iii) the local chain behavior.[26] Figure 1 summarizes the typical SAS profiles
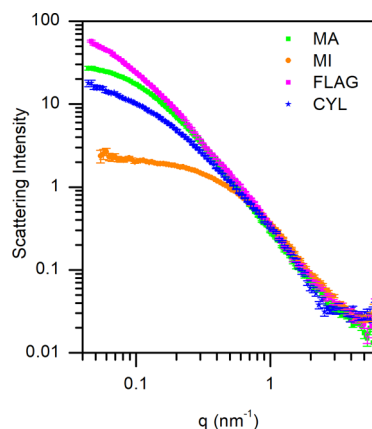

Figure 1. A direct comparison of the SAXS profiles of the four spider silk proteins at comparable protein concentration reveals differences in the overall shape of the protein (intensity normalized for protein concentration). The concentrations are 2.8, 2.7, 1.5, and 1.5 mg/mL for major ampullate (MA), minor ampullate (MI), flagelliform (FLAG), and cylindriform (CYL), respectively.

for the four principal spider silk proteins, indicating differences in the overall shape of the protein in solution comparing their SAS scattering curves. At the concentrations compared in Figure 1, we observed no structure factor (see Figure S1a−d for the concentration series, Supporting Information).

The extrapolated molecular weight (M_w) was 1.2 MDa for flagelliform (FLAG), indicating a dimeric state in solution by comparison to the 570 kDa estimated from the amino acid sequence.[27] The new and more complete flagelliform protein's gene sequences[28] suggest a molecular weight of 759.5 kDa for the monomeric units. Given the uncertainty on the extrapolated M_w estimate, we believed that FLAG was dimeric in solution.

Similarly, comparing our calculated molecular weight for major ampullate (MA) (527−560 kDa) and the observed weight from the SDS Page at 350 kDa,[29] we confirm that MA is also in a dimeric form. As a low protein concentration is used in this study, we hypothesize that the dimer of MA is formed from the C-terminal linkage.[30] In contrast to FLAG and MA, cylindriform and minor ampullate (CYL and MI, respectively) are in a monomeric state; for CYL, a M_w of 300−320 kDa is obtained from the SAXS data, and for MI, a M_w of 225 kDa is obtained.[a] Our results agree well with the literature where the molecular weights, calculated from their primary sequences, are 370−480 kDa for CYL[31] and 250−315 kDa for MI.[32] Interestingly, the full-length gene sequences for CYL[33] and MI[34] silks from *Araneus* suggest much lower molecular weights at 213.2 and 201.3 kDa, respectively. Overall, the new genomic and proteomic information[33−37] suggested molecular weights to be systematically lower than the experimentally determined one (e.g., this work and others[8,38,39]). Looking at the most studied silk, namely, the MA silks, the transcriptome[40] reveals 18 to 29 proteins identified as the spidroin. More interestingly, the experimentally agreed M_w of 250−350 kDa for MA might

be due to the oligomerization of smaller spidroins in the 80 to 90 kDa M_w range.[36] This implied a heterogeneous protein composition and, consequently, heterogeneity in the fiber properties. Note, however, that the heterogeneity was not reflected in the SAS data. All curves suggested monodispersed entities.

Shape Anisotropy. Two transformations can be applied to the SAS data presented in Figure 1: a Fourier transformation to obtain the pair distribution function $p(r)$ and a Kratky transformation to estimate the folding and flexibility of the silks.[26] The pair distribution function $p(r)$ for the silk proteins in aqueous solution (Figure 2) allows the estimation of the radius of gyration (R_g) and maximum size (D_{max}). Table 1 summarizes the findings.
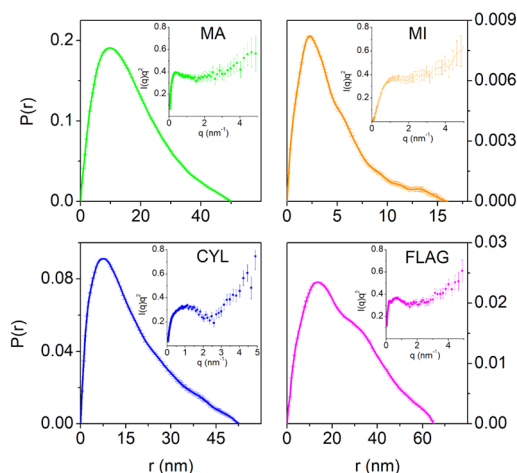

Figure 2. $p(r)$ curves for the four types of spider silk protein obtained from GNOM.[21] Assuming that each system was monodisperse, D_{max} was evaluated from the best fit, especially at low-q values. The $p(r)$ of all silks suggests an anisotropic/elongated structure. Interestingly, the $p(r)$ for the FLAG silk shows a multidomain or dumb-bell shape for the protein in solution.[41] The insets show the Kratky plots of the four silks, suggesting flexibility indicated by the increase in $I*q^2$ at high-q values.

For all four silks, we found that the ratio R_g/D_{max} was smaller than 0.39, a typical value for a spherically shaped particle,[42] thus indicating that they are anisotropic in solution. The shape of the $p(r)$ confirms an anisotropic shape for MA, MI, and CYL. At the same time, FLAG exhibits an extended structure and a $p(r)$ function typical of a dumb-bell-shaped protein or protein complex (Figure 2), perhaps driven by the bulky side chains favoring an extended structure of the protein in solution.[43]

Structural Plasticity: Scattering. The second transformation, the Kratky plot (Figure 2 inset), is indicative of the compactness of the scattering silks in solution. The four silks (Figure 2 insets) display a peak at lower-q values, as typically observed for partially folded proteins, suggesting that these silks are not merely random coils. Importantly, the increasing intensity at higher-q values is indicative of high molecular flexibility.[44]

The scattering data show that spider silk proteins in solution are elongated and flexible. To better understand the nature of the silk proteins' structures in solution, we analyzed their respective CD spectrum. Figure 3 shows the specific secondary structure profiles of the four silks. The results suggest that the silk proteins were not in a random coil conformation (as

Table 1. Data Summary

silk	function	repeat motif	secondary structure	M_w (kDa)	R_g (Å)[d]	ratio R_g /D_{max}	folding index	l_p (Å)[e]	R_{gc} (Å)[f]
			intrinsic properties of silk proteins in solution				behavior of silk proteins in solution		
MA	dragline, structural web	$(GA)_n/(A)_n$GPGGX/ GPGQQ	PP II[a]	527–565	133 ± 1	0.3	0.25 ± 0.01	17.2 ± 2.8	7.4 ± 3.4
MI	auxiliary spiral	$(GGX)_n/(GA)_n$	3_{10} helix	250–330[b]	60–80	0.24	0.54 ± 0.03	14.9 ± 0.9	8.8 ± 0.1
FLAG	core fibers of capture spiral	GPGGX/GGX	β-spiral, 3_1 helix	1200[c]	200 ± 11	~0.31	0.64 ± 0.03	19.0 ± 1.7	7.5 ± 3.4
CYL	egg sac	$(A)_n/(S)_n/(SA)_n$	α-helix-like	300–321	141 ± 5	0.27	0.99 ± 0.01	27.7 ± 0.5	13.5 ± 0.1

[a]Poly-proline II. [b]MI extrapolated from SANS. [c]Estimate from extrapolated $I(0)$. [d]R_g obtained from $p(r)$. [e]Persistence length. [f]Radius of the cross section.
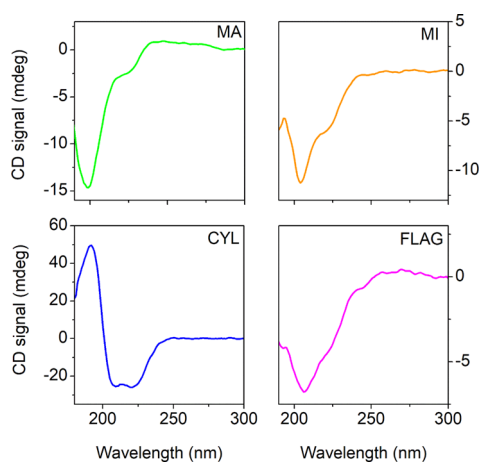

Figure 3. All stock solutions were measured by circular dichroism (CD) before the SAXS experiments to check the quality of the protein solutions and determine the folding state. The CD spectra of MA, MI, FLAG, and CYL silk proteins in H_2O confirm apparent structural differences between the four silks, as reported in the literature.[43] MA showed a strong negative peak at 199 nm, indicating a predominantly random coil dominated structure. It also showed a plateau at 217 nm, suggesting the presence of residual β-sheet-like structures (sheets and turns). The spectrum indicated a polyproline II conformation for MA.[43] The spectrum of MI suggested a 3_1-helix structure, showing a strong negative peak at around 208 nm. The spectrum of FLAG indicates a β-spiral conformation as predicted by Zhou et al.,[47] while CYL showed an α-helical dominated spectrum.[43]

previously proposed[45,46]), indicating partial folding for MA and FLAG and confirming folding for MI and CYL.[43,45] A question remains: How are the secondary structures related to the measured scattering data?

Structural Plasticity: Circular Dichroism. The CD spectra also allow us to probe the structural plasticity of silk proteins; based on the analysis of a wide range of silks and their processing systems,[45] we define a conformational flexibility marker termed the folding index. The index is defined as the ratio of the two CD minima found at around 220 (indicative of the amount of folding) and 200 nm (characteristic of the amount of disordered-like structures, see Figure 3). A folding index higher than 0.9 is typical for a folded helical structure (e.g., myoglobin), while a value below 0.5 denotes a partially to fully unfolded structure.[45]

We observe an increasing proportion of folded structures in the silk proteins from MA (0.249 ± 0.01), MI, FLAG to CYL (0.991 ± 0.01). In our previous study, this was linked to the content ratio of glycine and correlated with an increase in conformational flexibility.[45]

To explore now the generic intrinsic properties of silks (i.e., elongated particles, flexibility, and secondary structure profiles) and their interplay, we used a semiflexible model to describe the structural behavior of silk protein in solution. A Holtzer plot ($I(q)q$ as a function of q) confirmed this semiflexible nature of silk proteins (Figure S2, Supporting Information), and indeed, we were able to fit the SAXS curves to the form factor for a flexible worm-like chain (Figure S3, Supporting Information).[25] This model considers the molecular chain as an articulated series of rigid cylinders of length l_p (persistence length) and radius R_{gc} (radius of the cylinder cross section).

Our calculated persistence lengths, l_p, are in the range of 15–19 Å for MA, MI, and FLAG, and l_p is 27.8 Å for CYL (see Table 1). For comparison, a typical protein chain adopting random conformations yields a shorter l_p of 9.35 Å ($\approx$ three amino acids),[44] giving clear evidence that silks do not take such a random structural conformation. Importantly, we found that regardless of the disparity of l_p and R_{gc} (Table 1), the aspect ratio (l_p/R_{gc}) of the rigid units was remarkably consistent for the four silks between 2 and 2.5. The aspect ratio (l_p/R_{gc}) could be indicative of convergent evolution in silks' proteins molecular design regardless of chemistry and spinning.

To understand the origin of the local rigidity, we combine (where available) sequence knowledge and secondary structures. The known repetitive motifs may provide a starting point.[5] In the case of MA and FLAG with a core sequence of GPGXX, we estimate a reasonable persistence length of between 17 and 19 Å from the distance between two proline residues, which is in good agreement with our findings. Here, the proline residue plays a vital role as a chain disrupter and interestingly in exposing the GXX moieties to the solvent. In such conformers, one could expect more hydrophobic interactions.

For CYL, the l_p was 27.8 ± 0.5 Å, significantly larger than the other three silks. Lin et al.[48] solved the structure of the repeating domain of CYL and found that the repeating units consisted of five co-aligned helices in a supersecondary structure. The units reported were up to 200 amino acids with a total length of 30 Å, remarkably close to our findings for the l_p value. For MI, a new but partial NMR structure[49] confirmed the helical folding propensity (see Table 1) and an estimate repeat unit length of about 30 Å. Interestingly, the MI repeat unit was about twice the persistence length. This suggested a less rigid repeat unit as compared to CYL. The data indicated that the four silks behave like semiflexible polymers and that the origin of local chain rigidity (l_p) can be traced to sequence and folding.

Rauscher et al.,[50] in a seminal study, found that a minimum threshold of combined proline and glycine amino acid content appeared to be fulfilled by proteins forming such diverse
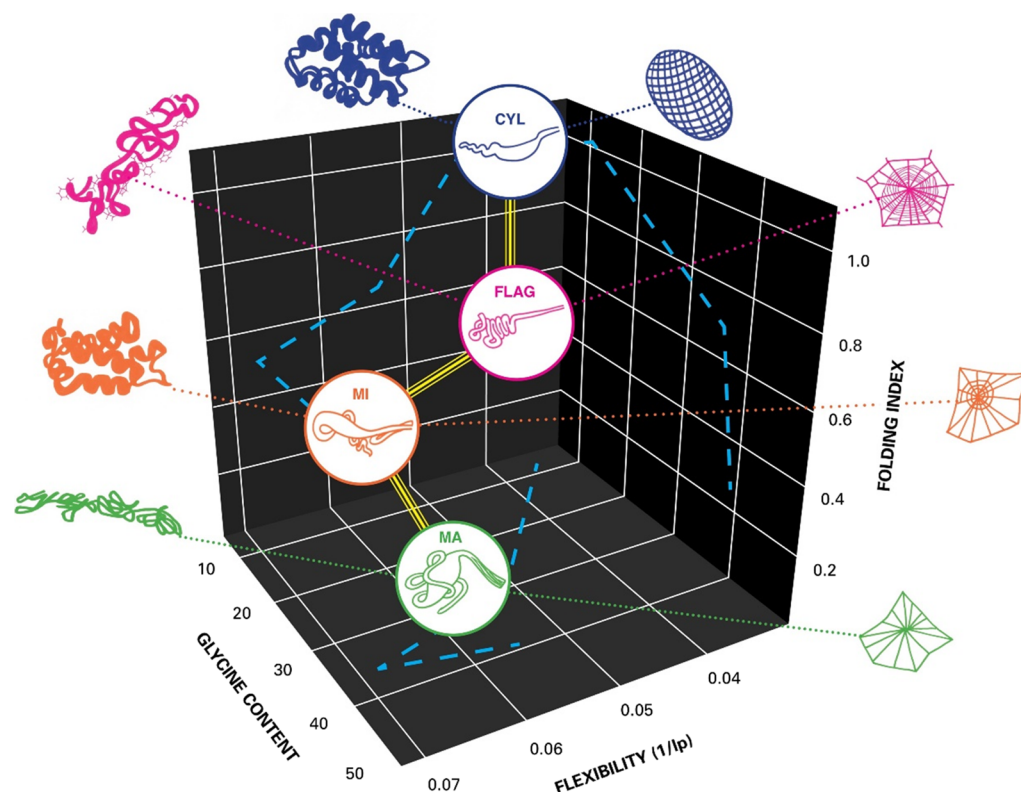
Figure 4. Combined plot of flexibility ($1/l_p$ obtained by scattering), glycine content (obtained from amino acid analysis), and folding index (derived from CD spectroscopy). Overlaid to these parameters are the glands from which the silks were extracted, the known motifs (see text) for the different silk proteins, and the final function (on the web or as a cocoon). From top to bottom: CYL (cocoon), FLAG (sticky spiral), MI (auxiliary nonsticky spiral), and MA (radial threads and dragline). The dotted blue lines are the 2D projections corresponding to the data presented in this work.

biomaterials as the human aorta, spider silk, and lizard eggshells. The combination of rigidity imparted by the proline and conformational plasticity conferred by the glycine residues suggested that maintaining a critical level of structural disorder is not only a fundamental requirement but may very well constitute the single most essential design principle of self-assembling elastic proteins. Recently, the role of the elusive polyproline II (PPII) helix conformation in the glycine-rich region in MA silk has been proposed as a soluble prefibrillar region that subsequently undergoes intramolecular interactions.[51] These findings unravel the importance of glycine-rich structures to mediate the initial step of and possibly explain how the extremely rapid process of β-sheet formation during spider silk assembly can be modulated to prevent catastrophic aggregation.

We, therefore, in Figure 4, attempted to capture the physicochemical prerequisites for the silk proteins and spinning process.

We combined the structural flexibility found by SAXS ($1/l_p$), conformational flexibility imparted by the glycine content, and the folding found by CD (folding index). Overlaid are the spinning apparatus, the structural motifs, and functions found in all four silks. Missing are of course the spinneret's evolution.[52] Given that FLAG must be the newest invention,[53−56] and MI is a sister clade to MA,[56−58] our results suggest that dimerization has evolved to enable the processing of disordered-like MA and FLAG silk proteins.

■ CONCLUSIONS

We measured, analyzed, and outlined the difference in overall molecular shape as well as the structure and the semiflexible nature of the four principal spider silk proteins in an attempt to determine the fundamental components of silk solution behavior. These new insights provide a unique window into the molecular origin of silk's ability to readily self-assemble and, in the process, mediate "low energy spinning". Our findings, summarized in Table 1, suggest that the nature and dimensions of the rigid segment tend to be strongly dependent on the chemical structure and local interaction of the protein's chain. The difference in the secondary structure content of the four native spider silk proteins (indicated by CD) is also reflected in the overall shape of each of the silk proteins suggesting a difference in the tertiary structure. The global anisotropy ratios for the four silks, however, are remarkably similar.

In a more generic context, we propose that secondary structures and their interactions into larger structures provide fibrous proteins with "handles" that ensure a correct hydrogen bond density in the ordered crystalline and disordered amorphous regions of the fibers. The controls could be a combination of protein concentration, terminal domain dimerization, ionic strength, and pH gradients, as well as a multiprotein component to spin successfully. From an evolutionary perspective, we propose that dimerization was introduced to help the processing of disordered-like silk proteins. We conclude that silks provide the student of molecular structure−property−function relationships with a

unique model material that can be studied in greater detail than most, perhaps any, other biological materials.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biomac.0c00819.

> Small-angle scattering concentration series; Holtzer plots; worm-like chain fits (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Imke Greving − *Institute of Materials Research, Helmholtz Zentrum Geesthacht, 21502 Geesthacht, Germany;*
Email: imke.greving@hzg.de

Ann E. Terry − *MAX IV Laboratory, 224 84 Lund, Sweden; Lund Institute of Advanced Neutron and X-ray Science, 223 70 Lund, Sweden;* Email: ann.terry@maxvi.lu.se

Cedric Dicko − *Pure and Applied Biochemistry, Department of Chemistry, Lund University, SE-221 00 Lund, Sweden; Lund Institute of Advanced Neutron and X-ray Science, 223 70 Lund, Sweden;* ⓞ orcid.org/0000-0001-6377-3500;
Email: cedric.dicko@tbiokem.lth.se

Authors

Chris Holland − *Department of Materials Science and Engineering, The University of Sheffield, Sheffield S1 3JD, United Kingdom;* ⓞ orcid.org/0000-0003-0913-2221

Maxime Boulet-Audet − *Department of Zoology, University of Oxford, Oxford OX1 2JD, United Kingdom*

Isabelle Grillo − *Institut Laue Langevin, 38000 Grenoble, France*

Fritz Vollrath − *Department of Zoology, University of Oxford, Oxford OX1 3SZ, United Kingdom*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.biomac.0c00819

Author Contributions

The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

Funding

Notes

The authors declare no competing financial interest.
◆Deceased (Isabelle Grillo).

■ ADDITIONAL NOTE

[a]MI M_w from SANS data, absolute intensity.

■ REFERENCES

(1) Vollrath, F.; Porter, D.; Holland, C. The science of silks. *MRS Bull.* **2013**, *38*, 73−80.

(2) Holland, C.; Vollrath, F.; Ryan, A. J.; Mykhaylyk, O. O. Silk and Synthetic Polymers: Reconciling 100 Degrees of Separation. *Adv. Mater.* **2012**, *24*, 105−109.

(3) Sparkes, J.; Holland, C. Analysis of the pressure requirements for silk spinning reveals a pultrusion dominated process. *Nat. Commun.* **2017**, *8*, 594.

(4) Lewis, R. V. Spider silk: the unraveling of a mystery. *Acc. Chem. Res.* **1992**, *25*, 392−398.

(5) Hayashi, C. Y.; Shipley, N. H.; Lewis, R. V. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *Int. J. Biol. Macromol.* **1999**, *24*, 271−275.

(6) Porter, D.; Vollrath, F. Silk as a Biomimetic Ideal for Structural Polymers. *Adv. Mater.* **2009**, *21*, 487−492.

(7) Vollrath, F.; Porter, D.; Holland, C. There are many more lessons still to be learned from spider silks. *Soft Matter* **2011**, *7*, 9595−9600.

(8) Parent, L. R.; Onofrei, D.; Xu, D.; Stengel, D.; Roehling, J. D.; Addison, J. B.; Forman, C.; Amin, S. A.; Cherry, B. R.; Yarger, J. L.; Gianneschi, N. C.; Holland, G. P. Hierarchical spidroin micellar nanoparticles as the fundamental precursors of spider silks. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 11507.

(9) Lin, T.-Y.; Masunaga, H.; Sato, R.; Malay, A. D.; Toyooka, K.; Hikima, T.; Numata, K. Liquid Crystalline Granules Align in a Hierarchical Structure To Produce Spider Dragline Microfibrils. *Biomacromolecules* **2017**, *18*, 1350−1355.

(10) Jin, H.-J.; Kaplan, D. L. Mechanism of silk processing in insects and spiders. *Nature* **2003**, *424*, 1057−1061.

(11) Knight, D. P.; Vollrath, F. Biological liquid crystal elastomers. *Philos. Trans. R. Soc. London, Ser. B* **2002**, *357*, 155−163.

(12) Walker, A. A.; Holland, C.; Sutherland, T. D. More than one way to spin a crystallite: multiple trajectories through liquid crystallinity to solid silk. *Proc. R. Soc. London, Ser. B* **2015**, *282*, 20150259.

(13) Greving, I.; Cai, M.; Vollrath, F.; Schniepp, H. C. Shear-Induced Self-Assembly of Native Silk Proteins into Fibrils Studied by Atomic Force Microscopy. *Biomacromolecules* **2012**, *13*, 676−682.

(14) Holland, C.; Urbach, J. S.; Blair, D. L. Direct visualization of shear dependent silk fibrillogenesis. *Soft Matter* **2012**, *8*, 2590−2594.

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

(16) Dicko, C.; Kenney, J. M.; Vollrath, F. β-silks: Enhancing and controlling aggregation. *Adv. Protein Chem.* **2006**, *73*, 17−53.

(17) Greving, I.; Dicko, C.; Terry, A.; Callow, P.; Vollrath, F. Small angle neutron scattering of native and reconstituted silk fibroin. *Soft Matter* **2010**, *6*, 4389−4395.

(18) Lees, J. G.; Smith, B. R.; Wien, F.; Miles, A. J.; Wallace, B. A. CDtool—an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Anal. Biochem.* **2004**, *332*, 285−289.

(19) Pernot, P.; Theveneau, P.; Giraud, T.; Fernandes, R. N.; Nurizzo, D.; Spruce, D.; Surr, J.; McSweeney, S.; Round, A.; Felisaz, F.; Foedinger, L.; Gobbo, A.; Huet, J.; Villard, C.; Cipriani, F. New beamline dedicated to solution scattering from biological macromolecules at the ESRF. *J. Phys.: Conf. Ser.* **2010**, *247*, 012009.

(20) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **2003**, *36*, 1277−1282.

(21) Svergun, D. I. Determination of the regularization parameter in indirect - transform methods using perceptual criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495−503.

(22) Hansen, S. BayesApp: a web site for indirect transformation of small-angle scattering data. *J. Appl. Crystallogr.* **2012**, *45*, 566−567.

(23) Svergun, D. I.; Semenyuk, A. V.; Feigin, L. A. Small-angle-scattering-data treatment by the regularization method. *Acta Crystallogr., Sect. A: Found. Adv.* **1988**, *44*, 244−250.

(24) Schmidt-Ott, A. Physical characterization of agglomerated aerosol particles through size-invariant fractal properties. In *Atmospheric Aerosols and Nucleation*; Berlin, Heidelberg, 1988;

Wagner, P. E.; Vali, G., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 1988; pp 113–115, DOI: 10.1007/3-540-50108-8_1028.

(25) Pedersen, J. S.; Schurtenberger, P. Scattering functions of semiflexible polymers with and without excluded volume effects. *Macromolecules* **1996**, *29*, 7602–7612.

(26) Rambo, R. P.; Tainer, J. A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **2013**, *496*, 477–481.

(27) Hayashi, C. Y.; Lewis, R. V. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *J. Mol. Biol.* **1998**, *275*, 773–784.

(28) Stellwagen, S. D.; Renberg, R. L. Toward Spider Glue: Long Read Scaffolding for Extreme Length and Repetitious Silk Family Genes AgSp1 and AgSp2 with Insights into Functional Adaptation. *G3: Genes, Genomes, Genet.* **2019**, *9*, 1909.

(29) Lombardi, S. J.; Kaplan, D. L. The amino acid composition of major ampullate gland silk (dragline) of Nephila clavipes (Araneae, Tetragnathidae). *J. Arachnol.* **1990**, *18*, 297–306.

(30) Hagn, F.; Thamm, C.; Scheibel, T.; Kessler, H. pH-Dependent Dimerization and Salt-Dependent Stabilization of the N-terminal Domain of Spider Dragline Silk-Implications for Fiber Formation. *Angew. Chem., Int. Ed.* **2011**, *50*, 310–313.

(31) Tian, M.; Lewis, R. V. Molecular characterization and evolutionary study of spider tubuliform (eggcase) silk protein. *Biochemistry* **2005**, *44*, 8006–8012.

(32) Colgin, M. A.; Lewis, R. V. Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like "spacer regions". *Protein Sci.* **1998**, *7*, 667–672.

(33) Wen, R.; Liu, X.; Meng, Q. Characterization of full-length tubuliform spidroin gene from Araneus ventricosus. *Int. J. Biol. Macromol.* **2017**, *105*, 702–710.

(34) Chen, G.; Liu, X.; Zhang, Y.; Lin, S.; Yang, Z.; Johansson, J.; Rising, A.; Meng, Q. Full-Length Minor Ampullate Spidroin Gene Sequence. *PLoS One* **2012**, *7*, No. e52293.

(35) dos Santos-Pinto, J. R. A.; Arcuri, H. A.; Lubec, G.; Palma, M. S. Structural characterization of the major ampullate silk spidroin-2 protein produced by the spider Nephila clavipes. *Biochim. Biophys. Acta, Proteins Proteomics* **2016**, *1864*, 1444–1454.

(36) dos Santos-Pinto, J. R. A.; Arcuri, H. A.; Priewalder, H.; Salles, H. C.; Palma, M. S.; Lubec, G. Structural Model for the Spider Silk Protein Spidroin-1. *J. Proteome Res.* **2015**, *14*, 3859–3870.

(37) Ruiz, V.; Jiang, P.; Müller, C.; Jorge, I.; Vázquez, J.; Ridruejo, Á.; Aznar-Cervantes, S. D.; Cenis, J. L.; Messeguer-Olmo, L.; Elices, M.; Guinea, G. V.; Pérez-Rigueiro, J. Preparation and characterization of Nephila clavipes tubuliform silk gut. *Soft Matter* **2019**, *15*, 2960–2970.

(38) Zhang, Y.; Zhao, A.-C.; Sima, Y.-H.; Lu, C.; Xiang, Z.-H.; Nakagaki, M. The molecular structures of major ampullate silk proteins of the wasp spider, Argiope bruennichi: A second blueprint for synthesizing de novo silk. *Comp. Biochem. Physiol., B: Biochem. Mol. Biol.* **2013**, *164*, 151–158.

(39) Sponner, A.; Schlott, B.; Vollrath, F.; Unger, E.; Grosse, F.; Weisshart, K. Characterization of the protein components of Nephila clavipes dragline silk. *Biochemistry* **2005**, *44*, 4727–4736.

(40) Whaite, A. D.; Wang, T.; Macdonald, J.; Cummins, S. F. Major ampullate silk gland transcriptomes and fibre proteomes of the golden orb-weavers, Nephila plumipes and Nephila pilipes (Araneae: Nephilidae). *PLoS One* **2018**, *13*, No. e0204243.

(41) Putnam, C. D.; Hammel, M.; Hura, G. L.; Tainer, J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **2007**, *40*, 191–285.

(42) Durand, E.; Waksman, G.; Receveur-Brechot, V. Structural insights into the membrane-extracted dimeric form of the ATPase TraB from the Escherichia coli pKM101 conjugation system. *BMC Struct. Biol.* **2011**, *11*, 4–4.

(43) Dicko, C.; Knight, D.; Kenney, J. M.; Vollrath, F. Secondary structures and conformational changes in flagelliform, cylindrical, major, and minor ampullate silk proteins. Temperature and concentration effects. *Biomacromolecules* **2004**, *5*, 2105–2115.

(44) Receveur-Brechot, V.; Durand, D. How Random are Intrinsically Disordered Proteins? A Small Angle Scattering Perspective. *Curr. Protein Pept. Sci.* **2012**, *13*, 55–75.

(45) Dicko, C.; Porter, D.; Bond, J.; Kenney, J. M.; Vollrath, F. Structural disorder in silk proteins reveals the emergence of elastomericity. *Biomacromolecules* **2008**, *9*, 216–221.

(46) Lefèvre, T.; Paquet-Mercier, F.; Rioux-Dubé, J.-F.; Pézolet, M. Structure of silk by raman spectromicroscopy: From the spinning glands to the fibers. *Biopolymers* **2012**, *97*, 322–336.

(47) Zhou, Y.; Wu, S.; Conticello, V. P. Genetically directed synthesis and spectroscopic analysis of a protein polymer derived from a flagelliform silk sequence. *Biomacromolecules* **2001**, *2*, 111–125.

(48) Lin, Z.; Huang, W.; Zhang, J.; Fan, J.-S.; Yang, D. Solution structure of eggcase silk protein and its implications for silk fiber formation. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8906–8911.

(49) Gao, Z.; Lin, Z.; Huang, W.; Lai, C. C.; Fan, J.-S.; Yang, D. Structural characterization of minor ampullate spidroin domains and their distinct roles in fibroin solubility and fiber formation. *PLoS One* **2013**, No. e56142.

(50) Rauscher, S.; Baud, S.; Miao, M.; Keeley, F. W.; Pomès, R. Proline and Glycine Control Protein Self-Organization into Elastomeric or Amyloid Fibrils. *Structure* **2006**, *14*, 1667–1676.

(51) Oktaviani, N. A.; Matsugami, A.; Malay, A. D.; Hayashi, F.; Kaplan, D. L.; Numata, K. Conformation and dynamics of soluble repetitive domain elucidates the initial β-sheet formation of spider silk. *Nat. Commun.* **2018**, *9*, 2121.

(52) Mariano-Martins, P.; Lo-Man-Hung, N.; Torres, T. T. Evolution of Spiders and Silk Spinning: Mini Review of the Morphology, Evolution, and Development of Spiders' Spinnerets. *Front. Ecol. Evol.* **2020**, *8*, 109.

(53) Blackledge, T. A.; Scharff, N.; Coddington, J. A.; Szüts, T.; Wenzel, J. W.; Hayashi, C. Y.; Agnarsson, I. Reconstructing web evolution and spider diversification in the molecular era. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 5229–5234.

(54) Bond, J. E.; Opell, B. D. Testing adaptive radiation and key innovation hypotheses in spiders. *Evolution* **1998**, *52*, 403–414.

(55) Craig, C. L. *Spiderwebs and silks: tracing evolution from molecules to genes to phenotypes*; Oxford University Press: New York, 2003.

(56) Garb, J. E.; Ayoub, N. A.; Hayashi, C. Y. Untangling spider silk evolution with spidroin terminal domains. *BMC Evol. Biol.* **2010**, *10*, 243.

(57) Vienneau-Hathaway, J. M.; Brassfield, E. R.; Lane, A. K.; Collin, M. A.; Correa-Garhwal, S. M.; Clarke, T. H.; Schwager, E. E.; Garb, J. E.; Hayashi, C. Y.; Ayoub, N. A. Duplication and concerted evolution of MiSp-encoding genes underlie the material properties of minor ampullate silks of cobweb weaving spiders. *BMC Evol. Biol.* **2017**, *17*, 78–78.

(58) Starrett, J.; Garb, J. E.; Kuelbs, A.; Azubuike, U. O.; Hayashi, C. Y. Early Events in the Evolution of Spider Silk Genes. *PLoS One* **2012**, *7*, No. e38084.