MATERIALS SCIENCE

Spider silk self-assembly via modular liquid-liquid phase separation and nanofibrillation

Ali D. Malay[1]*, Takehiro Suzuki[2], Takuya Katashima[1†], Nobuaki Kono[3], Kazuharu Arakawa[3], Keiji Numata[1,4]*

Spider silk fiber rapidly assembles from spidroin protein in soluble state via an incompletely understood mechanism. Here, we present an integrated model for silk formation that incorporates the effects of multiple chemical and physical gradients on the different spidroin functional domains. Central to the process is liquid-liquid phase separation (LLPS) that occurs in response to multivalent anions such as phosphate, mediated by the carboxyl-terminal and repetitive domains. Acidification coupled with LLPS triggers the swift self-assembly of nanofibril networks, facilitated by dimerization of the amino-terminal domain, and leads to a liquid-to-solid phase transition. Mechanical stress applied to the fibril structures yields macroscopic fibers with hierarchical organization and enriched for β-sheet conformations. Studies using native silk gland material corroborate our findings on spidroin phase separation. Our results suggest an intriguing parallel between silk assembly and other LLPS-mediated mechanisms, such as found in intracellular membraneless organelles and protein aggregation disorders.

INTRODUCTION

Spiders produce silks for a wide range of biological functions. Some types of spider silk, such as dragline, have extraordinarily high tensile strength and toughness that can rival or even surpass the performance of advanced artificial fibers. Spider silks are composed primarily of spidroins, large polypeptides that have a distinctive tripartite architecture featuring an extended central repetitive domain (which can comprise hundreds of repeat units in certain silk types), capped on either end by small globular terminal domains. The three domains fulfill distinct biochemical roles; the repeat sequences determine, to a large extent, the mechanical properties of the resultant silk fibers, whereas the terminal domains play crucial roles in orchestrating the silk assembly process in response to environmental cues (*1*–*4*).

The polymerization of dragline spider silk fibers from soluble spidroin precursors is a remarkable process, especially considering its rapidity and how it occurs completely under benign conditions within the spider's body. Before spinning, the spidroins reside as a highly concentrated protein solution (~50% protein) inside the sac of specialized silk glands (*1*, *2*, *5*). A number of models have been proposed to describe the molecular arrangement of the prefibrillar spidroin chains, including formation of protein micelle, a liquid crystalline state, or other granular structures (*6*–*9*). Fiber spinning involves sequential changes to protein conformation as the spidroins travel along an extended spinning duct connected to the gland's storage sac. Here, the spidroin molecules encounter multiple simultaneous changes in external conditions, including acidification (*2*, *10*), ion flux (primarily an exchange of sodium and chloride for potassium and phosphate ions) (*11*, *12*), CO_2 generation (*10*), dehydration (*2*, *11*), and increasing shear forces (*2*, *13*). In response to these external stimuli, the spidroins convert from a largely unfolded state [described as a loose ensemble of random coils, α helices, and 3_1 (poly-proline type II) helices in conformational flux] into solid fibers with hierarchical arrangement of nanofibrils, and that feature aligned nanocrystalline β-sheets (largely contributed by poly-alanine elements in the repeat domains) dispersed against a more amorphous background of protein chains (*2*, *5*, *14*, *15*).

Despite decades of research, we are still far from having a complete picture of how the native spinning process works. On the one hand, there is an abundance of data describing the precise molecular changes that occur within isolated spidroin domains in response to specific chemical stimuli (*10*, *15*–*20*). On the other hand, detailed information on the physical construction and geometry of the silk spinning apparatus help to elucidate the mechanisms for alignment of spidroin chains in the spinning ducts, generation of differential shear forces, etc. (*2*, *8*, *13*). However, there is still a considerable gap between these two levels of understanding; in particular, there is an inadequate understanding of the principal events that take place at the submicrometer level (i.e., mesoscale), which is the level at which certain critical phenomena occur, such as the rapid assembly of hierarchically ordered fibrils from the concentrated silk feedstock.

Additional questions remain that pose crucial challenges to future attempts to generate high-performance biomimetic silk fibers. For instance, while we have a compelling framework to explain the oriented flow of protein molecules during fiber spinning (as per the liquid crystalline model), how do the spidroins avoid premature and catastrophic aggregation, given their extreme concentrations in the silk glands, predominantly unfolded structure, and elevated hydrophobic content of the underlying amino acid sequences? How the multidomain spidroin proteins achieve coordinated structural transitions under simultaneous chemical and physical gradients has also not been rigorously explored to date. These considerations are important as it is becoming clear that to replicate the physical properties of natural silk in the laboratory, it is necessary to follow biomimetic principles (*1*, *21*). In light of these questions, we have conducted the present study on spider silk self-assembly, aimed at probing interactions between the different spidroin functional domains and the associated chemical and physical gradients that occur in vivo, with particular emphasis on probing mesoscale structural changes.

[1]Biomacromolecules Research Team, RIKEN Center for Sustainable Resource Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [2]Biomolecular Characterization Unit, RIKEN Center for Sustainable Resource Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [3]Institute for Advanced Biosciences, Keio University, 246-2 Mizukami, Kakuganji, Tsuruoka, Yamagata 997-0052, Japan. [4]Department of Material Chemistry, Kyoto University, Katsura, Nishikyo, Kyoto 615-8510, Japan.
*Corresponding author. Email: a.malay@riken.jp (A.D.M.); keiji.numata@riken.jp (K.N.)
†Present address: Department of Bioengineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

RESULTS

Generation of MaSp2 constructs

For this study, we developed a recombinant spidroin platform based on the dragline silk protein, major ampullate spidroin 2 (MaSp2). Along with the better-studied MaSp1, MaSp2 is a major component of spider dragline silk produced by web-building spiders. The two MaSp proteins share a similar architecture, with the repeat domains featuring alternating blocks of poly-Ala and Gly-rich regions. The Gly-rich regions of MaSp2 are interspersed with residues Tyr and Pro as well as Gln dyads (QQ) (*22*, *23*), and are predicted to adopt flexible β-turn configurations (*4*, *24*). Because of the lower hydrophobicity of its core repeat sequences compared to MaSp1, we hypothesized that MaSp2 would be more amenable to biochemical investigation owing to a higher solubility, especially in constructs with numerous repeat units. We used a rational design approach to generate chimeric constructs of MaSp2 bearing different domain combinations, including tripartite full-domain constructs featuring an N-terminal domain (NTD or N), tandem repeat domain (R*x*, where *x* denotes the number of repeat units), and C-terminal domain (CTD or C), as well as truncated domain constructs (Fig. 1 and fig. S1). Notably, the repeat regions were designed at the nucleotide level to facilitate the seamless expansion of repeats via iterated cycles of DNA amplification, restriction, and ligation (Fig. 1, A and B).

Analysis of the recombinant MaSp2 sequences (Fig. 1C) shows that the tandem repeat domain has a highly amphiphilic character, with the sequences alternating between hydrophobic poly-Ala stretches and the more hydrophilic Gly-rich portions. Overall, the repetitive regions exhibit low sequence complexity, with only six residue types represented (Gly, Ala, Pro, Gln, Ser, and Tyr). A high level of structural disorder is predicted for the repeat domains; however, in the full-domain constructs, these are flanked by the terminal domains (NTD and CTD) that each fold into globular α-helical bundles (*16*, *18*). The tandem repeat sequence and the latter part of the CTD bear a resemblance to prion-like sequences based on the distribution of polar and aromatic residues (*25*).

The recombinant MaSp2 constructs were expressed in *Escherichia coli* and purified under conditions that were aimed at preserving the native structural state. In particular, the host bacteria were lysed through a combination of enzymatic digestion via lysozyme and a mild detergent, thereby minimizing structural perturbations, as are likely to occur when using methods that rely on shear forces for cellular disruption. Purified MaSp2 was highly soluble and exhibited circular dichroism (CD) spectra characteristic of α-helical conformations (fig. S2A), as expected for properly folded terminal domains. Furthermore, full-domain MaSp2 demonstrated constitutive oligomer formation via the CTD (fig. S2B) (*18*) and inducible dimerization via the NTD in response to acidification (fig. S2C) (*16*), indicating that the terminal domains could preserve native biochemical properties.

Phosphate ions induce LLPS in MaSp2

The array of different MaSp2 constructs was used to systematically explore structural changes that occur in response to the different external stimuli associated with the in vivo silk spinning process. We found that the addition of potassium phosphate (KPi), buffered to neutral pH, produced an instantaneous turbidity in purified

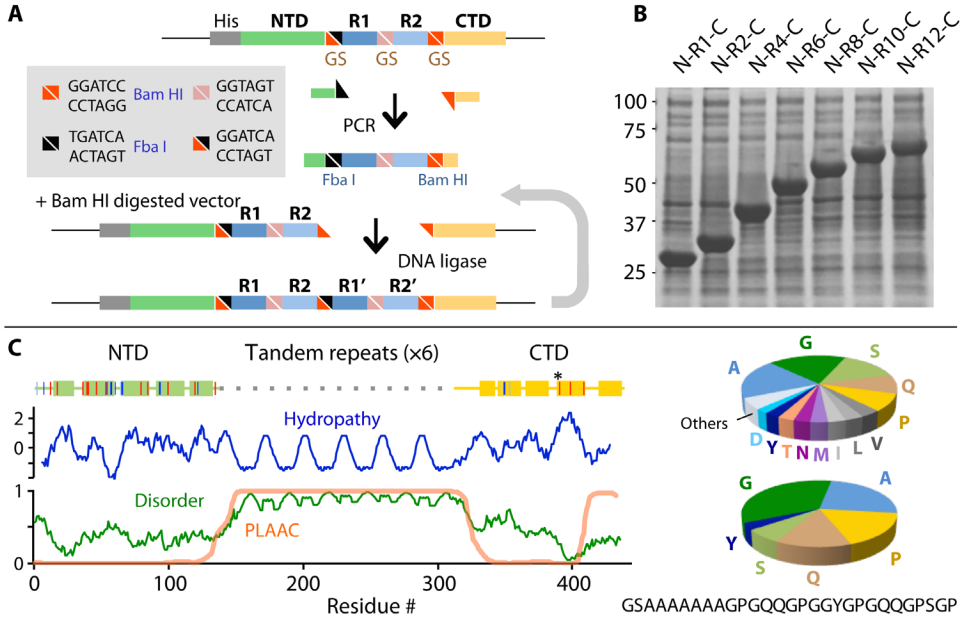

Fig. 1. Rational design of biomimetic MaSp2 constructs. (**A**) Strategy for iterated expansion of tandem repeat domains. R1 and R2 correspond to repeat modules encoding identical amino acid sequences but with differing nucleotide sequences. The inset shows sequences relevant to the cloning method, including Bam HI and Fba I restriction sites and Gly-Ser (GS) coding regions flanking the repeats. (**B**) Expression of recombinant MaSp2 for constructs bearing 1 to 12 repeats, visualized by SDS-PAGE. The overexpressed band in each lane corresponds to MaSp2. Molecular weights are indicated in kDa. (**C**) Amino acid sequence analysis of recombinant MaSp2. The primary structure of N-R6-C is illustrated on the left, with the overall architecture depicted on top, consisting of the NTD (green), repeat units (dotted line), and the CTD (yellow). The predicted α-helical regions in the terminal domains are depicted as solid blocks. The distribution of cationic (blue) and anionic residues (red) is shown, and an asterisk indicates the Cys residue in the CTD involved in disulfide bonding. Below are plots depicting the hydropathy (blue), predicted disordered regions via IUPred2 (green), and predicted prion-like regions using PLAAC (orange). Pie charts on the right depict the relative abundance of amino acids for the entire N-R6-C sequence (top) and for the 29-residue repeat module (bottom), demonstrating the low diversity of residue types.

MaSp2 solution above a critical concentration (Fig. 2A). Centrifugation of the turbid samples led to a separation into two clear phases: a low-density phase (LDP) and a high-density phase (HDP), with MaSp2 partitioning into the HDP to very high concentrations (Fig. 2, A and B). At the microscopic level, the initial cloudy samples exhibited myriad spherical droplets of heterogeneous size (~0.1 to 10 μm) coexisting with the bulk solution phase. These droplets were observed to fuse readily into larger bodies when they come into contact with one another, a hallmark of liquid-liquid phase separation (LLPS) or coacervation (Fig. 2C and movies S1 and S2) (*26*, *27*). We observed well-defined boundaries between the one-phase (solution) and two-phase regions of the sample when prepared under a concentration gradient (movie S3). By excluding water from the HDP, the KPi-dependent LLPS essentially constitutes a self-concentration step and is consistent with the dehydration of soluble silk dope seen during the natural spinning process (movie S4) (*11*).

We generated phase separation diagrams for the different MaSp2 domain combinations (N-R1-C, N-R6-C, N-R12-C, N-R6-*x*, *x*-R6-C, *x*-*x*-C, and *x*-R6-*x*) under constant pH (7.5) and temperature (23°C) (Fig. 2D and fig. S3). For full-domain MaSp2 variants (N-R*n*-C), demixing was initiated at low protein concentrations, e.g., at <10 μM (<0.5 mg ml^{-1}) for N-R6-C and N-R12-C in 0.5 M KPi, in the absence of additional crowding agents. Comparing the different truncation mutants (Fig. 2D), we identified the main drivers for LLPS as the CTD (seen by comparing the N-R6-C and N-R6-*x* diagrams), in synergy with the Rep domains, with the propensity for phase separation increasing with the number of tandem repeats (compare N-R1-C, N-R6-C, and N-R12-C). The isolated Rep domain (*x*-R6-*x*) could undergo LLPS, although to a much lesser degree than *x*-R6-C; CTD by itself (*x*-*x*-C) did not exhibit phase separation under the conditions of the assay, although it readily underwent LLPS at lower pH levels (see Fig. 3 results). NTD itself showed no apparent influence on the induction of LLPS (compare the identical diagrams for N-R6-C and *x*-R6-C).

We expanded our analysis of the ion specificity for LLPS induction, which revealed a strong dependence on the anionic species (Fig. 2E and fig. S4). Apart from phosphate (PO_4^{3-}), the presence of other multivalent kosmotropic anions such as citrate ($C_6H_5O_7^{3-}$) and sulfate (SO_4^{2-}) also induced efficient condensation of MaSp2 from solution, whereas salts containing monovalent anions such as chloride or acetate failed to do so. The identity of the cationic species, on the other hand, did not exert any apparent influence on phase

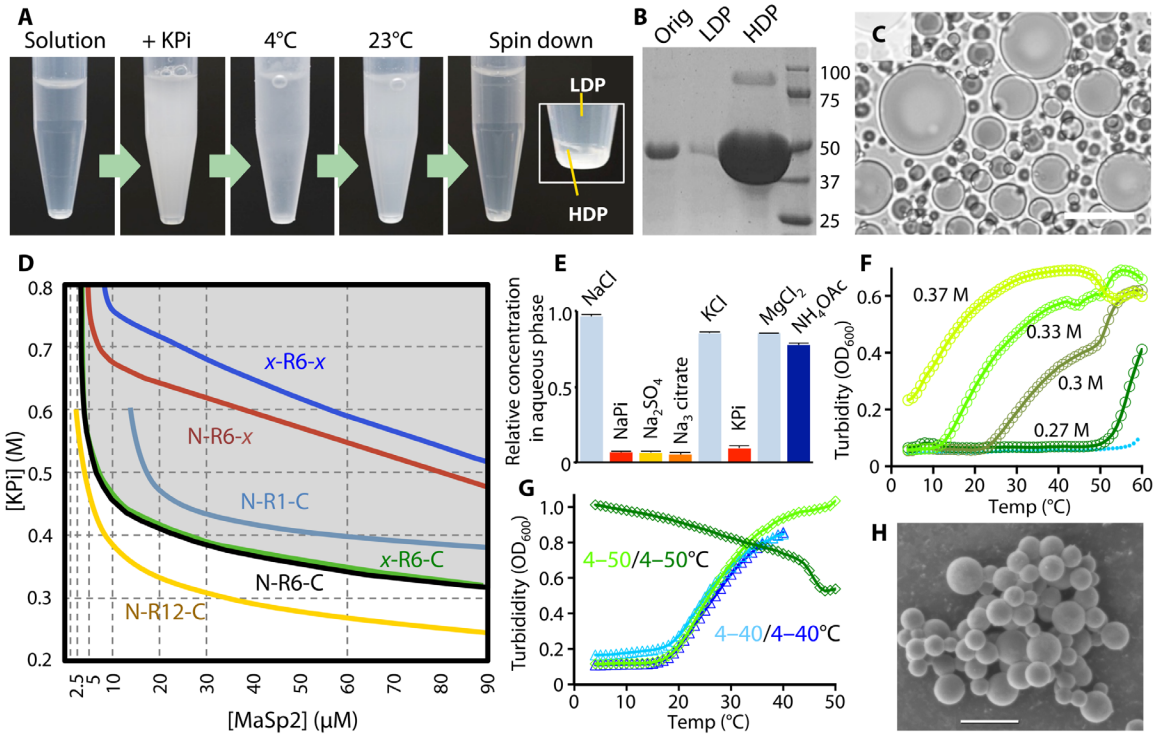

Fig. 2. Multivalent anions induce LLPS in MaSp2. (**A**) Purified MaSp2 [85 μM N-R6-C in 20 mM tris-HCl (pH 7.5), 150 mM NaCl] became turbid upon adding KPi, pH 7.3 (to 0.4 M), with the turbidity varying with temperature. Upon centrifugation, sample separated into LDP and HDP (inset). (**B**) SDS-PAGE of 1-μl aliquots from LDP and HDP in (A) shows the partitioning of MaSp2 into the HDP, with HDP:LDP ratio above 100:1. (**C**) Microscopic droplets formed immediately in a mixture of N-R6-C (4.5 mg/ml) and 1.0 M KPi (pH 7); scale bar, 20 μm. (**D**) Phase separation maps as a function of MaSp2 and KPi (pH 7.5) concentration, for the different domain constructs. All measurements were taken at 23°C against a background of 0.1 M NaCl. Two-phase conditions (LLPS) are on the upper-right section of each curve (illustrated with shading for N-R6-C). (**E**) Ion dependence of LLPS induction. Each sample contained N-R12-C (25 mg/ml) and 0.4 M of different salts as indicated. Samples were mixed and centrifuged, and the protein concentration in the upper surface (corresponding to LDP in case of LLPS) was measured by OD_{280}. Columns are colored by anion species. (**F** and **G**) LLPS transitions in N-R12-C (25 μM) with KPi, pH 7.5, were measured by temperature-dependent turbidity shifts. (F) LCST-type behavior of MaSp2, showing high sensitivity to KPi concentration (indicated). Light blue line shows measurements in the absence of KPi. (G) Samples were subjected to two heating cycles (with intermittent cooling), at either 4° to 40°C (blue) or 4° to 50°C (green), in 0.3 M KPi. Samples subjected to 4° to 40°C cycling displayed reversible LLPS, whereas heating to 50°C destroyed this reversibility (**H**) Scanning electron microscopy image showing solidified LLPS aggregates upon heating the MaSp2/KPi mixture to 70°C for 10 min. Scale bar, 10 μm.

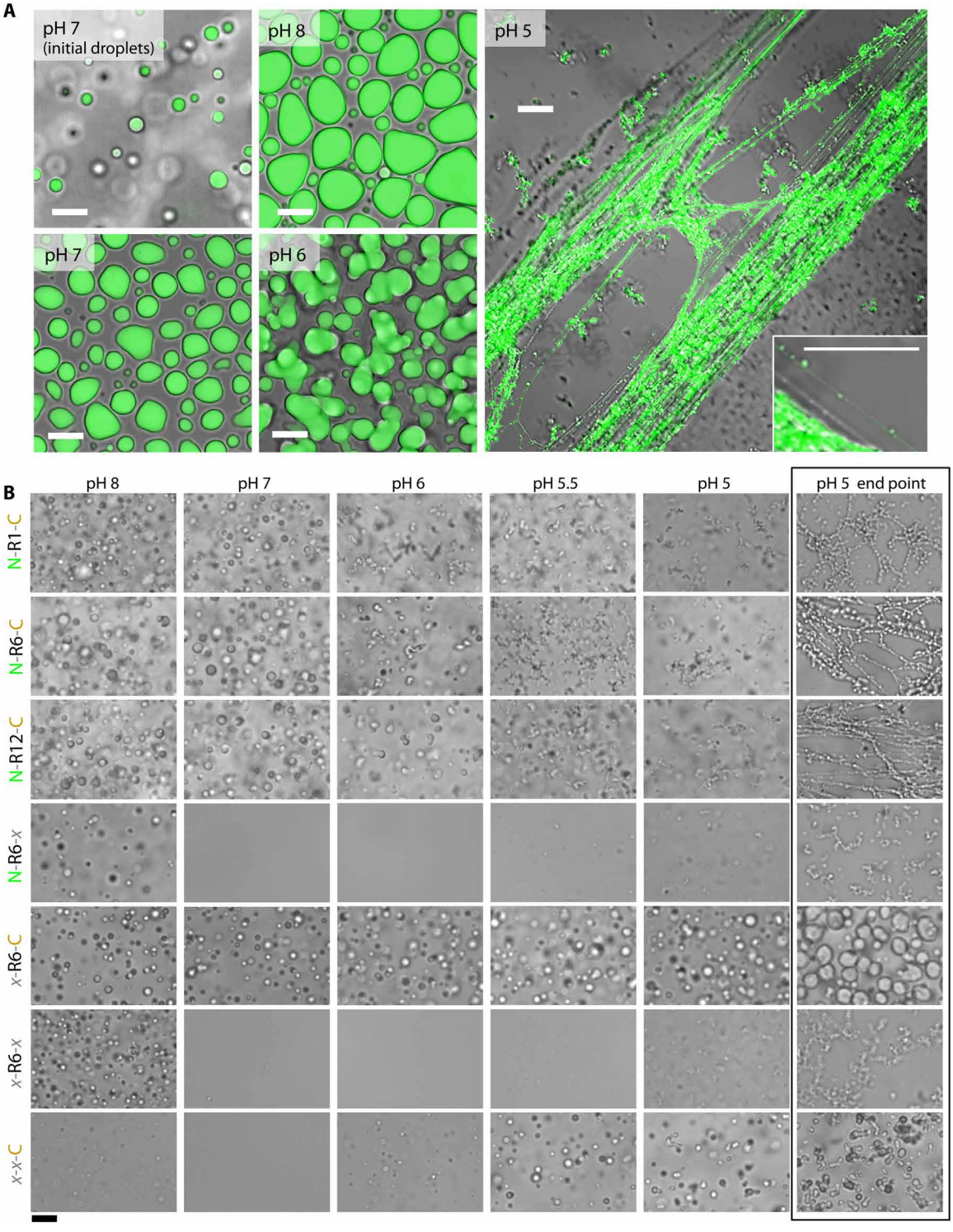

Fig. 3. Acidification triggers rapid self-assembly of MaSp2 nanofibrils. (**A**) N-R12-C labeled with DyLight 488 (10 to 20 mg ml^{-1} final concentration) was mixed into 0.5 M KPi at the indicated pH values and visualized by confocal laser scanning microscopy. Upon mixing of the components, MaSp2 rapidly separated from the aqueous fraction (green structures), eventually settling on the glass surface. At pH 7 and 8, the MaSp2 condensates appear as LLPS droplets undergoing dynamic fusion, while at pH 6, the resultant structures are static with a semisolid appearance. The reaction at pH 5 leads to rapid self-assembly of MaSp2 into frequently aligned, extended fibril networks. The inset shows an individual fibril at pH 5 with a diameter of ~100 nm. Scale bars, 10 μm. (**B**) Different MaSp2 constructs were assessed for their ability to undergo phosphate-induced LLPS and pH-induced fibril self-assembly. Purified MaSp2 (150 μM final concentration) was mixed into a drop of 0.5 M KPi at the indicated pH values on a glass slide. All images in the unboxed region (left) were taken within 30 s of mixing. The boxed region on the right shows the end point of the pH 5 reactions, taken 30 min after initiation. All images were taken at the same scale; scale bar, 10 μm.

separation. Temperature-dependent turbidity measurements (*27*) were used to investigate the MaSp2 phase transitions in greater detail (Fig. 2, F and G). The resulting turbidity profiles, featuring a clear-to-cloudy transition with increasing temperature, displayed an exquisite sensitivity to the phosphate ion concentration and were consistent with the phenomenon of lower critical solution temperature (LCST) (Fig. 2F) (*17*, *28*). The propensity for phosphate-induced phase separation was only minimally affected by the background levels of sodium chloride (fig. S4). The LLPS transition was found to be fully reversible within the range of 5° to 40°C (Fig. 2G); however, heating to 50°C destroyed this reversibility and instead resulted in solidified spherical aggregates (Fig. 2, G and H).

Acidification plus LLPS triggers nanofibril self-assembly

We further investigated the effects of pH decrease on MaSp2 in the context of ion-dependent phase separation. As expected, at pH 7 to 8, mixing of full-domain MaSp2 with KPi resulted in dynamic spherical droplets that coalesced and eventually sedimented onto the glass slide surface (Fig. 3A and fig. S5A). Lowering the pH, however, produced distinct changes in morphology. Addition of KPi, pH 6, to MaSp2 resulted in gel-like deposits that resembled protein droplets arrested in mid-fusion (Fig. 3A). Adding KPi at pH 5 produced the most marked results, namely, the rapid and spontaneous assembly of extended nanofibrillar protein networks. These self-assembled fibrils were frequently seen aligned along a single axis (as shown), or otherwise as randomly oriented fibril networks. These local variations in fibril orientation were likely a by-product of directional sample flow on the glass slide imparted by pipetting or during coverslip application before observation (Fig. 3A and fig. S5B).

The effect of even lower pH values on MaSp2 assembly was explored using citrate buffer in place of phosphate (fig. S5C). As was shown in Fig. 1H, citrate ions promote phase separation, presumably as a consequence of its multivalency. Here, we show that N-R12-C, when mixed with sodium citrate buffered at pH 5 or 4, produced fibrillar protein networks identical in appearance to those observed under KPi, pH 5. MaSp2 mixed with sodium citrate buffer at pH 3 failed to produce any fibrils, instead displaying profuse aggregated material characteristic of denatured protein samples.

Samples prepared using relatively low concentrations of MaSp2 mixed with KPi at pH 5 allowed the tracking of nanofibril assembly in real time and revealed a complex process of near-simultaneous liquid-phase condensation and fibril network formation (movie S5 and fig. S6). Notably, once formed, MaSp2 nanofibrils assembled under KPi at pH 5 were irreversible, with the fibrils maintaining their structural integrity even upon extensive washing with KPi, pH 8, or with Milli-Q water (fig. S7).

Sequence determinants of the ion- and pH-induced nanofibril self-assembly from MaSp2 condensates were assessed using the different domain combinations (Fig. 3B). Constructs bearing the full-domain complement (N-R1-C, N-R6-C, and N-R12-C) exhibited similar behaviors under 0.5 M KPi, forming condensed spherical liquid droplets at neutral pH, and fibril networks at pH 5.0, irrespective of the number of tandem repeats. NTD dimerization via pH decrease was crucial for fibril network assembly, as demonstrated by *x*-R6-C and the N^{A72R}-R6-C mutant, defective for NTD dimerization (*19*), with both proteins persisting in the LLPS state despite acidification (Fig. 3B and fig. S8A). The CTD, which plays a central role in LLPS, also affects the formation of the fibrillar networks: Constructs lacking CTD (N-R6-*x*, x-R6-*x*) produced only loose, fragmented aggregates at pH 5, unlike the tight networks seen with the full-domain constructs (Fig. 3B). Our findings show that the fibril self-assembly depends on the response of linked MaSp2 domains. N-R6-C fibril formation was inhibited by the addition of free NTD in a dose-dependent manner (fig. S8B). Crucially, an equimolar mix of N-R6-*x* and *x*-R6-C failed to yield any visible fibril structures at pH 5, indicating that the presence of NTD and CTD within the same polypeptide chain is essential for fibrillogenesis (fig. S8C). Our results thus demonstrate that the different domains play complementary roles in MaSp2 silk self-assembly, with the complete set of functional domains required to mimic the natural spinning system.

Probing liquid-to-solid transition in MaSp2 condensates

The recombinant MaSp2 system also enabled investigations into the liquid-to-solid phase transitions in spider silk (Fig. 4). Fluorescence recovery after photobleaching (FRAP) was used to probe changes in internal mobility within individual LLPS structures using labeled N-R12-C protein mixed with KPi in a range of pH values (Fig. 4A and fig. S9). The LLPS structures formed at pH 7 to 8 exhibited relatively fast recovery rates, indicating a quasi-liquid state, with apparent mobile fractions (MFs) of ≥75%. However, at pH 6, a steep decline in recovery rate was observed, even within sphere-like droplets, which is consistent with a highly viscous liquid state. The fibril network structures assembled at pH 5 to 5.5 showed minimal fluorescence recovery even over extended times, suggesting almost full loss of mobility in the MaSp2 chains in the fibrillar state.

Rheological tests based on creep measurements further probed the viscoelastic properties of phase-separated MaSp2 as a function of pH (Fig. 4B). In creep tests, solids and liquids show different behaviors: For solids, creep compliance stabilizes toward a constant value, while for liquids, a linear relationship is typically observed such that $J(t) = \eta_0^{-1}t$, where η_0 represents the zero shear viscosity. Because of $J(t)$ oscillations at short time scales (<1 s) arising from inertia of the apparatus, only compliance values for $t > 1$ s were considered in the calculations. The creep compliance values at pH 7 to 8 suggest that the LLPS structures exist in a liquid-like state, while lower pH values resulted in increasing levels of viscosity, with a viscoelastic solid state at pH 5 indicated by a $J(t)$ plateau at approximately 10^2 s. Zero shear viscosities were calculated from the slopes as 0.16 Pa·s at pH 8, 0.11 Pa·s at pH 7, 13 Pa·s at pH 6, and 4.1×10^2 Pa·s at pH 5.5.

Together, the results indicate that decreasing pH induces a liquid-to-solid transition in phase-separated MaSp2, which correlates with the observed self-assembly of fibril networks. This is supported by data on the mutant N^{A72R}-R6-C (inactive for pH-induced fibril assembly), showing minimal change in viscosity as a function of pH (Fig. 4B).

Generation of biomimetic silk fibers

On the basis of the preceding in vitro results, we developed a biomimetic method for fiber production using recombinant MaSp2 that incorporates the regimes of LLPS, acidification, and mechanical stress. In a typical experiment, 0.5 μl of concentrated N-R12-C was added onto a drop of 1.0 M KPi (pH 5) on a glass slide, which yielded a condensed mass of protein fibrils. By manual drawing in air, this condensed protein material could be processed into extended fibers with length of ~10 cm and diameter of ~10 to 20 μm (fig. S10). These fibers showed an internal structure consisting of unidirectionally aligned nanofibrils and thus represent a true hierarchical architecture (Fig. 4C

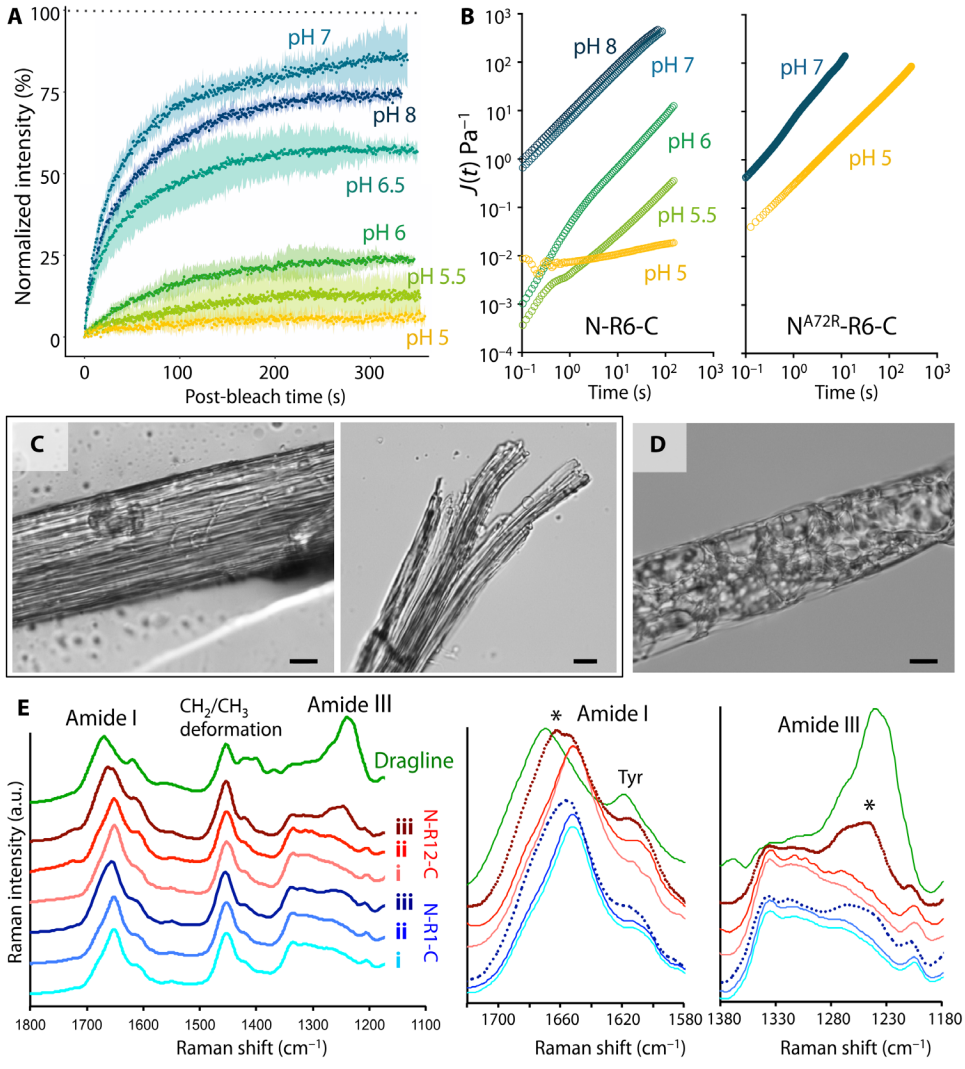

Fig. 4. Probing material state changes in MaSp2 condensates. (**A**) FRAP of MaSp2 condensates formed by mixing labeled N-R12-C in 0.15 M NaCl with 1.0 M KPi at the indicated pH values. Normalized fluorescence intensities are plotted against time elapsed after photobleaching, with the SDs depicted as shaded regions. The dotted line indicates mean intensity before bleaching. (**B**) Viscoelastic creep measurements as a function of pH. Creep compliance values, $J(t)$, were measured against time for wild-type N-R6-C (left) and N^{A72R}-R6-C (right) in 1.0 M KPi at the indicated pH values. (**C** and **D**) Light micrographs of MaSp2 fibers produced via biomimetic methods. Fibers were immersed in Milli-Q water before imaging, which produced some swelling. Scale bars, 10 μm. (C) Hierarchical structure of N-R12-C fibers, showing nanofibrils aligned along the longitudinal axis of the fiber, with a fracture surface shown in the right image. (D) Fibers generated from N^{A72R}-R12-C lacked a hierarchically aligned structure, with an interior appearing to consist of solidified LLPS droplets. (**E**) Representative Raman spectra for N-R1-C and N-R12-C structures under three conditions: (i) LLPS droplets (in 0.5 M KPi, pH 7), (ii) fibril networks (in 0.5 M KPi, pH 5), and (iii) biomimetic fibers prepared as in (C). Dragline silk fiber from *T. clavata* (green) is included as a reference. Right: Close-up views of the amide I and III regions. In the N-R12-C fiber, the shift in the amide I peak with a shoulder at ~1662 cm^{-1} and the amide III peak at ~1242 cm^{-1} signals the emergence of β-sheet structures (asterisks); similar but more modest peak shifts were observed for N-R1-C, which harbors a single repeat domain. The spectra were normalized and displaced along the *y* axis for ease of viewing. a.u., arbitrary units.

and fig. S10). On the other hand, ductile fibers produced from the N^{A72R}-R12-C mutant lacked any visible organization in internal structure, which instead resembled solidified LLPS droplets (Fig. 4D and fig. S10).

Raman spectroscopy was used to probe changes in the secondary structure of MaSp2 at different stages of the biomimetic spinning process (Fig. 4E). The addition of KPi, regardless of pH, did not by itself induce changes in protein conformation: Both LLPS and fibril network structures yielded spectra characteristic of α-helical conformations (~1655 cm^{-1}), consistent with signal contribution from the globular terminal domains, and similar to the spectra seen for native spider silk dope (*14*). In contrast, in the biomimetic fibers that were subjected to manual stretching (as in Fig. 4C), a greater proportion of β-sheet content was clearly apparent, seen as peak shifts in the amide I (~1662 cm^{-1}) and amide III (~1242 cm^{-1}) regions. Notably, the magnitude of the β-sheet amide peak shifts was proportional to the number of tandem repeats in the protein sequence (comparing spectra iii between N-R1-C and N-R12-C), indicating that the β-sheet conformations derive from structural changes in the Rep domains. We hypothesize that the emergence of β-sheets results from the mechanical stress imparted by fiber stretching that formed part of the biomimetic spinning protocol (Fig. 4E).

Investigating native spidroin behavior

Insights into the phase behavior of spidroins were corroborated with studies using native major ampullate (MA) silk gland material isolated from *Trichonephila clavata* spiders. On the basis of the results obtained from recombinant MaSp2, we predicted that native spidroins would exhibit two responses: liquid-phase condensation upon exposure to phosphate ions at neutral pH and spontaneous assembly into fibril structures in response to phosphate ions at acidic pH. We found both predictions to be true (Fig. 5). Addition of KPi at pH 7 to the isolated gland contents produced a turbid sample that could be separated into LDP and HDP fractions by centrifugation. Electrophoresis revealed differential separation of protein species between the two phases (Fig. 5A). Individual protein bands were analyzed by tandem mass spectrometry (MS/MS), which showed a specific partitioning of spidroin species into the HDP (Fig. 5B and data file S1), whereas the LDP contained various nonspidroin proteins and was virtually free of spidroin components. Incidentally, our results support the recent finding of multiple spidroin subtypes within the MA gland, besides MaSp1 and MaSp2 (data file S1) (*29*).

In terms of morphological changes, the MA gland material mixed with 0.3 M KPi led to instantaneous formation of microscopic condensates (representing spidroin material, based on the above); at pH 8, these formed dense liquid-like droplets (Fig. 5C), while at pH 4.5, extensive self-assembled fibril networks were observed (Fig. 5D), consistent with our predictions. The apparent difference in the pH value at which fibril formation is induced for the gland extract (pH 4.5) versus the recombinant MaSp2 (pH 5.0) might reflect differences in sample composition: the MA gland extract contains multiple spidroin species and possibly certain undefined buffering components that may counteract the effect of the KPi buffer; variations in pH response between different spider species also cannot be ruled out.

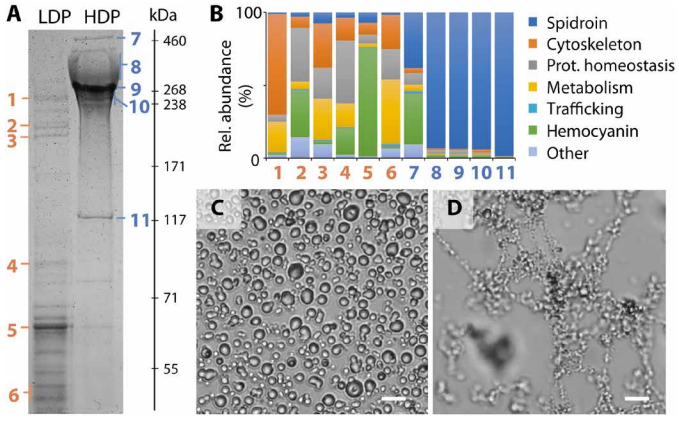

Fig. 5. Native spidroins undergo phosphate-dependent LLPS and fibril assembly. (**A** and **B**) Efficient sequestration of native spidroins from the aqueous phase was induced by the addition of KPi. (A) Harvested *T. clavata* MA gland material was mixed with KPi, pH 7 (at 0.3 M final concentration), leading to sample turbidity, which upon centrifugation separated into LDP and HDP fractions. SDS-PAGE revealed a sharp distinction between the constituents of the two fractions. Prominent bands (numbered 1 to 11) were selected for subsequent analysis. Molecular weights are indicated in kDa. (B) Protein composition in the LDP and HDP fractions. Gel slices from (A) were subjected to protein sequence identification and quantification via MS/MS and MASCOT analysis. The results are plotted as relative abundance of proteins in each of the 11 bands, grouped according to their general function. Details of the analysis are supplied in data file S1. Notably, the HDP was highly enriched for spidroins and comprise a number of different spidroin types beside MaSp1 and MaSp2, including homologs of MaSp-e, MaSp-g, and Sp-74867 as recently described (*29*). In contrast, the LDP was populated mostly with proteins involved in cytoskeletal functions, protein homeostasis (amino acid metabolism, protein synthesis, and degradation, etc.), as well as hemocyanin (the latter possibly a contaminant carried over from the spider hemolymph). (**C** and **D**) Morphological changes in concentrated MA gland material upon addition of 0.5 M KPi at pH 8 (C), showing liquid droplet formation, while addition of 0.5 M KPi at pH 4.5 (D) resulted in self-assembly of fibril networks. Scale bars, 10 μm.

DISCUSSION

The MaSp2 platform presented here is able to recapitulate, to a large extent, the structural transformations reported for native spidroins under biomimetic conditions. Overall, our findings suggest that the spider silk's ability to form hierarchically organized fibers is, by and large, an intrinsic property of the spidroin molecules themselves; with the appropriate external stimuli, mesoscale self-assembly occurs rapidly and with minimal intervention, provided that all the spidroin functional domains are present and in the native state (Fig. 6A). Central to this process is an LLPS transition induced by multivalent anions, a finding that is consistent with the presence of a positive gradient of phosphate ions along the spider silk spinning ducts. Although our study is focused on MaSp2, from the conservation of domain sequences, we hypothesize that similar mechanisms operate on other spidroin types as well. For instance, nuclear magnetic resonance studies have shown that isolated MaSp1 repeat domains respond to kosmotropic ions by adopting more compact conformations through hydrogen bond interactions facilitated by the Gly-rich sequences (*30*). We should note that phase separation behavior has previously been reported in other spidroin variants (mainly recombinant ADF3 from *Araneus diadematus*); however, their relevance to the native silk assembly mechanism has, to a large extent, not been explored (*17*, *31*).

From a polymer physics framework, the conformational behavior of protein chains in solution is governed by the interplay among the polypeptide backbone-, sidechain-, and solvent-mediated interactions. In the case of MaSp2, the introduction of phosphate ions likely alters the character of the aqueous environment such as to promote the exclusion of the bulk solvent in favor of molecular interactions between protein backbone and sidechain elements. The conformation of the spidroins in these states, e.g., whether they adopt extended arrangements and collapsed globule forms or undergo more subtle structural changes, and such questions related to "solvent quality" (*32*) are still largely unexplored, but could yield important insights into the ability of spider silk to assemble into hierarchical fibers. For instance, determining the extent of prealignment of chains within the condensed LLPS state could illuminate the mechanism by which individual protein chains come into register to enable formation of aligned and oriented β nanocrystals in the mature silk fiber, as opposed to just randomly oriented protein aggregates. Ultimately, theoretical models to explain the behavior of the spidroin chains at the molecular level would likely need to be relatively complex; the interplay between the different functional modules—the long disordered repeat sequences and the structurally and biochemically complex terminal domains—would require careful consideration.

From the literature, a common theme regarding phase-separating proteins is a high degree of structural disorder, often associated with the concept of intrinsically disordered proteins (IDPs). IDP sequences responsible for LLPS often bear repeating modules of low sequence

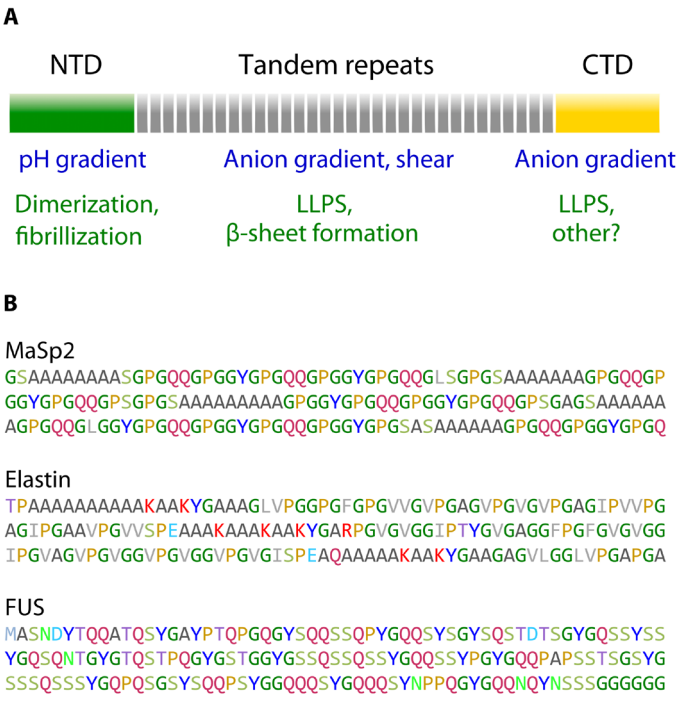
A
NTD
Tandem repeats
CTD
pH gradient
Anion gradient, shear
Anion gradient
Dimerization, fibrillization
LLPS, β-sheet formation
LLPS, other?
B
MaSp2
GSAAAAAAAAASGPGQQGPGGYGPGQQGPGGYGPGQQGLSGPGSAAAAAAAGPGQQGP
GGYGPGQQGPSGPGSAAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPSGAGSAAAAAA
AGPGQQGLGGYGPGQQGPGGYGPGQQGPGGYGPGSASAAAAAAGPGQQGPGGYGPGQ
Elastin
TPAAAAAAAAAAKAAKYGAAAGLVPGGPGFGPGVVGVPGAGVPGVGVPGAGIPVVPG
AGIPGAAVPGVVSPEAAAKAAAKAAKYGARPGVGVGGIPTYGVGAGGFPGFGVGVGG
IPGVAGVPGVGGVPGVGGVPGVGISPEAQAAAAAKAAKYGAAGAGVLGGLVPGAPGA
FUS
MASNDYTQQATQSYGAYPTQPGQGYSQQSSQPYGQQSYSGYSQSTDTSGYGQSSYSS
YGQSQNTGYGTQSTPQGYGSTGGYGSSQSSQSSYGQQSSYPGYGQQPAPSSTSGSYG
SSSQSSSYGQPQSGSYSQQPSYGGQQQSYGQQQSYNPPQGYGQQNQYNSSSGGGGGG

Fig. 6. Insights into sequence-based phase separation of MaSp2 toward silk assembly. (**A**) Division of labor encoded within the modular architecture of MaSp2. A representation of the different spidroin domains is shown above, while the specific external stimuli each domain responds to during fiber self-assembly (blue) and the corresponding changes in protein organization (green) are shown below. In the case of CTD, a structural change in response to pH decrease has also been proposed (*10*). The formation of native-like hierarchical silk fibers depends on cooperative action between the different modular units. (**B**) Sequence comparison between MaSp2 tandem repeats and two archetypal IDPs with LLPS functions. Representative sections of the MaSp2 tandem repeats from *T. clavipes* (UniProt code P46804), human elastin (UniProt code P15502), and the low-complexity region of human FUS RNA binding protein (UniProt code P35637) are shown.

complexity, which give rise to multivalent chain arrangements that provide numerous weak and transient bonding interactions between the repeat units, enabling the protein to form a condensed liquid state while avoiding catastrophic structural collapse into insoluble aggregates (*26*, *27*, *33*).

For MaSp2, the precise sequence features involved in the LLPS response are outstanding questions, particularly with regard to the contributions of the CTD, which, unlike many phase-separating proteins, maintains a well-ordered globular structure. For the repetitive regions, in contrast, a comparison with well-characterized LLPS proteins might yield important clues (Fig. 6B). The elastin precursor, for instance, is a disordered protein with a low-complexity sequence that consists of alternating hydrophobic blocks (enriched for Gly, Pro, Val, and Ala) and cross-linking blocks (poly-Ala stretches interspersed with Lys residues) (*34*). The ability of elastin to maintain a liquid state upon condensation, despite its strongly hydrophobic nature, has been attributed to the presence of structure-breaking Pro-Gly motifs distributed throughout the sequence (likewise a feature of the MaSp2 repeat domains). Like MaSp2, elastin-like proteins exhibit an LCST-type phase separation that is enhanced by kosmotropic anions, which are thought to exert their effect through direct local interactions with the water molecules in the protein hydration shell (*35*). Another prominent example is FUS protein, involved in inducible stress granule formation inside cells. The LLPS response of FUS is governed by a prion-like domain of low complexity that is enriched for residues Ser, Tyr, Gly, and Gln (likewise constituents of the Gly-rich region in MaSp2 repeats), with the phase separation behavior dependent on the distribution of the bulky Tyr and Gln polar residues within the sequence (*33*, *36*, *37*).

Notably, many proteins that undergo LLPS as part of normal biological function (such as mediating the formation of intracellular membraneless organelles) also show a marked tendency to undergo liquid-to-solid transition toward insoluble aggregates, often with pathological outcomes. These transitions form the basis for a wide range of neurodegenerative disorders, such as those associated with the deposition of tau protein in Alzheimer's disease, FUS, or hnRNPA1 in amyotrophic lateral sclerosis, among others (*27*). Our results suggest that for MaSp2, and perhaps for spidroins in general, a related mechanism might be involved, whereby the condensed liquid form (signified by LLPS) occupies an intermediate state between individual protein chains in solution and the mature fibrillar deposits. A main difference is that whereas in the pathological aggregate states, fibril deposition is a stochastic process (often modulated via deleterious mutations), in spider silk assembly, phase transition and fibril assembly are tightly orchestrated molecular-level events, mediated by interactions between the protein modules and the biochemical gradients found in the spinning ducts.

Our present findings suggest broad opportunities for future investigations related to spider silk self-assembly mechanisms. Mutagenesis studies on spidroin constructs combined with structural characterization will likely help to elucidate the LLPS phenomenon in greater detail. An assessment of spidroin biochemical behavior in light of the theoretical frameworks for protein phase separation and fibrillogenesis that are being developed in the evolving field of biomolecular condensates is expected to yield important insights into the nature of silk fiber formation.

MATERIALS AND METHODS

Amino acid sequence analysis

Overall sequence analysis was performed using ProtParam (ExPASy), including hydrophobicity analysis via ProtScale using Kyte and Doolittle parameters (window size, 11). Disorder prediction was carried out using IUPred2 (*38*), and PLAAC (*25*) was used to evaluate prion-like sequence features. Secondary structure prediction of the terminal domains was based on reported structures of similar spidroin domains (*16*, *18*).

Genetic constructs

Initial spidroin constructs were purchased from GenScript in the pET15b vector (Nde I/Xho I inserts), including isolated NTD and CTD domains from the different spider species and the initial full-domain MaSp2 constructs N-R1-C and N-R2-C. From optimization of expression and solubility screening (fig. S1), the full-domain MaSp2 constructs were designed as chimeras featuring NTD and repeat domains derived from MaSp2 of *Trichonephila clavipes* (formerly *Nephila clavipes*) and the CTD corresponding to MaSp2 of *Latrodectus hesperus*. Nucleotide sequences were optimized to reduce secondary structure formation and to facilitate the recursive strategy for tandem repeat amplification (see below). The N-R2-C construct was further modified by site-directed mutagenesis (see

below) to disable the Bam HI site on the plasmid to enable the repeat amplification strategy.

Amplification of tandem repeat coding sequences

Starting from the MaSp2 N-R2-C plasmid construct, the repeat domains were amplified by polymerase chain reaction (PCR) using KOD+ polymerase with the flanking primers *Rep_expand_F* (5′-CGAATGAAGTGTCTTATGGAGGCGCATGA-3′) and *Rep_expand_R* (5′-ACATATTGTGATGGTCCATAGCCACCGGA-3′); *Rep_expand_F* harbors a single base mismatch (underlined T) to generate an Fba I restriction site. The resultant PCR product was purified from a 1% agarose gel using a Qiagen gel purification kit, digested with Fba I and Bam HI restriction enzymes (Takara), and purified using a Qiagen PCR purification kit. The original N-R2-C plasmid was digested with Bam HI, treated with Quick CIP (NEB) to prevent plasmid recircularization, and purified from a 1% agarose gel using a Qiagen gel purification kit. The digested vector and insert were ligated using Ligation High v.2 (Toyobo) with 50 ng of vector and threefold molar excess of insert for 2 hours at room temperature, and the product was transformed into XL10-Gold ultracompetent *E. coli* cells (Stratagene). The results were screened by colony PCR using GoTaq DNA polymerase (Promega) with T7 primers, followed by DNA sequencing using the BigDye Terminator v3.1 reaction to confirm the correct insert orientation. The process was performed recursively, adding two extra tandem repeats per cycle to generate plasmid constructs with 4, 6, 8, 10, and 12 repeat sequences.

Site-directed mutagenesis

Truncation and point mutations were generated via PCR site-directed mutagenesis using primers with 11 to 18 complementary bases flanking the substitution site and 30 amplification cycles using KOD+ polymerase (Toyobo) followed by Dpn I digestion (Takara) and transformation into XL10-Gold ultracompetent *E. coli* cells (Stratagene). Correct products were confirmed by DNA sequencing using the BigDye Terminator v3.1 reaction (Thermo Fisher Scientific).

Expression of recombinant MaSp2

The plasmid of interest was transformed into *E. coli* strain BL21(DE3) and grown overnight as 100-ml preculture in LB medium (BD) with ampicillin (100 μg ml^{-1}; Sigma-Aldrich) at 37°C with shaking at 180 rpm. The preculture was used to inoculate 2 to 8 liters of main culture that was grown at 37°C with shaking until the OD_{600} (optical density at 600 nm) reached ~0.8, after which the temperature was lowered to 20°C and protein expression was induced overnight with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Wako). Cells were harvested by centrifugation, resuspended in lysis buffer [50 mM tris-HCl (pH 7.5), 10% glycerol, 0.1% Triton X-100], and stored at −80°C until further use.

Purification of recombinant MaSp2

Frozen cell slurry was thawed at 37°C and supplemented with hen egg lysozyme (Wako) at 100 mg per 2-liter culture, 250 U of TurboNuclease (Accelagen), and cOmplete EDTA-free protease inhibitor cocktail (Roche). After incubation for 1 hour at room temperature with stirring, the lysates were centrifuged at 8000 rpm for 30 min at 4°C, and the supernatant fraction was transferred to fresh Falcon tubes and centrifuged again for 30 min. The clear supernatant was loaded onto a 5-ml HisTrap column (GE Healthcare) and washed extensively with a solution containing 20 mM tris-HCl, 20 mM imidazole, and 500 mM NaCl (pH 7.5) using an AKTA Explorer (GE Healthcare) until baseline levels of absorbance were achieved. Bound protein was eluted using buffer containing 20 mM tris-HCl (pH 7.5), 500 mM imidazole, and 500 mM NaCl. The eluted product was concentrated and buffer-exchanged using in a VivaSpin centrifugal concentrator with 10-kDa cutoff (GE Healthcare) at 8000 rpm against 20 mM tris-HCl (pH 7.5), with 0.15 M NaCl at 4°C. The poly-histidine tag was cleaved by digestion with thrombin (Sigma-Aldrich) at 4°C overnight, and then the protein was further washed and concentrated in the same buffer using VivaSpin to a final volume of 0.1 to 0.3 ml. The protein concentration was calculated by measuring the absorbance at 280 nm using a NanoDrop instrument (Thermo Fisher Scientific). Protein purity was assessed by SDS–polyacrylamide gel electrophoresis (SDS-PAGE). All subsequent experiments using MaSp2 used either protein in freshly purified state or else from frozen stocks (concentrated MaSp2 flash-frozen in liquid nitrogen and stored at −80°C until further use).

SDS-polyacrylamide gel electrophoresis

Recombinant spidroin samples were electrophoresed on a Mini-Protean TGX 4 to 20% denaturing gel (Bio-Rad). Samples extracted from spider silk glands were run on a NuPAGE 3 to 8% tris-acetate gel (Thermo Fisher Scientific) under denaturing conditions. Gels were stained with 0.1% Coomassie brilliant blue in 40% ethanol and 10% acetic acid and destained with Milli-Q water.

Far-ultraviolet CD

CD spectra were measured at 25°C using a JASCO J820 instrument. Dilute MaSp2 at a final concentration of 3.7 μM was prepared in 10 mM sodium phosphate buffer (400 μl final volume), and spectra were recorded in 1-nm increments from 250 to 190 nm at a 100 nm min^{-1} rate with a 1-s response time in a Starna 21Q quartz cuvette. Experiments were performed three times, with identical results.

NTD dimerization assay

NTD dimerization as a function of pH was monitored based on tryptophan fluorescence shifts (*16*). Each 100-μl assay contained 5 μM recombinant MaSp2 with 150 mM NaCl, and the pH was controlled by means of a mixed buffer system containing 20 mM each of sodium acetate, MES, and Hepes, with the pH adjusted to values ranging from 4.75 to 7.5 (in increments of 0.25 pH units). The assays were prepared in black 96-well plates (Iwaki), and fluorescence spectra were obtained at 25°C using a SpectraMax M3 instrument (Molecular Devices) with excitation at 280 nm and emission from 300 to 400 nm. Response curves were generated by calculating the ratio of emission intensity between 339 and 351 nm and plotting the results against pH.

Buffers for ion-induced phase separation

For the LLPS experiments performed under specific pH values, the pH of the sample was controlled via the buffering capacity of either the phosphate or citrate components at 25°C. The different KPi buffers were prepared by mixing calculated ratios of 1 M stock components (dipotassium hydrogen phosphate, potassium dihydrogen phosphate, and phosphoric acid). Similarly, the citrate buffers were prepared by mixing calculated ratios of 1 M trisodium citrate and citric acid. The correct pH values were ascertained using a pH meter or pH paper.

Optical microscopy

Samples were observed at 23°C using an Olympus BX53 microscope equipped with a DP73 camera via Stream Essentials software version 1.9.3 (Olympus). For prolonged observations, samples were deposited on a glass slide with a customized chamber made using vinyl adhesive tape and overlaid with a coverslip.

Phase separation maps

Samples containing different combinations of protein concentration (2.5 to 90 μM range) and KPi concentration (0.2 to 0.8 M range) were evaluated for phase separation by optical microscopy at 23°C. Components were initially mixed in PCR tubes at 20-μl final volumes, and a 5-μl aliquot was applied onto a glass slide and immediately observed to detect the formation of LLPS droplets. Each data point ("y/n" LLPS) was tabulated onto a grid representing MaSp2 concentration versus KPi concentration, from which best-fit curves were manually constructed. Each sample point was prepared at least twice, which yielded identical results.

Turbidity assays

One-phase to two-phase LLPS transitions were measured as turbidity shifts against a temperature gradient using a JASCO J820 instrument equipped with a PTC-423L Peltier temperature controller (JASCO). Samples containing MaSp2 and KPi (400 μl total volume) were placed in a Starna 21Q quartz cuvette and allowed to equilibrate for 10 min. The experiment was initiated upon applying a temperature ramp of +1°C min^{-1}, and absorbance values were recorded at 1-min intervals at 600 nm.

Fluorescent labeling of MaSp2

One milligram of purified MaSp2 N-R6-C or N-R12-C was buffer-exchanged against a solution containing 20 mM sodium phosphate buffer (pH 8.5) and 150 mM NaCl using an Amicon Ultra [50-kDa molecular weight cutoff (MWCO)] and then reacted overnight at 4°C with 50 μg of DyLight 488 *N*-hydroxysuccinimide (NHS) ester (Thermo Fisher Scientific), which was followed by further centrifugation in an Amicon Ultra (50-kDa MWCO) to obtain concentrated fluorescently labeled protein samples. No differences in phase separation behavior or mesoscale structural morphology were observed between the unlabeled and labeled samples.

Confocal laser scanning microscopy

Samples were imaged on a Zeiss LSM 880 confocal scanning microscope using a 448-nm excitation laser and 40×/1.3 numerical aperture (NA) or 63×/1.4 NA oil immersion objectives. Purified MaSp2 and KPi at the indicated concentrations and pH values were mixed in a customized vinyl tape chamber on a glass slide, overlaid with a coverslip, and sealed with nail polish. Images were processed using Zen Black (Zeiss) and Fiji software.

Fluorescence recovery after photobleaching

Samples were prepared by pipetting 1.5 μl of N-R12-C (20 mg ml^{-1} in 0.15 M NaCl, with a 2:1 ratio of unlabeled:DyLight 488–labeled protein) into a 1.5-μl drop of 1.0 M KPi at the appropriate pH on a glass slide with a customized chamber. Intra-droplet FRAP was performed using a Zeiss LSM 880 confocal microscope with a 63×/1.4 NA oil immersion objective. For samples prepared at pH 6 to 8, approximately equal-sized spherical protein droplets (~5 to 10 μm in diameter) resting on the glass surface were selected for analysis. For samples prepared at pH 5 and 5.5, which formed fibrillar networks, discrete clumps of protein condensates of similar dimensions were analyzed. The sample fluorescence was bleached in a 0.55-μm-diameter spot in condensates with a size 5 to 10× the bleach diameter, and time-lapse images were collected at approximately 773-ms intervals for >5 min. For each test, three regions of interest were analyzed, including the bleached spot, the overall droplet region, and a background region, using Fiji. The results represent replicates of several experiments (pH 8, $N = 6$; pH 7, $N = 6$; pH 6.5, $N = 7$; pH 6, $N = 8$; pH 5.5, $N = 6$; pH 5, $N = 6$). The data were analyzed using easyFRAP (*39*), with full-scale normalized curves used to calculate the MF and half-maximal recovery time ($t_{1/2}$) following biexponential curve fitting.

Rheological analyses

The time development of strain $\gamma(t, \sigma)$ for phase-separated MaSp2 under constant stress, σ, was measured at 25°C with the stress-controlled rheometer MCR 301 (Anton Paar). All measurements were performed using a cone-and-plate fixture. The diameter of the cone was 10 mm, and the cone angle was 2°. The applied stress ranged from 1.0 to 5.0 Pa, which was preliminarily checked to be within the range of linear viscoelasticity. From the obtained strain data, the creep compliance, $J(t)$, was evaluated using the equation

$$J(t) = \gamma(t, \sigma) / \sigma$$

Purified MaSp2 was pipetted onto the plate (16 μl of 116.8 mg ml^{-1}), and 4 μl of 1.0 M KPi at the appropriate pH was added. The creep measurements were initiated approximately 10 min after the cone fixture was set.

For the pH 5 experiment, we conducted oscillatory measurements to investigate the time development of storage and loss moduli (G' and G'', respectively) with strain $\gamma = 1\%$ and angular frequency $\omega = 10$ rad·s^{-1} (before creep test initiation).

Biomimetic fiber formation

In a typical experiment, a small volume (~0.5 μl) of >100 mg ml^{-1} recombinant MaSp2 was pipetted onto a glass slide and covered with 4 μl of 1.0 M KPi (pH 5), turning the protein sample turbid. Using forceps, the sample was carefully pulled from one end of the condensed protein mass and extended unidirectionally in the air environment. The resultant solid fibers routinely exceeded 10 cm in length, although premature breakage sometimes occurred.

Raman spectroscopy

Confocal laser Raman spectra were recorded using JASCO NRS-4100 equipped with a DU420-OE charge-coupled device detector (Andor). Samples were placed on a glass slide with a customized chamber made from vinyl adhesive tape, upon which a coverslip was placed and sealed with nail polish. A 532-nm excitation source was used, with the beam focused using a 100× Plan Achromat objective lens (oil), a grating of 900 grooves mm^{-1}, and a 100 μm × 8000 μm slit size. Spectra were acquired from 500 to 2500 cm^{-1} at 10-mW beam intensity with typical exposures of 3 × 30 s. No sign of sample deterioration was observed under these conditions. All fiber samples were immersed in 0.5 M KPi (pH 5) to ensure a consistent background environment. The collected spectra were calibrated internally to the peak maximum at 1453 cm^{-1}, corresponding to deformation vibrations of CH_2/CH_3 groups. At least eight spectra were collected for each sample, which showed similar results.

Scanning electron microscopy

Samples were placed on an aluminum stub with conductive carbon tape, sputtered with gold using Smart Coater (JEOL), and visualized on a JCM-6000 instrument (JEOL) at 5 kV in secondary electron imaging mode under high vacuum.

Native spidroin purification

Adult female *T. clavata* spiders were collected from areas around Tokyo. CO_2-anaesthesized individuals were sacrificed and immediately dissected under 150 mM NaCl and 0.01% NaN_3 to obtain the pair of MA glands. Proteins from the MA gland lumen were obtained by cutting off the spinning duct region and gently extruding the contents of the ampulla into a 100-µl drop of 150 mM NaCl at 4°C using forceps. The dense sample was solubilized by slow pipetting followed by centrifugation for 5 min at 21,000*g* at 4°C to remove insoluble debris. The supernatant fraction was concentrated using an Amicon Ultra 0.5-ml device (50-kDa MWCO) to a final volume of <50 µl. Phase separation–mediated purification was achieved by the addition of KPi (pH 8.0) to a final concentration of 0.3 M, followed by centrifugation at 21,000*g* for 5 min at 4°C to produce LDP and HDP fractions. The HDP was solubilized in 50 µl of 150 mM NaCl, and the fractions were analyzed by SDS-PAGE.

MS of spider silk gland material

Following electrophoresis, protein bands of interest were excised, destained, and digested with trypsin [trypsin treated with N-tosyl-L-phenylalanine chloromethyl ketone (TPCK), Worthington Biochemical Co.]. The resultant peptides were analyzed by liquid chromatography–MS/MS using Easy-nLC 1000 and Q Exactive (Thermo Fisher Scientific GmbH, Bremen, Germany). The peptide mixture was separated with a linear gradient from 0 to 80% solvent B (100% acetonitrile/0.1% formic acid) over 20 min at a flow rate of 300 nl min^{-1} using a nano-ESI spray column (100 mm length × 75 µm internal diameter, 3 µm opening, NTCC analytical column C18; Nikkyo Technos). The mass spectrometer was operated in the positive ion mode, and spectra were acquired using a data-dependent top 10 method. The acquired spectra were analyzed using Mascot software version 2.5 (Matrix Science) (*40*) by searching against a database consisting of Araneae sequences and bacterial sequences from NCBI_nr and Swiss-Prot, respectively, that was supplemented by full-length and partial spidroin amino acid sequences from *T. clavata* that were generated in house. The following parameters were used: type of search, MS/MS ion search; enzyme, trypsin; fixed modification, none; variable modifications, acetyl (Protein N-term), Gln->pyro-Glu (N-term Q), oxidation (M), propionamide (C); mass values, monoisotopic; peptide mass tolerance: ±15 ppm; fragment mass tolerance: ±30 mmu; max missed cleavages: 3; and instrument type, ESI-TRAP. Protein abundance values were calculated from the sum of the ion intensities of peptides using Proteome Discoverer Software version 2.2 (Thermo Fisher Scientific); human-derived contaminants such as keratin were excluded from the analysis using the built-in filtering function. In the case of spidroin repetitive sequences, ion intensity values were normalized against the number of repeats in the full-length sequences. To minimize false-positive results, a cutoff value of 100 was applied to the Mascot scores.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/45/eabb6030/DC1

REFERENCES AND NOTES

1. A. Rising, J. Johansson, Toward spinning artificial spider silk. *Nat. Chem. Biol.* **11**, 309–315 (2015).
2. F. Vollrath, D. P. Knight, Liquid crystalline spinning of spider silk. *Nature* **410**, 541–548 (2001).
3. M. Heim, D. Keerl, T. Scheibel, Spider silk: From soluble protein to extraordinary fiber. *Angew. Chem. Int. Ed. Engl.* **48**, 3584–3596 (2009).
4. J. L. Yarger, B. R. Cherry, A. Van Der Vaart, Uncovering the structure–function relationship in spider silk. *Nat. Rev. Mater.* **3**, 1–11 (2018).
5. D. H. Hijirida, K. G. Do, C. Michal, S. Wong, D. Zax, L. W. Jelinski, ^{13}C NMR of *Nephila clavipes* major ampullate silk gland. *Biophys. J.* **71**, 3442–3447 (1996).
6. H.-J. Jin, D. L. Kaplan, Mechanism of silk processing in insects and spiders. *Nature* **424**, 1057–1061 (2003).
7. T. Y. Lin, H. Masunaga, R. Sato, A. D. Malay, K. Toyooka, T. Hikima, K. Numata, Liquid crystalline granules align in a hierarchical structure to produce spider dragline microfibrils. *Biomacromolecules* **18**, 1350–1355 (2017).
8. F. Vollrath, D. P. Knight, Structure and function of the silk production pathway in the spider *Nephila edulis*. *Int. J. Biol. Macromol.* **24**, 243–249 (1999).
9. L. R. Parent, D. Onofrei, D. Xu, D. Stengel, J. D. Roehling, J. B. Addison, C. Forman, S. A. Amin, B. R. Cherry, J. L. Yarger, N. C. Gianneschi, G. P. Holland, Hierarchical spidroin micellar nanoparticles as the fundamental precursors of spider silks. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 11507–11512 (2018).
10. M. Andersson, G. Chen, M. Otikovs, M. Landreh, K. Nordling, N. Kronqvist, P. Westermark, H. Jornvall, S. Knight, Y. Ridderstrale, L. Holm, Q. Meng, K. Jaudzems, M. Chesler, J. Johansson, A. Rising, Carbonic anhydrase generates CO_2 and H^+ that drive spider silk formation via opposite effects on the terminal domains. *PLOS Biol.* **12**, e1001921 (2014).
11. E. K. Tillinghast, S. F. Chase, M. A. Townley, Water extraction by the major ampullate duct during silk formation in the spider, *Argiope aurantia* Lucas. *J. Insect Physiol.* **30**, 591–596 (1984).
12. D. P. Knight, F. Vollrath, Changes in element composition along the spinning duct in a *Nephila* spider. *Naturwissenschaften* **88**, 179–182 (2001).
13. J. Sparkes, C. Holland, Analysis of the pressure requirements for silk spinning reveals a pultrusion dominated process. *Nat. Commun.* **8**, 594 (2017).
14. T. Lefèvre, S. Boudreault, C. Cloutier, M. Pézolet, Conformational and orientational transformation of silk proteins in the major ampullate gland of *Nephila clavipes* spiders. *Biomacromolecules* **9**, 2399–2407 (2008).
15. N. A. Oktaviani, A. Matsugami, A. D. Malay, F. Hayashi, D. L. Kaplan, K. Numata, Conformation and dynamics of soluble repetitive domain elucidates the initial β-sheet formation of spider silk. *Nat. Commun.* **9**, 2121 (2018).
16. G. Askarieh, M. Hedhammar, K. Nordling, A. Saenz, C. Casals, A. Rising, J. Johansson, S. D. Knight, Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **465**, 236–238 (2010).
17. J. H. Exler, D. Hümmerich, T. Scheibel, The amphiphilic properties of spider silks are important for spinning. *Angew. Chem. Int. Ed.* **46**, 3559–3562 (2007).
18. F. Hagn, L. Eisoldt, J. G. Hardy, C. Vendrely, M. Coles, T. Scheibel, H. Kessler, A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **465**, 239–242 (2010).
19. K. Jaudzems, G. Askarieh, M. Landreh, K. Nordling, M. Hedhammar, H. Jornvall, A. Rising, S. D. Knight, J. Johansson, pH-dependent dimerization of spider silk N-terminal domain requires relocation of a wedged tryptophan side chain. *J. Mol. Biol.* **422**, 477–487 (2012).
20. S. Rammensee, U. Slotta, T. Scheibel, A. R. Bausch, Assembly mechanism of recombinant spider silk proteins. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 6590–6595 (2008).
21. M. Andersson, Q. Jia, A. Abella, X.-Y. Lee, M. Landreh, P. Purhonen, H. Hebert, M. Tenje, C. V. Robinson, Q. Meng, G. R. Plaza, J. Johansson, A. Rising, Biomimetic spinning of artificial spider silk from a chimeric minispidroin. *Nat. Chem. Biol.* **13**, 262–264 (2017).
22. M. B. Hinman, R. V. Lewis, Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *J. Biol. Chem.* **267**, 19320–19324 (1992).
23. A. D. Malay, K. Arakawa, K. Numata, Analysis of repetitive amino acid motifs reveals the essential features of spider dragline silk proteins. *PLOS ONE* **12**, e0183397 (2017).
24. J. E. Jenkins, M. S. Creager, E. B. Butler, R. V. Lewis, J. L. Yarger, G. P. Holland, Solid-state NMR evidence for elastin-like beta-turn structure in spider dragline silk. *Chem. Commun.* **46**, 6714–6716 (2010).
25. A. K. Lancaster, A. Nutter-Upham, S. Lindquist, O. D. King, PLAAC: A web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* **30**, 2501–2502 (2014).
26. S. Alberti, A. Gladfelter, T. Mittag, Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 (2019).
27. Y. Shin, C. P. Brangwynne, Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).

28. F. G. Quiroz, A. Chilkoti, Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nat. Mater.* **14**, 1164–1171 (2015).
29. P. L. Babb, N. F. Lahens, S. M. Correa-Garhwal, D. N. Nicholson, E. J. Kim, J. B. Hogenesch, M. Kuntner, L. Higgins, C. Y. Hayashi, I. Agnarsson, B. F. Voight, The *Nephila clavipes* genome highlights the diversity of spider silk genes and their complex expression. *Nat. Genet.* **49**, 895–903 (2017).
30. N. A. Oktaviani, A. Matsugami, F. Hayashi, K. Numata, Ion effects on the conformation and dynamics of repetitive domains of a spider silk protein: Implications for solubility and β-sheet formation. *Chem. Commun.* **55**, 9761–9764 (2019).
31. P. Mohammadi, A. S. Aranko, L. Lemetti, Z. Cenev, Q. Zhou, S. Virtanen, C. P. Landowski, M. Penttilä, W. J. Fischer, W. Wagermaier, M. B. Linder, Phase transitions as intermediate steps in the formation of molecularly engineered protein fibers. *Commun. Biol.* **1**, 86 (2018).
32. A. S. Holehouse, R. V. Pappu, Collapse transitions of proteins and the interplay among backbone, sidechain, and solvent interactions. *Annu. Rev. Biophys.* **47**, 19–39 (2018).
33. G. L. Dignon, R. B. Best, J. Mittal, Biomolecular phase separation: From molecular driving forces to macroscopic properties. *Annu. Rev. Phys. Chem.* **71**, 53–75 (2020).
34. L. D. Muiznieks, S. Sharpe, R. Pomès, F. W. Keeley, Role of liquid–liquid phase separation in assembly of elastin and other extracellular matrix proteins. *J. Mol. Biol.* **430**, 4741–4753 (2018).
35. Y. Zhang, P. S. Cremer, Chemistry of Hofmeister anions and osmolytes. *Annu. Rev. Phys. Chem.* **61**, 63–83 (2010).
36. Y. Lin, S. L. Currie, M. K. Rosen, Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J. Biol. Chem.* **292**, 19110–19120 (2017).
37. A. C. Murthy, G. L. Dignon, Y. Kan, G. H. Zerze, S. H. Parekh, J. Mittal, N. L. Fawzi, Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **26**, 637–648 (2019).
38. B. Mészáros, G. Erdős, Z. Dosztányi, IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* **46**, W329–W337 (2018).
39. M. A. Rapsomaniki, P. Kotsantis, I.-E. Symeonidou, N.-N. Giakoumakis, S. Taraviras, Z. Lygerou, easyFRAP: An interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data. *Bioinformatics* **28**, 1800–1801 (2012).
40. D. N. Perkins, D. J. C. Pappin, D. M. Creasy, J. S. Cottrell, Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).

Acknowledgments: This work was supported by the Impulsing Paradigm Change through Disruptive Technologies Program (ImPACT), L-INSIGHT (Kyoto University), and the RIKEN Engineering Network. **Funding:** JST ERATO grant number JPMJER1602. **Author contributions:** A.D.M. and K.N. designed the research. T.S. performed the MS analysis. T.K. performed the rheological analysis. N.K. and K.A. analyzed and assigned the MS data. A.D.M. performed the other experiments. A.D.M. and K.N. prepared the manuscript draft, and all the authors confirmed the final manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 5 March 2020
Accepted 16 September 2020
Published 4 November 2020
10.1126/sciadv.abb6030

Citation: A. D. Malay, T. Suzuki, T. Katashima, N. Kono, K. Arakawa, K. Numata, Spider silk self-assembly via modular liquid-liquid phase separation and nanofibrillation. *Sci. Adv.* **6**, eabb6030 (2020).