Supplementary information for the manuscript:

The fractal self-assembly of the silk protein sericin

- Tejas S. Khire, Joydip Kundu, Subhas C. Kundu, Vamsi K. Yadavalli

Natural sericin can be recovered from the silk cocoon by various degumming processes such as boiling at 95 ^{o}C for 120 minutes [1], using water under pressure in an autoclave at 110-120 ^{o}C for different times [2], alkaline and soap degumming using 1.0 M NaOH solution, 0.02 M Na_2CO_3 or Marseille soap [3-6], acid degumming by boiling in citric, tartaric or succinic acid for 30 minutes [7] and enzymatic degumming, in which cut pieces of the cocoons are immersed in a buffer solution containing different amounts of enzyme, before deactivation of the enzymes at 85^{o} C for 10 minutes.[8] The size of the resulting sericin depends on the temperature, pH, and processing conditions. In this work, we extracted sericin from fresh cocoons by the process outlined below.

Sericin isolation

Sericin was obtained from the cocoons of two silk worms, *B. mori* and *A. mylitta,* following the protocols modified from Dash et al. [3] and Mandal et al. [9] respectively. In brief, the cocoons were cut into small pieces, degummed using alkali and heated in a 1:100 v/v solution of 0.02M sodium carbonate for about 30 minutes.[5] The water-soluble sericin is removed from cocoon pieces while the water-insoluble fibroin forms a mass of fibers. The sericin solution was then dialyzed several times to remove the salt and to concentrate it. Sericin was also obtained by boiling the cocoons in deionized water under pressure using an autoclave at 110-120 ^{o}C for 120

minutes [2]. The sericin solution was then concentrated against PEG 6000 using a 3.5 kDa dialysis membrane to the desired concentration. The molecular weight distribution of the sericin obtained using this method was verified using SDS PAGE. Sericin powder was obtained by lyophilizing the concentrated sericin solution, and it was stored at 4 ^{o}C until use.

Size exclusion chromatography on the sericin protein was performed using a high resolution Sephacryl S200 column. The crude sericin solution in DI water was separated into different size fractions using an AKTA Prime chromatograph (GE Healthcare). Fractions obtained after size exclusion were collected and further used for imaging.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis

8% sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) was performed according to Laemmli [10]. About 0.1-0.2 mg of the sericin samples that were isolated from the cocoons of *Bombyx mori* and *Antheraea mylitta* using alkali (0.02M Na_2CO_3 solution for 30 minutes) and hot boil degumming (autoclaved for 2 hours using steam sterilization conditions) were incubated with Laemmli sample loading buffer. The proteins were loaded onto a 5% stacking gel cast on the top of an 8% SDS polyacrylamide gel and electrophoresed. The gel was run at a constant voltage of 80V for approximately 3 hours so that the protein could be resolved efficiently. After electrophoresis, the gel was stained with Coomassie Brilliant Blue R-250 (Sigma) and destained in a methanol/water 1:1 solution that contained 20% acetic acid.

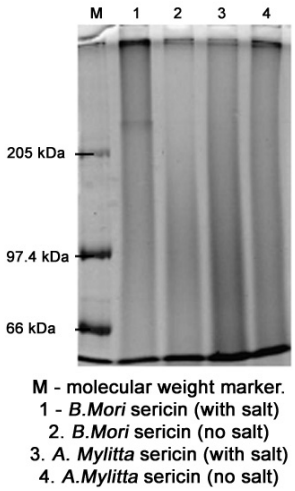

Figure S1: 8% SDS-PAGE analysis of sericin from cocoons of *A. mylitta* and *B. mori*. M- Molecular weight marker, 1- *B. mori* sericin (with salt) 2- *B. mori* sericin (no salt) 3- *A. mylitta* sericin (with salt), 4- *A. mylitta* sericin (no salt).

Dynamic Light Scattering Experiments:

Materials (polymers, proteins, colloids, or nanoparticles) scatter incoming laser light. Due to the random motion of these particles, the scattered light intensity fluctuates in time. Processing the fluctuating signal with autocorrelation yields the particle's diffusion coefficient, from which the equivalent spherical particle size is calculated using the Stokes-Einstein equation. Molecular weights can also be calculated from these fluctuations or from the scattered intensity. To estimate the size distribution and polydispersity of the sericin, dynamic light scattering experiments were performed on the protein samples from *B. mori* and *A. mylitta*.

Experiments were conducted on a 90 Plus Particle size analyzer, (Brookhaven Instruments Corporation, NY). A sample volume of 1-3 mL (concentration ~ 0.1 mg/ml) was used. Measurements were taken at 90° at 25^{o}C. The distribution is presented in Figure S2.

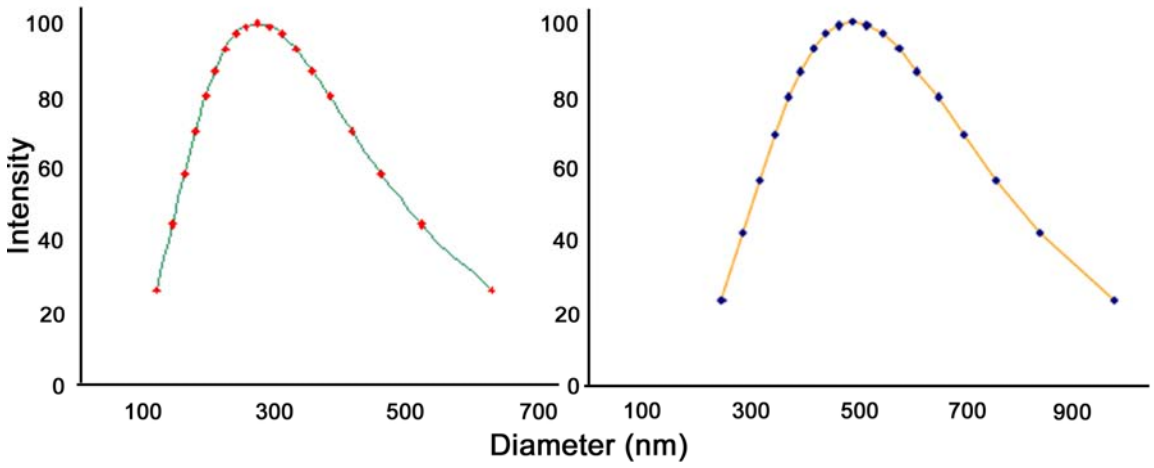

Figure S2: Dynamic light scattering experiments showing the size distributions of the two varieties of sericin (L) – *A. mylitta* (R) – *B. mori* studied.

Table 1: Analysis of the Dynamic light scattering data is presented below

Sample	Effective Diameter (nm)	Average diameter (nm)	Polydispersity	Half Width (nm)
B. mori	268	278 ± 6	0.3	148
A. mylitta	493	421 ± 83	0.2	216

The average diameter of the *B. mori* sericin protein dispersions in water were smaller compared to the *A. mylitta* sericin proteins in water. These observations are in good agreement with the height images obtained via AFM imaging. The *B. mori* sericin protein structures as observed under AFM were comparatively smaller than the *A. mylitta* sericin protein structures (page 8).

Results and Discussion

Under non-reducing conditions, the *Bombyx mori* sericin isolated by alkali (Na_2CO_3 salt solution) yielded a fraction larger than 200 kDa, whereas the sericin isolated by hot boil degumming showed no distinct fraction above 200 kDa. The *B. mori* sericin isolated by hot boil degumming showed a continuous banding pattern in the region between 30-100 kDa that resulted in the formation of a smear on the gel. Alkali and hot boil degumming of *A. mylitta* sericin, however, showed it to possess a continuous distribution of bands between 97 kDa and 29 kDa (and some bands that were >97 kDa and <29 kDa). Sericin obtained by degumming in hot water produces fractions of sericin polypeptides separating; causing a smear to appear in the SDS PAGE experiments. Taking these SDS–PAGE results together suggests that sericin represents a family of proteins with a diverse distribution of molecular weights. These results are very similar to those reported by Zhang et al. 2004 [2] and Wu et al., 2007 [11]. The sericin is imaged from the consolidation of these different polypeptides. Suitable control experiments were carried out to verify the presence of impurity during extraction and imaging process and it appears that the self assembled structures formed are solely due to the silk sericin protein.

Sericin drying experiments

A few drops of the sericin solutions (10-20 μl) were deposited on a freshly cleaved mica surface and allowed to dry under constant humidity. We observed that speeding up the regular drying process by heating the substrate or by forced convective drying using a stream of N_2 resulted in the formation of aggregates and clusters, indicating that the aggregation and self-assembly process was time-dependent.

The effect of the nature of the surface

The underlying substrate is an important factor that may direct the nature of assembly of sericin. In these experiments, we initially used silicon as the surface in which the sericin solutions were dried and subsequently imaged. However, due to the hydrophobicity of silicon (despite treatment with piranha solution), the solutions tended to dry unevenly, resulting in aggregated clusters and the formation of a film in many cases (Figure S3). We therefore used mica as the substrate of choice in subsequent sericin drying experiments. The hydrophilic and atomically flat nature of freshly cleaved mica permitted us to spread the solution and to dry it without concentration gradients, yielding high-resolution AFM images.

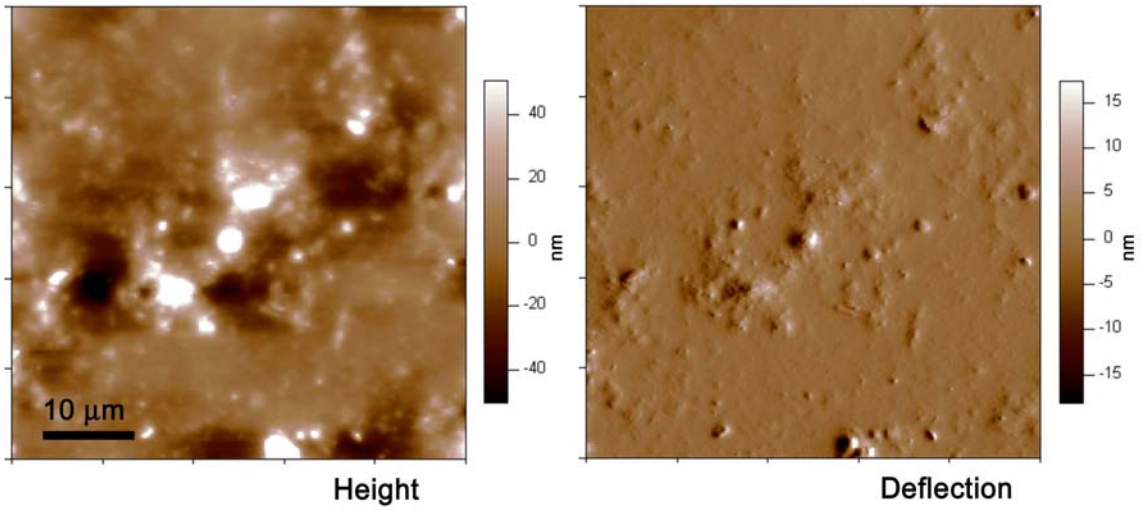

Figure S3: AFM scan of uneven sericin film formed on a silicon surface.

Finally, we verified that the lyophilization process did not change the behavior of sericin. We prepared liquid sericin freshly, by degumming from the cocoons and dialysis, and used this solution without any subsequent processing. After dilution, the solution was dried and imaged using AFM as described above. The images revealed similar branched patterns, indicating that the lyophilization process was not contributing the assembly process that we observed (Figure

S4). Figure S5 shows some other variants of the two categories of sericin images described in the manuscript.

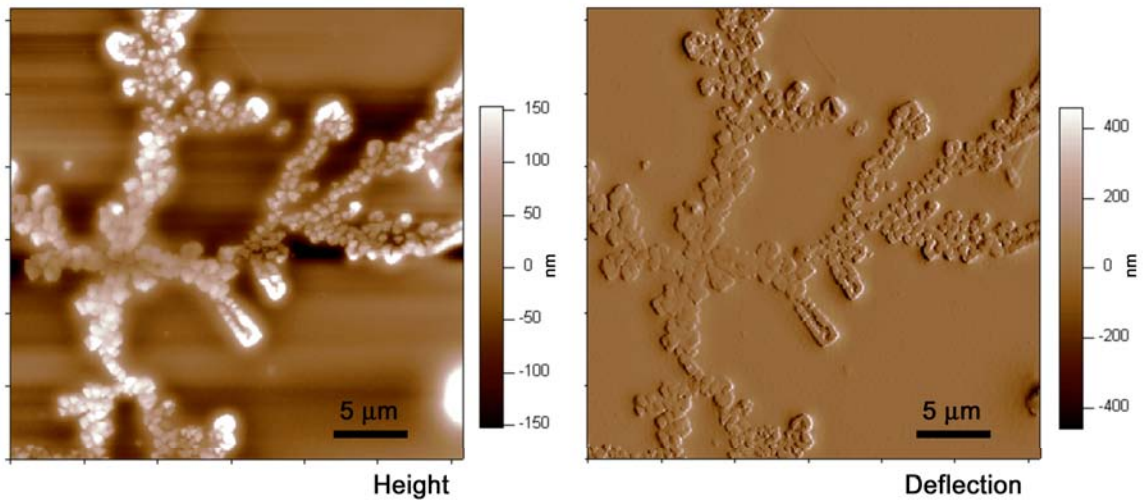

Figure S4: An AFM scan of a sericin sample derived from *B. mori* (0.1 mg/ml), before lyophilization.

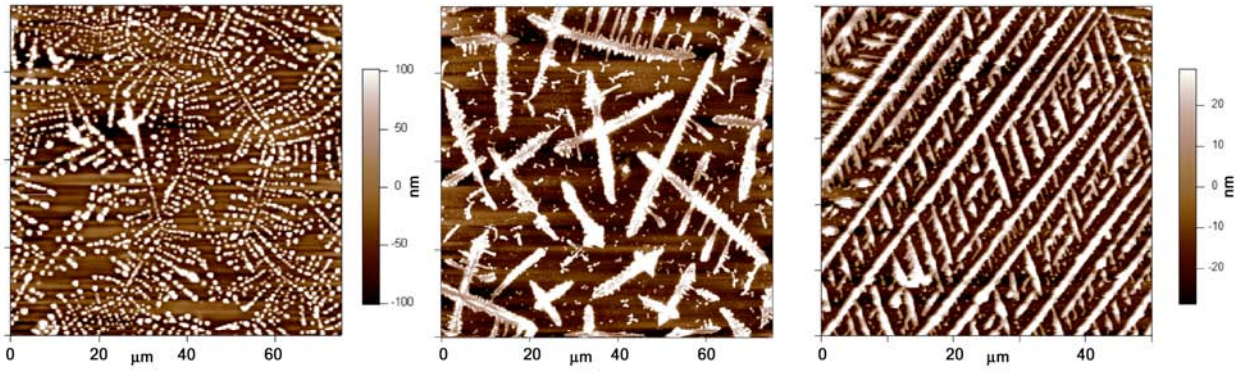

Figure S5: Variants of the fractal structures observed for sericin. The AFM images shown depict *B. mori* sericin (Right) and *A. mylitta* sericin (Left) and (Middle)

References:

1. N. Kato, S. Sato, A. Yamanaka, H. Yamada, N. Fuwa and M. Nomura, *Biosci. Biotechnol. Biochem.*, 1998, **62**, 145-147.
2. Y. Q. Zhang, M. L. Tao, W. D. Shen, Y. Z. Zhou, Y. Ding, Y. Ma and W. L. Zhou, *Biomaterials*, 2004, **25**, 3751-3759.
3. R. Dash, M. Mandal, S. K. Ghosh and S. C. Kundu, *Mol. Cell. Biochem.*, 2008, **311**, 111-119.
4. S. C. Kundu, B. C. Dash, R. Dash and D. L. Kaplan, *Prog. Polym. Sci.*, 2008, **33**, 998-1012.
5. H. Yamada, H. Nakao, Y. Takasu and K. Tsubouchi, *Mater. Sci. Eng., C*, 2001, **14**, 41-46.
6. Y. Q. Zhang, Y. Ma, Y. Y. Xia, W. D. Shen, J. P. Mao and R. Y. Xue, *J. Controlled Release*, 2006, **115**, 307-315.
7. A. Kurioka, F. Kurioka and M. Yamazaki, *Biosci. Biotechnol. Biochem.*, 2004, **68**, 774-780.
8. G. Freddi, R. Mossotti and R. Innocenti, *J. Biotechnol.*, 2003, **106**, 101-112.
9. B. B. Mandal and S. C. Kundu, *Acta Biomater.*, 2009, **5**, 2579-2590.
10. U. K. Laemmli, *Nature*, 1970, **227**, 680-685.
11. J. H. Wu, Z. Wang and S. Y. Xu, *Food Chem.*, 2007, **103**, 1255-1262.