# Acid-induced assembly of a reconstituted silk protein system

A. Pasha Tabatabai,[1] Katie M. Weigandt,[2] and Daniel L. Blair[1,*]

[1]*Georgetown University Department of Physics and Institute for Soft Matter Synthesis and Metrology, Washington DC 20057, USA*

[2]*Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*

(Received 19 January 2017; published 10 August 2017)

Silk cocoons are reconstituted into an aqueous suspension, and protein stability is investigated by comparing the protein's response to hydrochloric acid and sodium chloride. Aggregation occurs for systems mixed with hydrochloric acid, while sodium chloride over the same range of concentrations does not cause aggregation. We measure the structures present on the protein and aggregate length scales in these solutions using both optical and small-angle neutron scattering, while mass spectrometry techniques shed light on a possible mechanism for aggregate formation. We find that the introduction of acid modulates the aggregate size and pervaded volume of the protein, an effect that is not observed with salt.

DOI: 10.1103/PhysRevE.96.022405

## I. INTRODUCTION

Silk protein materials are the focus of many current biomedical applications because they are strong and biocompatible [1,2]. Nonfibrous silk forms can now be constructed by taking advantage of protein reconstitution techniques where prespun cocoon fibers are chemically dissolved into solution in large quantities [3]. However, the protein network morphologies of reconstituted silk-based materials are different than that of the prespun fiber. Consequently, the optimization of new silk materials requires a deeper understanding of the structures of these new systems.

A novel class of gel-like materials are formed by adjusting the pH of reconstituted silk solutions. Acidic domains are formed either through the incorporation of acid (pH-gel) or by electrolysis when running a DC electric current through the solution (e-gel) [3–9]. In both e-gels and pH-gels, proteins assemble while maintaining the random coil or $\alpha$-helical protein confirmation of the reconstituted solution [4–7]. In prespun fibers, the fiber strength is attributed to the ordering of $\beta$-sheets; however, interesting bulk rheological properties are seen in the more amorphous e-gels [10–12].

The mechanism by which non-$\beta$-sheet silk assembles into materials is not well understood, and progress is needed to improve reconstituted silk material design. Current microscopic insight in pH-induced assembly is primarily found through lyophilized state electron microscopy, but protein scale information, particularly in solution, is lacking [5,6,13]. One measurement technique that can resolve protein and subprotein length scales is small-angle neutron scattering (SANS), but SANS has not been utilized extensively in the silk community [14–16].

In this work, we focus on the structural changes of reconstituted silk fibroin protein using SANS when assembled into macroscopic structure. We incorporate hydrochloric acid (HCl) and work at low silk concentrations that form aggregates instead of gels as a first step toward better understanding silk protein organization into larger-scale structures. The mechanism of aggregation is investigated by determining the structure of the protein through measurement of the fractal dimension, and comparisons are made to silk solutions containing equivalent concentrations of sodium chloride (NaCl) where no aggregation is observed.

## II. METHODS AND MATERIALS

### A. Reconstituted protein

Silk protein solutions are reconstituted from *Bombyx mori* silk cocoons following the method detailed by Rockwood *et al.* [3]. Cocoons are boiled in aqueous sodium carbonate for 10 min to remove the globular sericin, leaving behind only the insoluble structural fibroin proteins that are washed in deionized water. The fibers are soaked in an aqueous solution of lithium bromide (LiBr) at 70°C for 2 h to solubilize the protein through denaturation and to cleave the disulfide bond connecting the heavy and light chain [17]. The LiBr-protein solution is dialyzed against cycled deionized water for 48 h in a dialysis bag with a molecular weight cutoff of 10 kDa to remove the solubilized ions. Undissolved fibers are removed by centrifugation, and protein aggregates larger than 0.45 $\mu$m are removed through filtration. The resultant reconstituted silk solution is composed entirely of the heavy chain fibroin protein and will now be referred to as silk protein [17]. For neutron scattering measurements, aqueous silk solutions are subsequently dialyzed against deuterium oxide ($D_2O$) to an $H_2O$:$D_2O$ solvent volume ratio of 5:95.

Silk is stable in an unbuffered solution for weeks to months. Gel electrophoresis shows that silk proteins after reconstitution are polydisperse but have a well-characterized molecular weight distribution that ranges from 100 → 700 kDa [17]. Subnative-sized protein fragments (< 390 kDa) exist due to degradation, whereas sizes greater than the native molecular weight (> 390 kDa) are attributed to unavoidable aggregation [17]. From the molecular weight distribution, a conservative estimate for the overlap concentration $c^* = 40$ mg ml$^{-1}$ is determined by approximating the Kuhn length and assuming $\theta$-solvent conditions [17]. Using the molecular weight for the native silk of 390 kDa, the overlap concentration is equivalent to $c^* = 100$ $\mu$M; concentrations of protein $c_P$ for structural measurements are maintained in the dilute regime $c_P < c^*$.

Despite the protein polydispersity, the molar extinction coefficient at 280 nm $\varepsilon_{280} = 441\,030$ cm$^{-1}$M$^{-1}$ is used given

*daniel.blair@georgetown.edu

 

the amino acid sequence of the native protein [18]. Ultraviolet-visible light (UV-Vis) spectroscopy measures the absorbance of light at 280 nm from proteins of the entire molecular weight distribution, and the use of $\varepsilon_{280}$ for the native undegraded protein converts the absorbance into an effective undegraded protein concentration.

### B. Mixing silk with HCl or NaCl

NaCl and HCl are chosen because they are common, ionize completely at our experimental concentrations, and share a common anion. HCl and NaCl solutions of concentration $c_H$ and $c_N$, respectively, are prepared by dilution in either $H_2O$ or $D_2O$. For neutron scattering, HCl stock solutions are prepared by dilution of a 37% HCl in $H_2O$ assay into $D_2O$. At $c_H < 5$ mM, solutions of HCl in $D_2O$ have a scattering length density equivalent to 100% $D_2O$; further dilutions of this stock HCl solution do not appreciably change the scattering length density. Acid or salt solution is added to the $D_2O$ protein solution at a 1:1 volume ratio resulting in an $H_2O$:$D_2O$ solvent volume ratio of 2.5:97.5.

The exact stoichiometric interaction between HCl or NaCl and silk protein is unknown so we define the molar equivalent (ME) as the ratio of the moles of added compound to the effective number of moles of native silk in solution determined with UV-Vis and $\varepsilon_{280}$. Solutions are prepared in deionized water with background NaCl present in trace amounts and at pH 9 ([HCl] $= 10^{-9}$ M). Since $c_H, c_N \geqslant 0.1$ mM and are much larger than their respective background concentrations, ME is equivalent to the absolute ion concentration per protein.

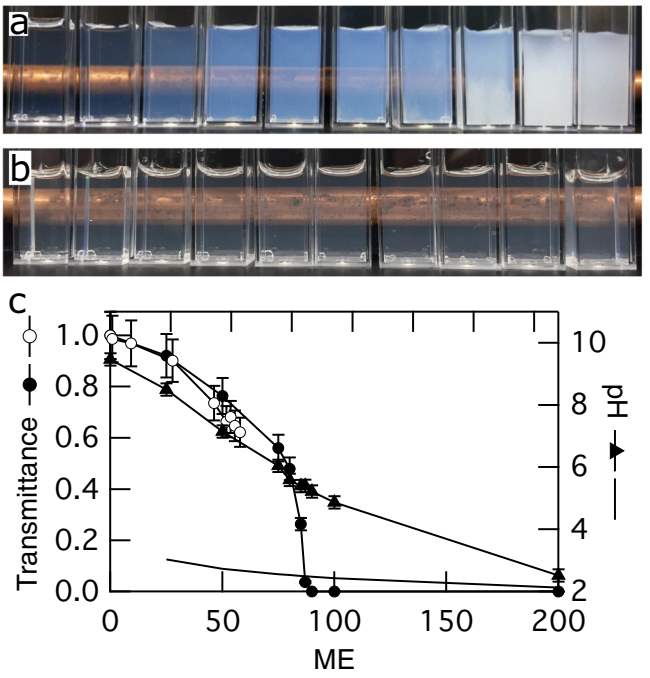


FIG. 1. 37 $\mu$M silk protein solutions with (a) $c_H = 0 \rightarrow 7.4$ mM (left to right) ($0 \leqslant$ ME $\leqslant 200$) and (b) NaCl over the same range $0 \leqslant$ ME $\leqslant 200$. (c) Data corresponding to samples from (a): the normalized transmittance for each sample at 488 nm (●), the measured pH (▲), and the expected pH value (−). Transmittance for the samples used in the neutron scattering experiment (○).

### C. Measuring ion concentrations

The concentration of lithium (Li) and bromine (Br) ions are measured with inductively coupled plasma mass spectrometry (ICP-MS) for the isotopes Li-6, Li-7, Br-79, and Br-81; reported values are a sum of these isotopes. The count rates from an ICP-MS measurement are converted into a concentration by measurements of known LiBr concentrations; the conversion between count rates and concentration is confirmed to be linear in the measured range and least squares fitting gives the conversion factor with 95% confidence. Data plotted are the best estimates, and error bars correspond to the propagation of uncertainties in the conversion factors.

Using centrifugation filters with a molecular weight cutoff of 10 kDa, elutions of silk solutions containing HCl or NaCl are collected. The protein and ion concentrations in the elutions are measured with UV-Vis spectroscopy and ICP-MS, respectively. Changes in elution ion concentrations are representative of changes in mobile ion concentrations in the bulk.

## III. RESULTS AND DISCUSSION

### A. Aggregate formation

The turbidity of silk solutions increase as $c_H$ increases from $0 \rightarrow 7.4$ mM [Fig. 1(a)]. The protein concentration $c_P = 37$ $\mu$M gives $0 \leqslant$ ME $\leqslant 200$. Transmittance $T$ decreases rapidly from $T = 60\%$ to $T = 0\%$ in the range of $75 \leqslant$ ME $\leqslant 90$ and remains at zero with any additional acid [Fig. 1(c)]. In contrast, NaCl has no visible effect on the sample turbidity at $0 \leqslant$ ME $\leqslant 200$ [Fig. 1(b)]. The difference in total optical scattering shows that HCl and NaCl affect reconstituted silk differently. Interprotein aggregation is a consequence of HCl-induced association.

The pH of each of the samples in Fig. 1(a) are measured directly with a pH probe and decrease smoothly from $9 \rightarrow 2$ with increasing ME [Fig. 1(c)]. Since ~1% of the amino acids on the protein are ionizable, the approximation for the pH of the solution pH $= -\log(c_H)$ predicts a decrease in pH from $3 \rightarrow 2$. However, the measured pH differs significantly from the predicted value due to the strong association of $H^+$ with the protein. When ME $= 200$, the measured pH and the expected pH value begin to converge; the protein is saturated with $H^+$ and unable to continue to buffer the solution. Consequently, any additional $H^+$ (ME $> 200$) will remain in solution. Because the measured pH does not decrease stepwise as ionization occurs, we suspect that HCl-induced aggregation is not due to protonation of amino acids nor due to a minimization of charge repulsion [8,9,13,19].

Changes in $c_H$ result in changes in $T$. Likewise, changes in $c_H$ alter the intensities of SANS spectra on both the aggregate and protein length scales [Fig. 2]. An increase in $c_H$ results in an increase in the scattering intensity $I$ for wave vector $q$ in the low-$q$ ($q < 7 \times 10^{-3}$ Å$^{-1}$) regime. The low-$q$ increase signifies an increase in the total aggregate volume that is consistent with turbidity measurements [Fig. 1]. High-$q$ ($q > 4 \times 10^{-2}$Å$^{-1}$) scattering is independent of $c_H$, while a

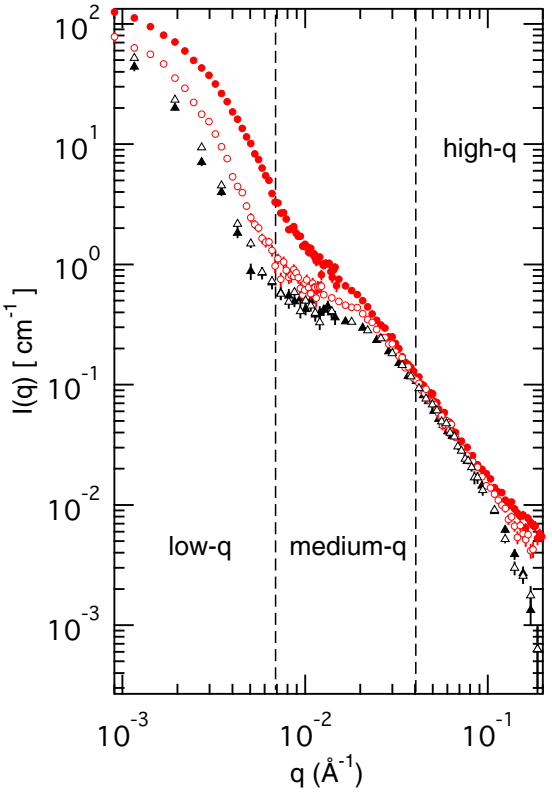


FIG. 2. Scattering intensity $I(q)$ of deuterated silk samples at $c_P = 37\ \mu$M: 0.5 mM NaCl (△), 1.0 mM NaCl (▲), 0.5 mM HCl (○), and 1.0 mM HCl (●).

slope change between high-$q$ and medium-$q$ ($7 \times 10^{-3}$ Å$^{-1}$ $< q < 4 \times 10^{-2}$ Å$^{-1}$) provides the characteristic length scale for a single protein.

In contrast, $I(q)$ are independent of changes in $c_N$; unchanged low-$q$ scattering is consistent with the lack of turbidity in Fig. 1(b). Instead of inducing aggregation, charge screening effects from NaCl appear at $q > 0.06$ Å$^{-1}$ as observed by a slight increase in the slope of the NaCl curves in Fig. 2. The presence of $q$-dependent scattering at low-$q$ for these spectra is a consequence of unavoidable aggregate byproducts from reconstitution. However, *changes* in aggregates are only induced by HCl.

Because $c_H$ and $c_N$ span the same range and aggregates form only from increases in $c_H$, aggregation is only induced by HCl. Aggregation is not a consequence of ionic strength or charge screening; incorporation of HCl is needed before interprotein associations can exist.

### B. Reconstituted protein stability

In spun fibers, silk is folded into multiple $\beta$-sheets through hydrogen bonding: a dipole-dipole interaction between polar amine N-H and carbonyl C=O groups on the protein backbone. However, reconstituted silk is predominantly an unstructured random coil where N-H and C=O are unassociated [8,17]. We find that the random coil configuration is likely stable because of residual lithium ions from the reconstitution process that associate with C=O analogous to a charge-dipole interaction. A comparison of lithium concentrations before and after

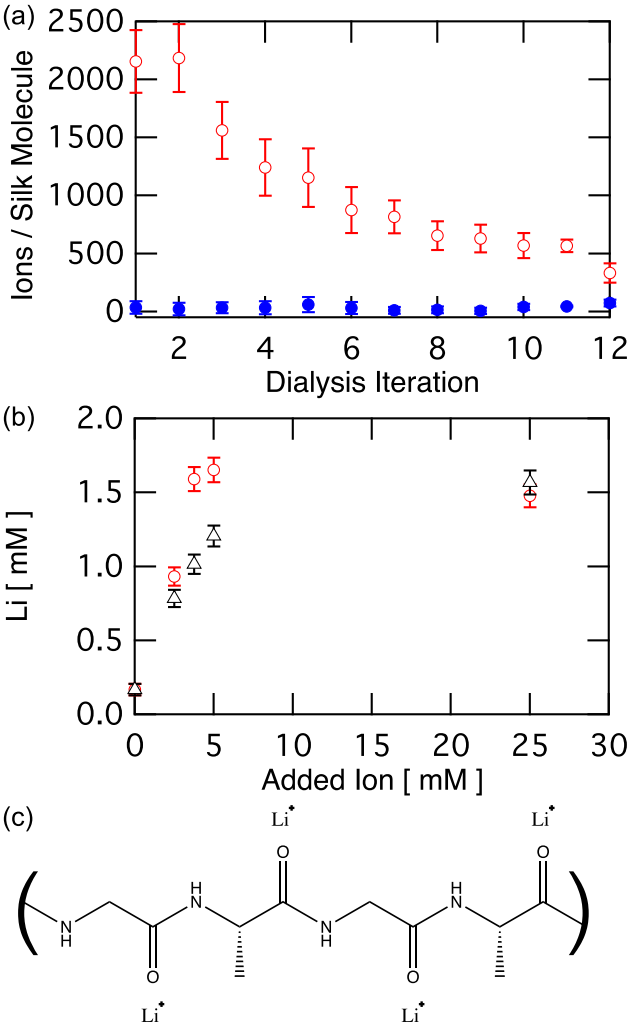


FIG. 3. (a) Lithium (○) and bromine (●) ions per protein during twelve extra dialysis iterations beyond the normal reconstitution protocol. Iteration number one represents the ion concentrations present after the standard reconstitution protocol. (b) Lithium ions in the elution with the addition of either HCl (○) or NaCl (△). (c) A common segment of the protein sequence (G-A-G-A) with lithium ions drawn to associate with carbonyl oxygens.

HCl or NaCl addition indicates a mechanism for aggregate formation.

After the reconstitution of silk at $c_P = 100\ \mu$M, the number of $Li^+$ per protein is measured to be $2200 \pm 300$ [Fig. 3(a)]; reconstituting silk at $c_P = 1\ \mu$M has no significant effect on this ratio. The native silk protein has 5263 amino acids, therefore there is one $Li^+$ for every $2.4 \pm 0.3$ amino acids. Given that the native protein has a molecular weight of 390 kDa, the average mass of an amino acid is 74 g. Hence, our results suggest that a dehydrated silk film has mass ratio of $3.9\% \pm 0.5\%$ lithium. It is unclear why a previous experiment measured no $Li^+$ content in a silk film [7].

The ratio of $Li^+$ per amino acid is equivalent to one $Li^+$ for every $2.4 \pm 0.3$ carbonyl oxygens and is consistent with molecular dynamics simulations and density functional theory calculations for $Li^+$ associations with electronegative oxygens on different polymer chemistries [20–22]. Therefore, $Li^+$ are

present at the right stoichiometric ratio to be associated to the carbonyl oxygens on the protein backbone. $Li^+$ has previously been found through experiments and computation to have a high binding affinity with the carbonyl oxygen on the protein mimetic molecule N-methyl-acetamide as well as with the amino acids [22,23]. Associations of $Li^+$ to the carbonyl is reasonable since chaotropic LiBr is chosen in the reconstitution process to denature the protein, i.e., disassociate N-H and C=O groups. Silk's stability and inability to self-assemble is likely caused by the continued complexation of the carbonyl with $Li^+$ [24]. Meanwhile, the $Br^-$ concentration is very low; presumably all $Br^-$ are removed during dialysis. A counterion for $Li^+$ likely exists to maintain electroneutrality, and further studies are needed to determine the chemical composition of the counterion.

To confirm that $Li^+$ are associated to the protein, the concentration of $Li^+$ per protein is measured during additional dialysis iterations. The silk solution $V_{silk} = 15$ ml is placed back into a 10 kDa dialysis cassette and set in a reservoir with volume $V_{res} = 3000$ ml of gently stirred deionized water. After 24 h, an aliquot of the silk solution is collected, and the reservoir water is replaced with new deionized water. After 12 iterations, the number of $Li^+$ per protein decreases approximately by a factor of 10, whereas $Br^-$ concentrations remain at zero [Fig. 3(a)].

The concentration of $Li^+$ inside the cassette ($c_1$) is greater than the concentration in the reservoir initially. After 24 h, the new equilibrium concentration of $Li^+$ in the cassette $c_2 = c_1 \times V_{silk}/V_{res}$. Following 12 dialysis iterations, the recursive relation gives $c_{12} = c_1 \times (V_{silk}/V_{res})^{11}$ or $c_{12}/c_1 \sim (1/200)^{11}$. The expected concentration ratio $c_{12}/c_1$ is significantly different than the measured ratio $c_{12}/c_1 \sim 1/10$. Therefore, it must be the case that $Li^+$ has an affinity to the silk protein.

We vary $0\ \text{mM} \leqslant c_H, c_N \leqslant 25\ \text{mM}$ in silk solutions prepared with the standard reconstitution protocol to determine if $H^+$ and $Na^+$ ions displace $Li^+$ from the protein. As $c_H$ or $c_N$ increases, the concentration of free $Li^+$ increases [Fig. 3(b)]. HCl is better at displacing $Li^+$ than NaCl at low concentrations, but there is no effective difference at 25 mM. Because both HCl and NaCl remove $Li^+$ from the protein, we know that both compounds interact with the protein. $H^+$ associated to the protein is a likely mechanism for the discrepancy between the measured pH value and the expected value shown in Fig. 1.

Since $Li^+$ is known to associate with carbonyls, as previously mentioned, it is likely that both $H^+$ and $Na^+$ also associate with the carbonyls to conserve charge. The displacement of $Li^+$ with either $H^+$ or $Na^+$ is reasonable given that density functional theory calculations provide roughly equivalent affinities of $H^+$, $Na^+$, and $Li^+$ for carbonyl oxygens [22,25,26]. It is not only the removal of $Li^+$ that causes aggregation but specifically the removal of $Li^+$ by $H^+$ that is needed to destabilize the protein and induce aggregation.

### C. Acid-induced structures

Our results establish a link between aggregate formation and acid. HCl-induced protein aggregates are now measured to quantify HCl-dependent changes in structure. Silk protein solutions at a final concentration of $c_P = 54\ \mu$M protein in $D_2O$ are mixed with HCl in the range of $0\ \text{mM} \leqslant c_H \leqslant 3.125$ mM, giving $0 \leqslant \text{ME} \leqslant 58$. The transmittance of these samples is a function of ME and ranges from 100% → 50% [Fig. 1].

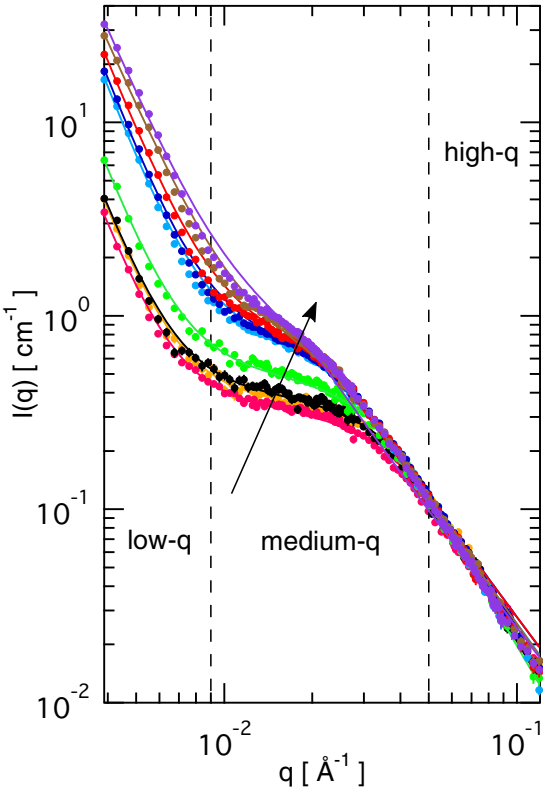


FIG. 4. $I(q)$ changes monotonically as ME increases along the arrow from $0 \rightarrow 58$ except for the two lowest values of ME: ME = 0 (●) and ME = 1.9 (●). $I(q)$ for each value of ME are fit with a double Guinier-Porod model and are plotted as the solid lines.

Changes in the scattering intensity $I(q)$ reveal ME dependent structural changes [Fig. 4]. At high-$q$ ($q > 4 \times 10^{-2}$ Å$^{-1}$), $I(q)$ is unchanged over the range of ME; the protein has a constant fractal dimension irrespective of the aggregate conditions. Medium-$q$ scattering ($9 \times 10^{-3}$ Å$^{-1} < q < 4 \times 10^{-2}$ Å$^{-1}$) exhibits an ME-dependent slope change that likely exists due to the superposition of scattering from the aggregate and protein length scales. Although byproducts of reconstitution are seen in the $q$ dependence at low-$q$ ($q < 9 \times 10^{-3}$ Å$^{-1}$) for ME = 0, increases in ME manifest themselves as a vertical shift in $I(q)$; more aggregates are formed with increasing ME as seen in Fig. 1. ME does not affect the low-$q$ slope; the internal structure of an aggregate remains constant.

$I(q)$ in Fig. 4 are a result of scattering from both the individual proteins and aggregates. To minimize assumptions during data fitting, each length scale is interpreted using the Guinier-Porod model in three dimensions,

$$I(q) = \begin{cases} G \cdot \exp\left(\frac{-q^2 R_G^2}{3}\right), & \text{if } q \leqslant q_1, \\ \frac{D}{q^d}, & \text{if } q \geqslant q_1, \end{cases} \quad (1)$$

where G is the Guinier scaling parameter, $D$ the Porod scaling factor, $d$ the Porod exponent, and characteristic length $q_1$

[27]. Both of these equations and their derivatives must be continuous at $q = q_1$; therefore, the following relations must hold:

$$q_1 = \frac{1}{R_G}\left(\frac{3d}{2}\right)^{\frac{1}{2}},$$

$$D = G\exp\left(-\frac{d}{2}\right)\left(\frac{3d}{2}\right)^{\frac{d}{2}}\frac{1}{R_G^d}. \quad (2)$$

The Guinier regime of the protein and the Porod regime of the aggregate are not well separated, so both the aggregate and the protein are fit simultaneously using the same functional form, giving

$$I(q)_{\text{total}} = I(q)_{\text{protein}} + I(q)_{\text{aggregate}}. \quad (3)$$

The use of the double Guinier-Porod model is necessary to interpret the protein $R_G$. Any artifacts of this fitting method may arise in the absolute value of $R_G$, but the qualitative changes in $R_G$ with $c_H$ will be captured.

This double Guinier-Porod fit has six free variables: $R_G$, $G$, and $d$ for both the protein and the aggregate. To minimize the number of free fitting variables, dynamic light scattering (DLS) is used to identify the hydrodynamic radii $R_H$ of the aggregates, which are used as the aggregate $R_G$, and to confirm that the system is ergodic. The reported values of $R_H$ are the ensemble average $\langle R_H \rangle$ and decrease from 145 → 95 nm with increasing ME [Fig. 5(a)]. Since the increase in turbidity coincides with a decrease in aggregate size, the number of aggregates must increase with ME. It is likely that the original ME = 0 aggregates persist, and the number of smaller aggregates increases with ME.

Additionally, $d_{\text{protein}}$ is determined by fitting only over the protein lengthscale ($q > 1.5 \times 10^{-2}$ Å$^{-1}$) using Eq. (1) and is independent of ME [Fig. 5]. Taking the distributions of best estimates, we find $d_{\text{protein}} = 2.12 \pm 0.02$; over this range of ME, the fractal dimension of the protein is equivalent to a polymer on the bad side of $\theta$-solvent conditions. The measured fractal dimension near $\theta$-solvent conditions is consistent with previous measurements of random coil structure in reconstituted silk [7,17,28].

Each scattering curve in Fig. 4 is fit using Eqs. (1)–(3) and the fitting results are plotted as solid lines. Double Guinier-Porod fitting provides the best estimates and uncertainty for the Porod exponent of the aggregate $d_{\text{aggregate}}$ and $R_G$ of the protein. All values extracted from fits represent an ensemble average $\langle ... \rangle$; therefore, the quoted value of $R_G = \langle R_G \rangle$ takes into account the protein polydispersity. As ME increases, $R_G$ for the protein increases from 40 → 70 Å [Fig. 5(b)]. The value of $R_G$ for ME = 0 is consistent with the only other known neutron study of reconstituted silk [16]. Measured values of $R_G$ are smaller than previous predictions by Partlow *et al.*, because the SANS results are a weighted average over the polydisperse protein and it is possible that their approximation of the Kuhn length was too large [17]. $d_{\text{aggregate}} \sim 4$ for most of the values of ME but varies slightly at high values of ME when the fitting quality in Fig. 4 starts to decline [Fig. 5(a)]; the aggregates amongst all values of ME are similarly structured and dense.

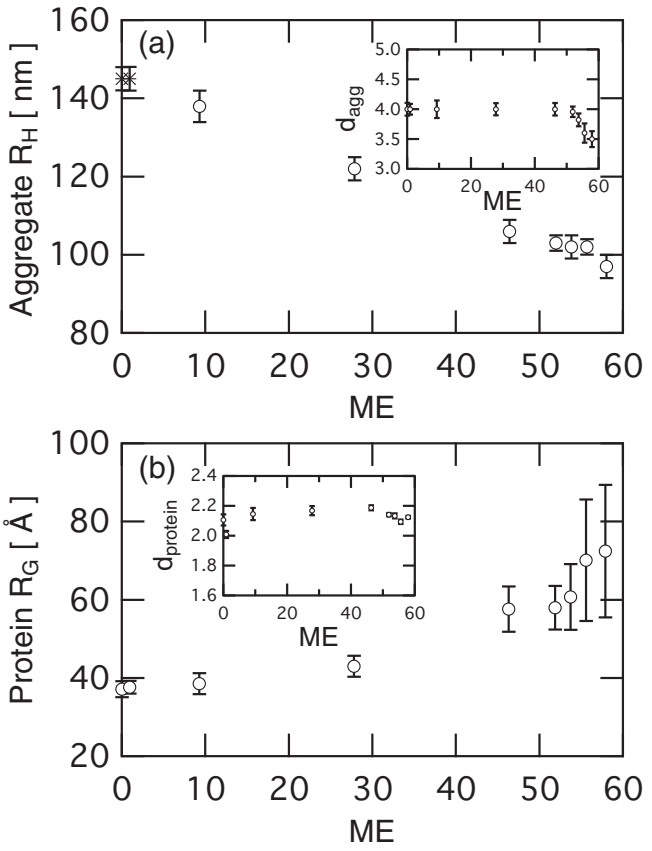


FIG. 5. (a) Aggregate $R_H$ as determined from DLS as a function of ME. The measured samples for ME > 9 (○) are used to approximate the hydrodynamic radii of the two unmeasured samples (∗) by linear interpolation, and the error bars correspond to the propagation of uncertainty in the interpolation. (a-inset) $d_{\text{aggregate}}$ determined by the double Guinier-Porod fit of I($q$). (b) Protein $R_G$ as determined through the double Guinier-Porod fit. (b-inset) $d_{\text{protein}}$ determined by the single Guinier-Porod fit.

The measured values of $R_H$ are close to the experimentally accessible length scale of SANS. The beginning of a turnover at low-$q$ is seen in Fig. 2 for samples with an extended $q$ range, as expected from the DLS data, but the complete low-$q$ plateau is still out of the range of SANS. The low-$q$ plateau is accessible with ultrasmall angle neutron scattering (USANS). However, combining SANS and DLS minimizes stability and aging effects of the protein that could be observed during a long USANS experiment.

## IV. CONCLUSIONS

Silk protein in reconstituted solutions are stabilized by $Li^+$ associated to a fraction of carbonyl oxygens. The addition of either HCl or NaCl displaces bound $Li^+$, but multiprotein structures are capable of forming only in the presence of HCl. The removal of $Li^+$ from the C=O dipole suggests that aggregation involves the reassociation of C=O and N-H as is seen in the silk fiber [29]. However, the replacement of $Li^+$ with $H^+$ leads us to suggest an alternative bond: $H^+$ facilitates a bond between two separate C=O for an effective dipole-charge-dipole interaction. This complex is

called a bifurcated hydrogen bond and is observed in biological systems [30]. It will be important to distinguish between hydrogen and bifurcated hydrogen bonding because the two types of interactions have different binding strengths [30].

While bonds cannot be seen explicitly, changes in the structure of the protein are indirect evidence of changing protein-protein interactions. We show that ME increases the protein $R_G$ in an aggregate but not the protein structure defined by $d_{\text{protein}}$. At first glance, these results appear to violate mass conservation, since the size of a particle can not increase without changing its density. However, these results are possible if either (1) the boundary between one protein and its aggregated partner is undefined because of intercalation, or (2) the proteins are extending. Because $d_{\text{aggregate}} \sim 4$, the most likely scenario is that two proteins intercalate, allowing the formation of a dense aggregate. Importantly, we have shown that reconstituted proteins do not form ordered structures in an aggregate, and that aggregates exhibit a characteristic size that depends on ME.

We have shown that SANS is a useful technique for measuring the protein scale structure of silk aggregates in solution. Extension of this work to higher concentration protein solutions ($c_P > c^*$), where networks form instead of aggregates, is necessary so that both SANS and rheology can be performed; our results provide an important step towards the correlation of silk material microstructure with mechanical properties. Additionally, rheological measurements of the silk network may be able to test our hypothesis regarding the bonding mechanism.

## ACKNOWLEDGMENTS

This work was supported by the Air Force Office of Scientific Research through Grant No. FA9550-07-1-0130. A.P.T. thanks the Walter G. Mayer Endowed Scholarship Fund for support. This work utilized facilities supported in part by the National Science Foundation under Agreement No. DMR-1508249. This work benefited from the use of the SasView application, originally developed under NSF Award No. DMR-0520547. SasView contains code developed with funding from the European Union's Horizon 2020 research and innovation program under the SINE2020 project, Grant Agreement No. 654000.

## APPENDIX A: SMALL-ANGLE NEUTRON SCATTERING

Neutron scattering experiments are performed on the 30-m SANS instruments at the National Institute of Standards and Technology Center for Neutron Research [31–35]. Measurements of the scattering intensity $I$ as a function of wave vector $q$ for the range $10^{-3}$ Å$^{-1}$ $< q <$ 0.5 Å$^{-1}$ are accomplished by using both 6 Å wavelength neutrons at detector positions of 1 m, 4 m, and 13 m and 8.09 Å lens focused neutrons at a detector position of 15.3 m. Deuterated silk solutions are pipetted into rectangular quartz cuvettes and mixed rapidly with an equal volume of HCl or NaCl solution. To ensure we measure the steady-state structure, samples equilibrate for 12 h before the measurement at ambient conditions.

Data are reduced using the Igor macros, and the $q$-independent background scattering intensity for each sample is averaged and subtracted from each curve before plotting so that data represent scattering from only the protein [36]. Curve fitting is performed using SasView and custom code where a nonlinear least squares regression is fit iteratively for the best estimate of each fitting parameter as well as the error associated with 95% confidence for each value [37].

## APPENDIX B: DYNAMIC LIGHT SCATTERING

The hydrodynamic radius $R_H$ of an aggregate is determined by dynamic light scattering on the same silk solutions from SANS over the range $5 \times 10^{-4}$ Å$^{-1}$ $< q < 25 \times 10^{-4}$ Å$^{-1}$ and have minima at $q = 20 \times 10^{-4}$ Å$^{-1}$. Reported values of $R_H$ are calculated at $q = 20 \times 10^{-4}$ Å$^{-1}$, and a range of $R_H$ is determined through five measurements of the same sample.

## APPENDIX C: TRANSMITTANCE

Transmittance values are calculated by measuring the absorbance $A$ of the silk solution at 488 nm as measured by UV-Vis spectroscopy. The transmittance

$$T = 10^{-(A_{\text{sample}} - A_{\text{water}})} \tag{C1}$$

is normalized by the transmittance of the control sample with zero added acid (ME = 0). Error bars represent the standard deviation of three measurements of the same sample. All measurements have a path length of 1.0 cm. Changing the path length shifts the onset of turbidity, but this shift can be rescaled by the path length.

[1] G. H. Altman, F. Diaz, C. Jakuba, T. Calabro, R. L. Horan, J. Chen, H. Lu, J. Richmond, and D. L. Kaplan, Biomaterials **24**, 401 (2003).
[2] Y. Wang, H. J. Kim, G. Vunjak-Novakovic, and D. L. Kaplan, Biomaterials **27**, 6064 (2006).
[3] D. N. Rockwood, R. C. Preda, T. Yucel, X. Wang, M. L. Lovett, and D. L. Kaplan, Nat. Protoc. **6**, 1612 (2011).
[4] T. Yucel, N. Kojic, G. G. Leisk, T. J. Lo, and D. L. Kaplan, J. Struct. Biol. **170**, 406 (2010).
[5] G. G. Leisk, T. J. Lo, T. Yucel, Q. Lu, and D. L. Kaplan, Adv. Mater. **22**, 711 (2010).
[6] Q. Lu, Y. Huang, M. Li, B. Zuo, S. Lu, J. Wang, H. Zhu, and D. L. Kaplan, Acta Biomater. **7**, 2394 (2011).
[7] A. Matsumoto, J. Chen, A. L. Collette, U.-J. Kim, G. H. Altman, P. Cebe, and D. L. Kaplan, J. Phys. Chem. B **110**, 21630 (2006).
[8] U.-J. Kim, J. Park, C. Li, H.-J. Jin, R. Valluzzi, and D. L. Kaplan, Biomacromolecules **5**, 786 (2004).

[9] A. E. Terry, D. P. Knight, D. Porter, and F. Vollrath, Biomacromolecules **5**, 768 (2004).
[10] A. P. Tabatabai, D. L. Kaplan, and D. L. Blair, Soft Matter **11**, 756 (2015).
[11] T. Lefèvre, M.-E. Rousseau, and M. Pézolet, Biophys. J. **92**, 2885 (2007).
[12] T. Lefèvre, F. Paquet-Mercier, J.-F. Rioux-Dubé, and M. Pézolet, Biopolymers **97**, 322 (2011).
[13] Y. Lin, X. Xia, K. Shang, R. Elia, W. Huang, P. Cebe, G. Leisk, F. Omenetto, and D. L. Kaplan, Biomacromolecules **14**, 2629 (2013).
[14] K. M. Weigandt, D. C. Pozzo, and L. Porcar, Soft Matter **5**, 4321 (2009).
[15] K. M. Weigandt, L. Porcar, and D. C. Pozzo, Soft Matter **7**, 9992 (2011).
[16] I. Greving, C. Dicko, A. Terry, P. Callow, and F. Vollrath, Soft Matter **6**, 4389 (2010).
[17] B. P. Partlow, A. P. Tabatabai, G. G. Leisk, P. Cebe, D. Blair, and D. Kaplan, Macromol. Biosci. **16**, 666 (2016).
[18] C.-Z. Zhou, F. Confalonieri, M. Jacquet, R. Perasso, Z.-G. Li, and J. Janin, Proteins: Struct. Funct. Genet. **44**, 119 (2001).
[19] C. J. Cadwallader and F. O. Howitt, Trans. Faraday Soc. **42**, 642 (1946).
[20] M. A. Webb, Y. Jung, D. M. Pesko, and B. M. Savoie, ACS Cent. Sci. **1**, 198 (2015).
[21] N. Russo, M. Toscano, and A. Grand, J. Phys. Chem. B **105**, 4735 (2001).
[22] J. Jover, R. Bosque, and J. Sales, Dalton Trans. **45**, 6441 (2008).
[23] N. Manin, M. C. da Silva, I. Zdravkovic, O. Eliseeva, A. Dyshin, O. Yasar, D. R. Salahub, A. M. Kolker, M. G. Kiselev, and S. Y. Noskov, Phys. Chem. Chem. Phys. **18**, 4191 (2016).
[24] S. R. Koebley, D. Thorpe, P. Pang, P. Chrisochoides, I. Greving, F. Vollrath, and H. C. Schniepp, Biomacromolecules **16**, 2796 (2015).
[25] G. S. Gorman, J. P. Speir, and C. A. Turner, J. Am. Chem. Soc. **114**, 3986 (1992).
[26] T. Wyttenbach, M. Witt, and M. T. Bowers, J. Am. Chem. Soc. **122**, 3458 (2000).
[27] B. Hammouda, J. Appl. Crystallogr. **43**, 716 (2010).
[28] C. Mo, C. Holland, D. Porter, Z. Shao, and F. Vollrath, Biomacromolecules **10**, 2724 (2009).
[29] M. Boulet-Audet, F. Vollrath, and C. Holland, Phys. Chem. Chem. Phys. **13**, 3979 (2011).
[30] E. S. Feldblum and I. T. Arkin, Proc. Natl. Acad. Sci. USA **111**, 4085 (2014).
[31] C. J. Glinka, J. G. Barker, B. Hammouda, S. Krueger, J. J. Moyer, and W. J. Orts, J. Appl. Crystallogr. **31**, 430 (1998).
[32] D. F. R. Mildner, B. Hammouda, and S. R. Kline, J. Appl. Crystallogr. **38**, 979 (2005).
[33] Y. T. Cheng, D. F. R. Mildner, H. H. Chen-Mayer, V. A. Sharov, and C. J. Glinka, J. Appl. Crystallogr. **33**, 1253 (2000).
[34] J. C. Cook, C. J. Glinka, and I. G. Schröder, Rev. Sci. Instrum. **76**, 025108 (2005).
[35] J. G. Barker, C. J. Glinka, J. J. Moyer, M. H. Kim, A. R. Drews, and M. Agamalian, J. Appl. Crystallogr. **38**, 1004 (2005).
[36] S. R. Kline, J. Appl. Crystallogr. **39**, 895 (2006).
[37] M. Doucet, J. H. Cho, G. Alina, S. King, and P. Butler, SasView Version 4, Zenodo (2016).