

Spider Silk: Rapid, Bottom-Up Self-Assembly of MaSp1 into Hierarchically Structured Fibers Through Biomimetic Processing

Ali D. Malay,* Nur Alia Oktaviani, Jianming Chen, and Keiji Numata*

Spider dragline silk's exceptional mechanical performance, coupled with its benign production pathway, have inspired the design of diverse smart materials. Remarkably, relatively little is known about the self-assembly process of major ampullate spidroin 1 (MaSp1) – the main protein constituent of dragline fiber – during its rapid conversion from soluble precursor protein into hierarchically structured mature silk fibers. Herein, the biomimetic potential of MaSp1 is explored using a new recombinant platform with full domain representation. MaSp1 is found to undergo efficient liquid-liquid phase separation (LLPS) in response to kosmotropic ions, with notably higher propensity compared to MaSp2, and with MaSp1 exhibiting a more complex solubility profile relative to levels of sodium chloride. To further the understanding of silk spinning, the rapid assembly of mesoscale structures is monitored in real-time in response to ion and pH gradients, revealing a progression from LLPS droplets toward aligned hierarchical fibers. Furthermore, using this biomimetic system, insoluble macroscopic MaSp1 fibers can be readily generated, wherein the application of mechanical deformation triggers the emergence of β -sheet secondary structures. The insights gained from this study can have broader application in efforts to mimic the formation of biomaterials with spatially organized features across multiple length scales.

1. Introduction

There is an increasingly urgent need to develop new, greener technologies that will help shift away from an overreliance on nonrenewable or environmentally unsound components for material production and processing. Natural biomaterials, which employ robust, energy-efficient self-organization mechanisms that have been refined through deep evolutionary time, can be a particularly rich source of inspiration for envisioning alternative modes of production. A particularly exciting insight is that in many classes of biomaterials, including externally processed structural fibers, common design principles tend to be shared in common, including the use of stimuli-responsive biomolecular building blocks, storage of precursor molecules in a condensed and inert state, and employment of sensitive mechanisms to trigger rapid self-organization into hierarchically structured materials in response to environmental changes.^[1]

Spider dragline silk is a multicomponent protein-based fiber that has long been the subject of great interest, primarily because of its extraordinary tensile properties, which routinely surpass the performance of advanced synthetic materials.^[2] Equally remarkable is the natural spinning process that directs the conversion of the component soluble silk proteins (spidroins) into hierarchically organized insoluble fibers via a sophisticated self-assembly mechanism, all taking place under benign conditions and with great rapidity. In brief, dragline silk is composed mainly of major ampullate spidroins (MaSps, of which several subtypes exist: MaSp1, MaSp2, MaSp3, etc.), which are high molecular weight proteins with a modular architecture consisting of a long repetitive core region composed of short iterated amino acid motifs (which largely determine the material properties of the mature fiber), which is capped by conserved globular terminal domains that act as environmental sensors and structural switches that play critical roles in orchestrating the self-assembly process. During fiber formation, the MaSp precursor polypeptides stored within the major ampullate gland migrate through a spinning duct toward the exterior environment where they encounter simultaneously occurring changes in the chemical and physical conditions, including

A. D. Malay, N. A. Oktaviani, J. Chen^[+], K. Numata
Center for Sustainable Resource Science
RIKEN
2-1 Hirosawa, Wako, Saitama 351-0198, Japan
E-mail: a.malay@riken.jp; keiji.numata@riken.jp
K. Numata
Department of Material Chemistry
Graduate School of Engineering
Kyoto University
Katsura, Nishikyo-ku, Kyoto 615–8510, Japan

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adfm.202408175>

[+]Present address: Research Institute for Intelligent Wearable Systems and Research Centre of Textiles for Future Fashion, School of Fashion and Textiles, The Hong Kong Polytechnic University, Kowloon, Hong Kong

© 2024 The Author(s). Advanced Functional Materials published by Wiley-VCH GmbH. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](#) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1002/adfm.202408175

a pH gradient (acidification from pH 7.2 to pH 5.7 or lower), changes in ion composition (exchange from chaotropic to kosmotropic ion conditions), removal of bulk water, and shear and elongational flow effects brought about by changes in the internal duct geometry.^[3] These changes trigger precise biochemical responses in the individual MaSp domains, including the dimerization and partial loss of structure in the N- and C-terminal domains (NTD and CTD)^[3a,4] respectively, in response to acidification, or the conversion of the disordered core repetitive sequences into more aligned conformations punctuated by β -sheet nanocrystals characteristic of the mature fiber, in response to the effects of physical deformation.^[3c,5]

Recently, we showed that for MaSp2, a main driver of self-assembly is liquid-liquid phase separation (LLPS) that occurs in response to critical levels of phosphate,^[6] a kosmotropic anion that is enriched in the latter parts of the natural spinning apparatus.^[3b] Combined with acidification, LLPS could trigger the efficient formation of nanofibrillar MaSp2 networks that form the basis for hierarchically organized artificial silk fibers reminiscent of natural spider silk.^[6] It was suggested that LLPS could prime spidroins for assembly by promoting the alignment of aggregation-prone sequences prior to fiber formation.^[7]

Remarkably, while MaSp1 – considered the predominant and archetypal constituent of dragline silk^[2,8] – has been the subject of numerous studies, including as a constituent in biomimetic fiber spinning,^[9] its self-assembly behavior has so far been relatively little explored,^[10] especially at the mesoscale and at short time intervals that are most relevant to the emergence of hierarchical architectures in the native silk fiber. A reason for this relative lack of insight might be related to difficulties in producing recombinant MaSp1 bearing numerous tandem repeats in the natively folded state.

To address such issues, we developed a biomimetic platform based on MaSp1 from *Trichonephila clavipes* that encodes the full set of functional domains, including 6 tandem repeat elements and the terminal domains. We analyzed the solution structure of MaSp1 and mapped out the parameters governing its solubility, LLPS, and self-assembly into higher-order hierarchical structures in response to biomimetic gradients. Moreover, we provide the first visualization of the initial, rapid steps of phase separation and fibrillogenesis in response to the native-like chemical gradients. These insights were integrated toward the production of insoluble macroscopic MaSp1 fibers with aligned fibrillar organization and featuring deformation-inducible emergence of β -sheet conformations. The results from this study are expected to guide future efforts to produce more environmentally friendly supramolecular biomaterials with enhanced mechanical performance.

2. Results

2.1. Design of Recombinant MaSp1

We designed a new recombinant MaSp1 construct, termed N-R6-C, incorporating the three functional domains of MaSp1 from *T. clavipes*: the N-terminal domain (NTD or N), repetitive region (Rep or R6), and C-terminal domain (CTD or C) (Figure 1a). Rep consists of six tandem repeats of GGQGAGAAAAAGGAGQG-GYGGLGSQAGRGGL, representing a 33-residue consensus

deduced from the native sequence. It is useful to compare the Rep region of MaSp1 with that of the previously characterized MaSp2 construct (Figure 1b). Whereas both sequences feature alternating poly-Ala and Gly-rich regions, the latter consisting largely of GX and GGX motifs (including GGY), MaSp1 can be distinguished by the inclusion of the bulky hydrophobic Leu and the cationic Arg residue; in contrast, MaSp2 incorporates numerous Pro residues and the characteristic di-Gln (QQ) motif.^[8b] Consequently, the Rep regions of MaSp1 and MaSp2 produce different amphiphilicity profiles, with MaSp1 showing frequent alternations of hydrophilic and hydrophobic character, whereas MaSp2 displays a more uniform polar or hydrophilic profile within the Gly-rich regions (Figure 1b).

Analysis of the overall MaSp1 primary structure predicts a high level of intrinsic disorder in the Rep regions and the adjacent portion of the CTD (Figure 1c)^[11] consistent with previous reports.^[12] Furthermore, implementation of the FuzDrop method^[13] predicts a high probability for liquid-phase condensation (LLPS) for MaSp1 N-R6-C.

2.2. MaSp1 Behavior in Solution

MaSp1 N-R6-C was successfully expressed as a fusion downstream of the NusA protein, the latter being an established solubility enhancer.^[15] Under nonreducing conditions, the purified MaSp1 moiety migrated as a dimer with an effective molecular weight of 81.6 kDa, consistent with disulfide-linked dimerization via the CTD (Figure S1, Supporting Information)^[4c] An interesting observation was the crucial dependence of the MaSp1 solubility on background NaCl levels. For example, at 3.0 mg mL⁻¹ protein in 20 mM Tris-HCl, pH 7.5, NaCl levels ranging from 200–1000 mM was required for full solubility, with MaSp1 reversibly partitioning out of solution outside this range (Figure 2a). These NaCl-dependent solubility effects in MaSp1 differed markedly from observations with MaSp2, where the solubility was found to be largely independent of salt concentration.^[6]

Far-UV circular dichroism (CD) was used to probe the structure and stability of MaSp1 N-R6-C, as a function of pH (Figure 2b). CD spectra of the solutions prepared at either a pH of 7.0 or a pH of 5.0 were characteristic of α -helices at 20 °C, as expected from the presence of the two largely α -helical terminal domains.^[4a,c] The sample at pH 5.0, however, showed lower spectral amplitudes compared to pH 7.0 at this temperature, likely reflecting unfolding of the CTD upon acidification.^[3a] Conversely, the secondary structure of MaSp1 was more sensitive to thermal perturbation at pH 7.0, with a drastic loss in α -helical signal between 40 and 50 °C, while at pH 5.0 the thermally-induced secondary structure loss was less pronounced and more gradual; this result likely reflects the gain in stability of NTD upon dimerization at acidic conditions.^[3a] Overall, the results are consistent with the complementary effects of pH on the stabilities of the NTD and CTD.^[3a,16]

¹H-¹⁵N HSQC NMR was performed to gain further insight into the structure of MaSp1, including the repetitive region (Figure 2c). The spectra for N-R6-C (gray peaks) showed a wide overall amide proton dispersion (>1 ppm), indicating the presence of well-ordered structure, as contributed by the terminal

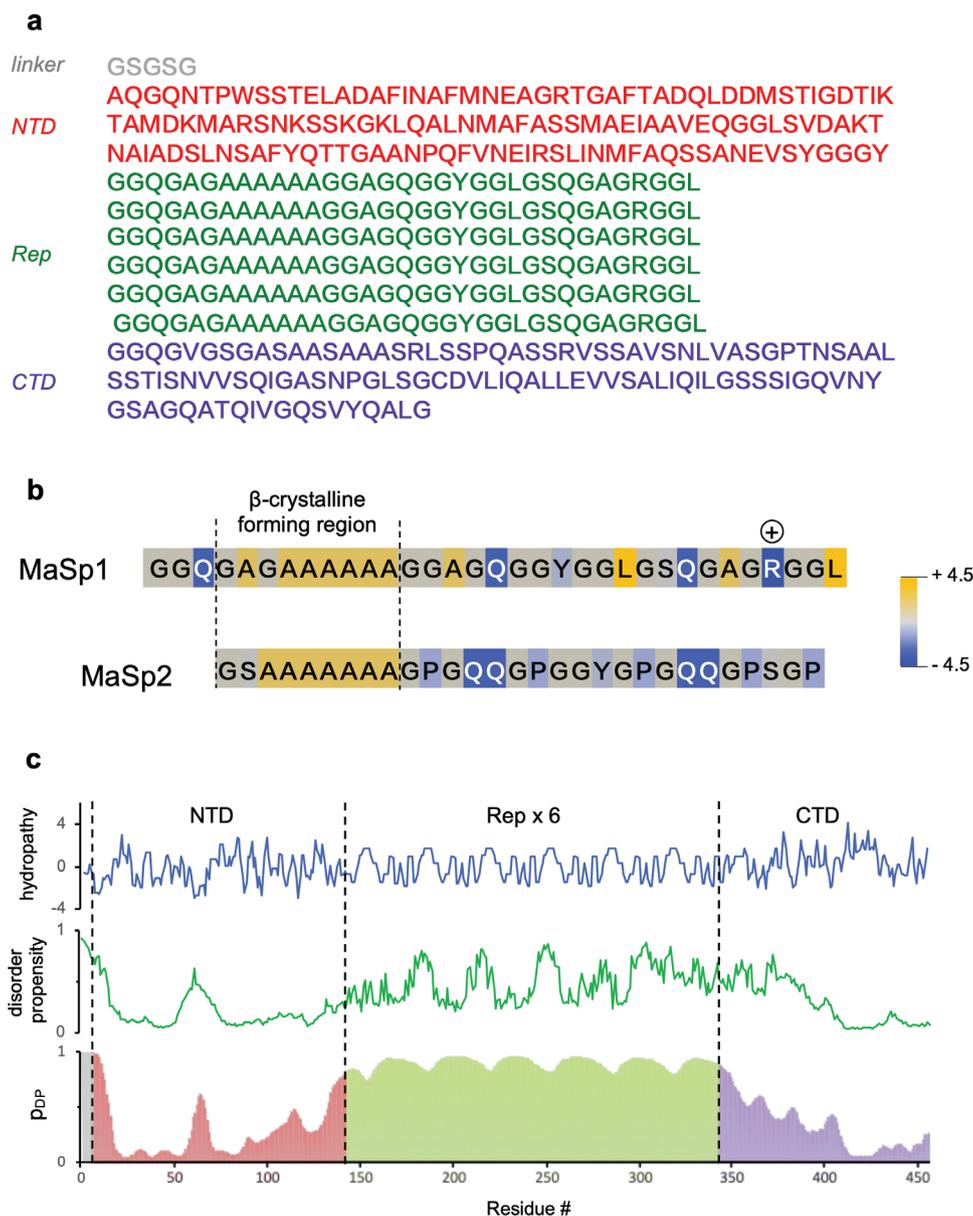


Figure 1. Design of the recombinant MaSp1 platform. a) Amino acid sequence of recombinant MaSp1 construct N-R6-C used in the study. The sequence is derived from native MaSp1 from *T. clavipes* and features 6 consensus tandem repeats of 33 residues each and flanked by N- and C-terminal domains. b) Comparison of the MaSp1 and MaSp2 consensus repeat sequences. The background color corresponds to the hydrophobicity of each residue according to Kyte–Doolittle,^[14] while the cationic charge on the arginine (R) in MaSp1 is also indicated. MaSp1 features a larger proportion of hydrophobic residues that punctuate the glycine-rich region, [leucine (L) and alanine (A)], whereas proline (P) and di-glutamine (QQ) residues feature prominently in MaSp2, to present a more uniformly hydrophilic profile outside the poly-Ala region. (c) Analysis of the MaSp1 N-R6-C sequence via different metrics, with domain boundaries indicated by dotted lines. *Top panel*, hydropathy plot (Kyte–Doolittle, window = 3); *middle panel*, structural disorder prediction via fDPnn^[11]; *bottom panel*, LLPS propensity via FuzDrop.^[13]

domains. Interestingly, despite the high molecular weight of N-R6-C (81.6 kDa, considering dimer formation via the CTD), relatively narrow linewidth signals were observed, suggesting that each domain maintains independent molecular motion, as suggested.^[10c] The terminal domains of N-R6-C could maintain their native folded structure, as observed from the good superimposition of the signals with the respective spectra from the isolated terminal domains (Figure S2, Supporting Information). However, the CTD signals were considerably broader than those

for NTD, possibly due to the larger structure of the CTD dimer or due to its relatively low solubility in aqueous buffer.^[3a,17] On the other hand, backbone signals attributed to the repetitive regions indicate highly dynamic conformations, as indicated by the narrow amide proton dispersion (Figure 2c, inside the red dotted ellipses), consistent with previous reports on dragline spidroins.^[10e,12b,18] Overlaying the N-R6-C spectra with data from MaSp1 4-tandem repeat R4 (yellow peaks) provides an approximate fingerprint of the different repeat domain assignments,

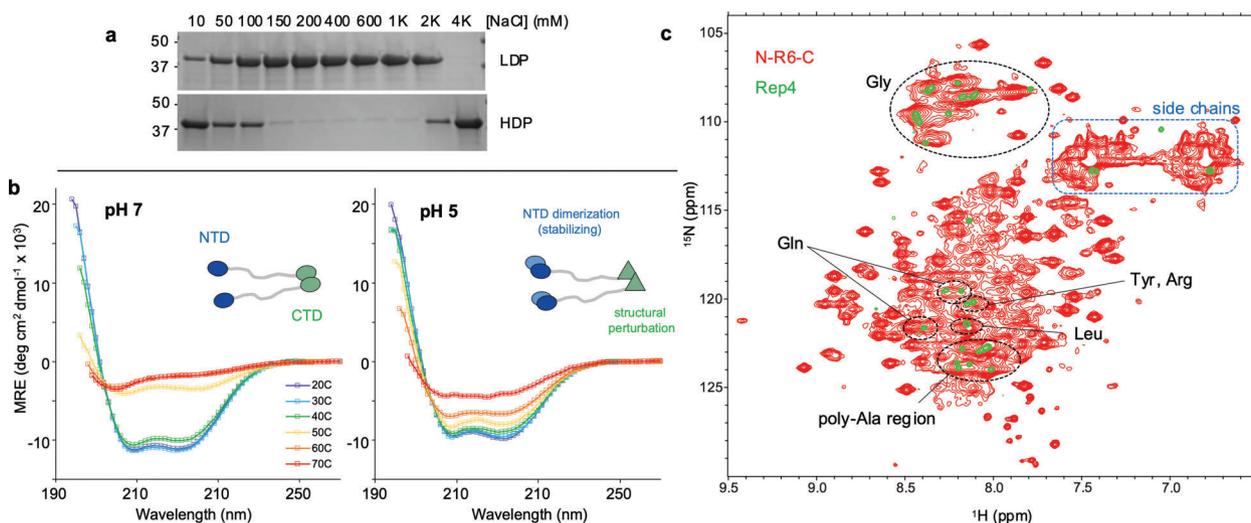


Figure 2. MaSp1 structure and behavior in the prefibrillar solution state. a) Dependence of MaSp1 solubility on NaCl concentration. MaSp1 N-R6-C (at a final concentration of 3.0 mg mL^{-1}) maintains full solubility at NaCl concentrations between $\approx 150 \text{ mM}$ and 1.0 M ; outside of this range, MaSp1 reversibly partitions out of solution, to accumulate in a high-density phase (HDP), with a concomitant depletion from the aqueous phase or LDP. b) Far-UV circular dichroism of dilute MaSp1 N-R6-C at pH 7.0 and at pH 5.0, taken as a function of temperature at 10-degree intervals from 20–70 °C. The terminal domains contribute α -helical signals, with apparent trade-offs in structural stability between the NTD and CTD under neutral and acidic conditions. c) 2D NMR ^1H - ^{15}N HSQC analysis of MaSp1 N-R6-C (red peaks) overlaid with spectra from the isolated MaSp1 tandem repeat (green). The NMR spectra are consistent with generalized disorder in the repetitive regions, coupled with well-structured terminal domains (see also Figure S2, Supporting Information).

including clustered signals derived from the various glycine residues, as well as the alanine residues comprising the polyalanine stretches.

2.3. Phase Separation Triggered by Ion Perturbations

Consistent with our predictions, recombinant MaSp1 exhibited a high propensity for LLPS, under conditions approximating the physicochemical changes encountered in the spider's natural spinning apparatus. MaSp1 in solution exposed to potassium phosphate (KP_i) at pH 7.0 above a certain concentration threshold immediately became turbid. Microscopically, this turbidity corresponded to a two-phase state characterized by numerous spherical liquid droplets of protein that could readily undergo fusion and surface wetting (i.e., LLPS) (Figure 3; Movie S1, Supporting Information). Eventually, the protein droplets settled via density to form merged condensate structures on the glass surface. We mapped the MaSp1 LLPS propensity under isothermal conditions with respect to the molar concentrations of protein and of the KP_i trigger, to thereby delineate the boundaries of the 1-phase and 2-phase regions (Figure 3a). Strikingly, MaSp1 was found to have a much greater LLPS propensity than MaSp2 with the equivalent N-R6-C architecture under identical conditions^[6] (see the red dotted curve in Figure 3a). Illustrative of this discrepancy, under 0.5 M KP_i conditions, MaSp1 N-R6-C displayed an LLPS concentration threshold (i.e., apparent saturation concentration) of $\approx 8 \text{ }\mu\text{M}$, whereas for MaSp2 N-R6-C the boundary for spontaneous phase separation was $\approx 80 \text{ }\mu\text{M}$, otherwise maintaining a one-phase solution state at lower protein levels.

To further elucidate the requirements for MaSp1 phase separation, we screened an array of ion compounds (Table S1, Supporting Information), which showed that LLPS could be triggered under a wide range of conditions. Interestingly, the choice of anion

species was found to largely dictate the LLPS propensity, while the cation partner seemed to have little effect, at least under the conditions tested. In summary, we observed that phosphate, sulfate, citrate, maleate and fluoride anions could efficiently induce LLPS; incidentally, such anions are classified as kosmotropic in character, such as defined by the Hofmeister series,^[19] suggesting that LLPS of MaSp1 at least partially arose in response to perturbations to the water–water and water–solute interactions. Notably, varying the ion species could create different concentration boundaries for LLPS, as seen for instance in the response of MaSp1 to trisodium citrate (Figure 3a, green curve), which produced lower concentration thresholds for phase separation (i.e., higher propensity) compared to the effect of KP_i under the same conditions.

As discussed, recombinant MaSp1 had a narrow solubility window with respect to the NaCl concentration, outside of which it partitioned from solution (see Figure 2a), and we confirmed that this phenomenon likewise proceeded via LLPS, at both the low- and high-salt extremes (Figure S3, Supporting Information). While the shapes and dimensions of the condensates formed under the high-NaCl conditions resembled those under KP_i -induced LLPS conditions, the droplets formed under the NaCl-deficient conditions appeared larger and had a more spherical morphologies, suggesting a greater degree of fluidity.

2.4. Effect of pH Gradient

During spider silk assembly, the response of spidroins to decreasing pH is a well-characterized phenomenon, enabling the dimerization of NTDs to produce interconnected polypeptide chains. To study the response of recombinant MaSp1 to acidification, we first confirmed the NTD-mediated dimerization of dilute N-R6-C

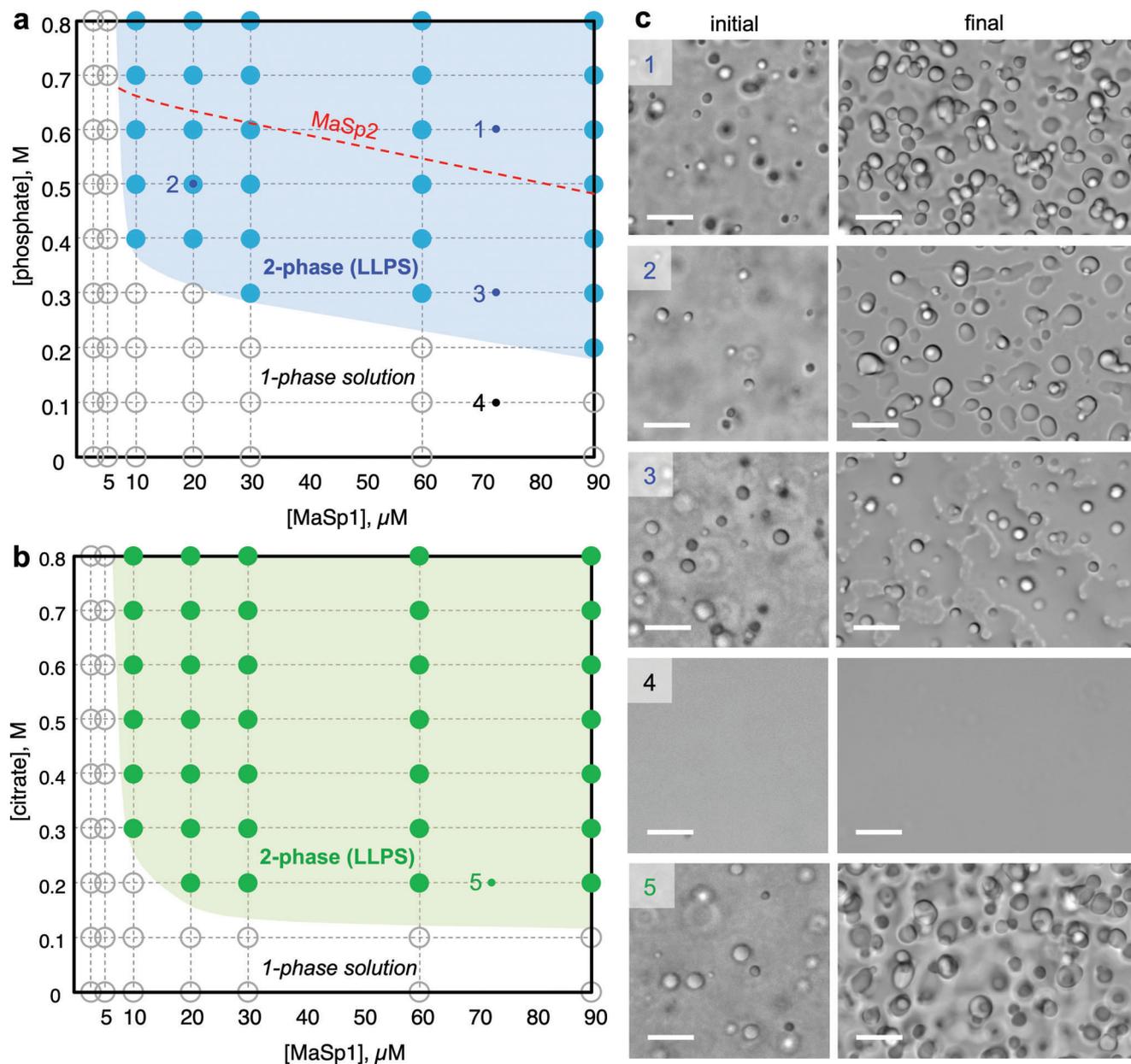


Figure 3. MaSp1 exhibits a high propensity for liquid–liquid phase separation. a,b) Mapping the LLPS propensity of MaSp1 N-R6-C as a function of the protein concentration (x-axis), and anion concentration (y-axis). Concentrations correspond to final the values after 1:1 mixing of protein and the chemical trigger, namely (a) KP_i and (b) tri-sodium citrate. The samples were monitored via microscopy at each condition corresponding to the intersections between the vertical and horizontal dotted lines, and scored for the presence (filled circles) or absence (open circles) of liquid protein droplets, from which best-fit estimate curves (shaded regions) were constructed. The phase separation map of MaSp2 N-R6-C against KP_i pH 7.0 is also shown in (a) as a red dotted line for comparison, which shows a much lower tendency for LLPS. c) Micrographs showing the representative MaSp1 condensate morphology obtained under the indicated conditions on the maps (numbered labels), taken immediately upon reaction initiation (“initial”) and in the final deposited state, 30 min after initiation. Scale bars = 10 μm .

via the fluorescence shift method.^[4a,20] **Figure 4a** shows that the dimerization response is strongly modulated by the background levels of NaCl, with midpoints of dimerization estimated at pH values of ≈ 6.25 and pH 5.75 under 25 mM and 300 mM NaCl, respectively, which was in agreement with previous findings obtained using NTD in isolation.^[4a,b,21]

We proceeded to monitor the mesoscale self-assembly of recombinant MaSp1 as a function of pH, under LLPS conditions,

which revealed profound changes in condensate morphology (Figure 4b). Upon mixing MaSp1 and KP_i in the pH range of 8.0–6.5, dynamic liquid-like LLPS droplets were produced, which merged continuously and deposited onto the glass slide surface on the timescale of a few minutes. Strikingly, at pH values less than or equal to 6.0, the MaSp1 condensates exhibited highly altered behaviors, which were characterized by rapid self-assembly into protein fibril networks with submicron features.

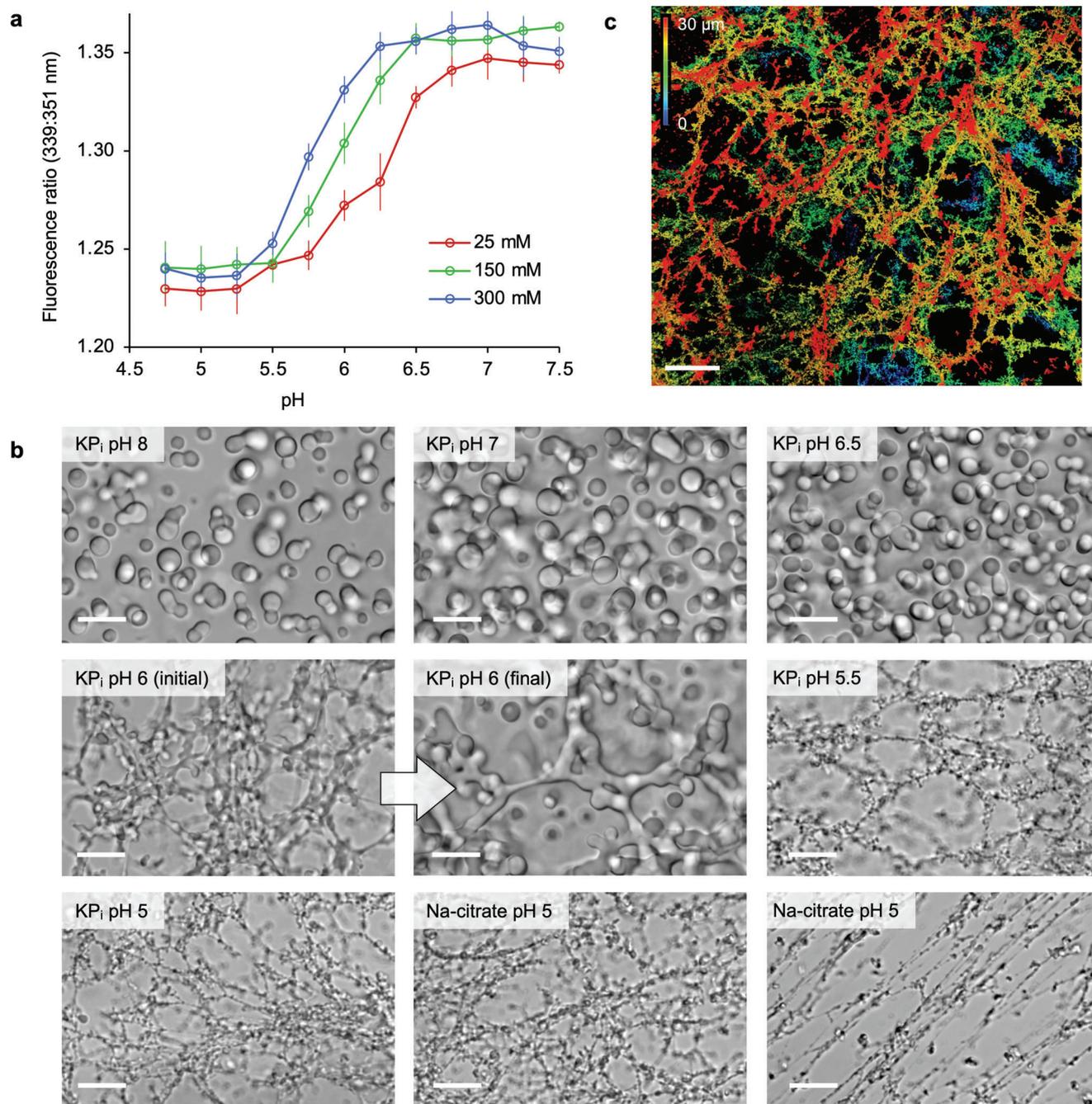


Figure 4. Acidification drives self-assembly of MaSp1 into submicron fibrils. a) Monitoring pH-directed dimerization of MaSp1 N-R6-C via the NTD via fluorescence emission shift assays. Curves at three different NaCl ambient concentrations (25, 150, 300 mM NaCl) are shown. b) Confocal light micrographs showing mesoscale condensate structures of MaSp1 that formed spontaneously upon mixing equal volumes of 20 mg mL⁻¹ of N-R6-C solution and 600 mM of KP_i at the indicated pH values (with a background of 500 mM NaCl). The morphology of the MaSp1 condensates is highly sensitive to the pH conditions, with acidification leading to self-assembly of submicron fibril networks. Also shown are structures obtained using sodium citrate buffer (pH 5.0), which produces identical results to KP_i pH 5.0. Scale bars = 10 μm. (c) 3D reconstruction of self-assembled MaSp1 fibrils showing propagation of the network across three spatial dimensions. The image is colored by z-depth with a range of 0–30 μm. Scale bar = 50 μm.

At the transitional pH of 6.0, mixing of MaSp1 and KP_i initially led to the rapid formation of fine structural networks; interestingly, within a matter of minutes these structures underwent gradual swelling, resulting in lobed structures with a “soft” appearance. Under more acidic conditions (pH 5.5–5.0), the imme-

diately assembly of submicron fibril networks encompassing the entire sample field was observed, which did not undergo any further transitions. Spontaneous alignment of the fibrils in parallel orientation was often observed, likely driven by capillary flow of the sample on the slide following placement of the coverslip. We

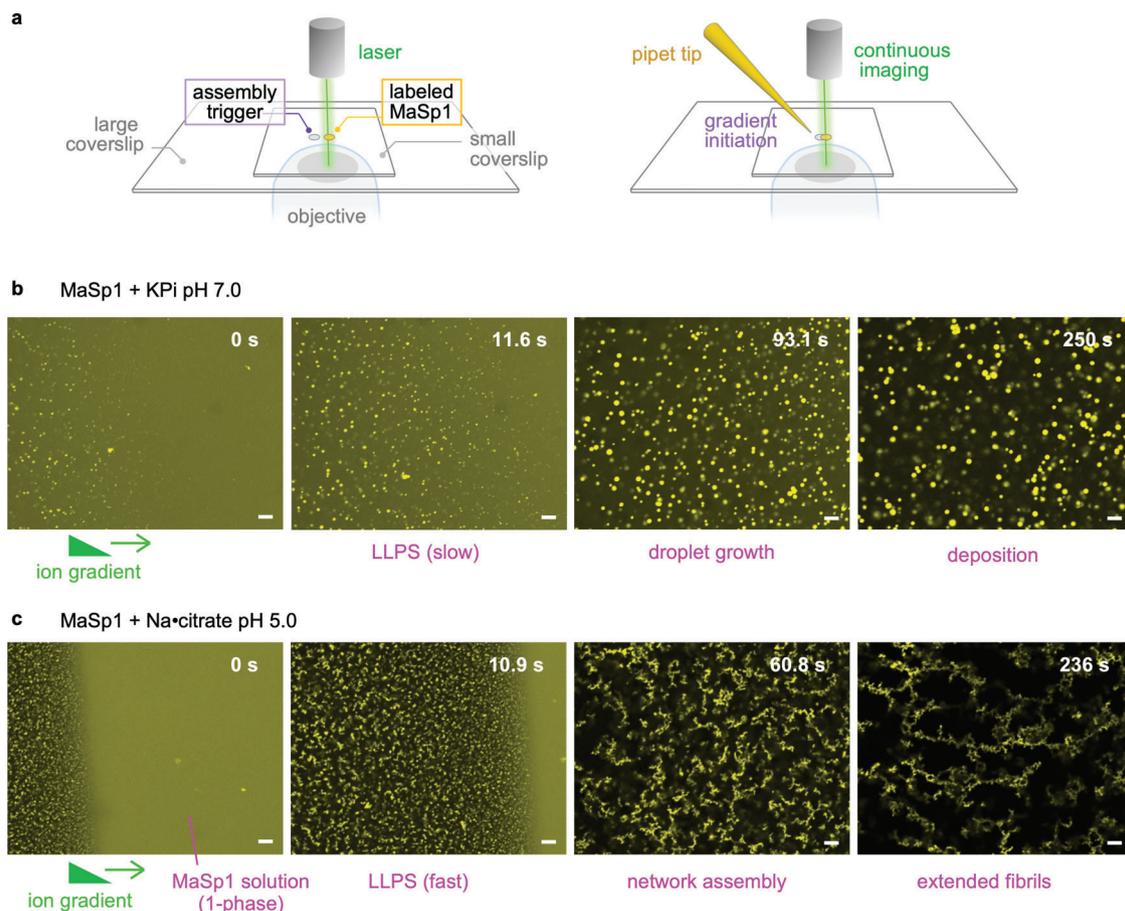


Figure 5. Probing real-time self-assembly of MaSp1 in response to biomimetic gradients. a) Schematic illustration of the method for observing MaSp1 bottom-up self-assembly in response to ion and pH gradients. Two drops were placed adjacent to each other between glass coverslips, one containing dilute, fluorescently labeled N-R6-C and the other containing the chemical trigger solution (left). Application of pressure using a micropipette tip was used to induce the two drops to merge, thus initiating an ion/pH gradient to migrate across the protein sample, with continuous imaging using a confocal fluorescence microscope. b,c) Monitoring the rapid changes of MaSp1 condensate morphology upon exposure to the mobile ion/pH gradients. The zero timepoint arbitrarily denotes the timing of the first stable image following gradient initiation. Scale bars = 10 μm . Panel (b) shows the effects of a KP_i (pH 7.0) gradient, while panel (c) shows the effects of a sodium citrate (pH 5.0) gradient. In both cases, the chemical gradient moves from left to right along the frames, which were extracted from Movies S3 and S4 (Supporting Information). Scale bar = 10 μm .

also tested sodium citrate, pH 5.0, as the trigger for assembly, which produced MaSp1 fibril networks that were indistinguishable from those formed with KP_i at pH 5 (by taking this phenomenon into account, we substituted citrate-based systems for phosphate in subsequent low-pH experiments, to exploit the favorable buffering capacity of citrate compared to phosphate under acidic conditions). Figure 4c shows a 3D reconstruction of the submicron fibril network structures computed from a series of stacked images, corresponding to a 30 μm section along the z-axis, which shows the fibrils making numerous interconnections along the three spatial dimensions (see also Movie S2, Supporting Information).

2.5. Real-Time Observation of Rapid MaSp1 Self-Assembly

A major challenge in studying the self-assembly of MaSp1 was that the initial events occurred very rapidly upon component mixing, making their observation inaccessible through conventional

microscopy. To address this issue, we developed a method for visualizing rapid mesoscale transitions in recombinant MaSp1 that occurred during the early stages of LLPS and fibrillogenesis (Figure 5). The method is shown in Figure 5a, and involved tracking the assembly of fluorescently-conjugated MaSp1 in real-time upon exposure to ion/pH gradients in a semi-controlled environment (details are provided in the Experimental Section). Typical results of the real-time monitoring performed at pH values of 7 and 5 are presented in Movies S3 and S4 (Supporting Information), respectively, with magnified details shown in Figure 5b,c.

Prior to exposure to the chemical trigger, uniform fluorescence was observed in the sample, corresponding to a homogeneous solution of labeled MaSp1. In Figure 5b and Movie S3 (Supporting Information), as the gradient of KP_i (pH 7.0) was initiated (sweep left to right), MaSp1 LLPS was induced upon contact, as shown by the appearance of bright yellow spots that coalesced into larger structures. This process was accompanied by the concomitant and gradual loss of fluorescence from the surrounding medium, as MaSp1 was progressively depleted from the bulk

solution, with the entire process taking over one minute. In contrast, at a pH of 5.0, MaSp1 phase separation proceeded nearly instantaneously as the ion front encountered the protein solution (Figure 5c; Movie S4, Supporting Information), seen as a sharp boundary migrating from left to right; in the wake of this boundary, there was an almost immediate depletion of fluorescence in the aqueous phase and its concomitant accumulation in numerous condensate bodies. The condensates formed at a pH of 5.0 appeared to be sticky and subsequently agglutinated into elongated structures that grew into a mesh-like network, culminating in extended fibrils that were predominantly oriented along the direction of sample flow.

Overall, our observations revealed major differences in MaSp1 phase separation behavior under the high and low pH regimes, with the pH 5.0 results suggesting higher rates of nucleation of condensate formation and suggesting an intriguing link between LLPS and self-assembly of hierarchically organized fibrillar structures.

2.6. Biomimetic Production of Hierarchically Organized MaSp1 Fibers

Based on the insights established in the previous sections, we aimed to generate macroscopic recombinant MaSp1 fibers with hierarchically organized substructures, reflecting the native dragline fiber organization (Figure 6). We focused on three main parameters, namely ion-dependent phase separation, acidification, and tensile deformation to mimic the pultrusion process encountered by the nascent silk fibers during the natural fiber spinning. Briefly, a highly concentrated solution of MaSp1 N-R-6-C (120–150 mg mL⁻¹ “silk dope”) was processed by successive exposure to 1.0 M citrate-phosphate buffer (CPB) at pH 7.0, which leads to LLPS, followed by 1.0 M of CPB at pH 5.5, which induced fibrillar self-assembly via the NTD. The viscous condensate mass was then subjected to unidirectional deformation through manual drawing using forceps (in buffer, followed by in air) to result in high-aspect-ratio solid fibers with a diameter of 10–15 μm. The morphological changes at the microscopic level are shown in Figure 6a.

Analyses of the biomimetic MaSp1 fibers by scanning electron microscopy (Figure 6b) and by confocal optical microscopy in water (Figure 6c) revealed the presence of a true hierarchical architecture, consisting of densely packed nanoscale fibrils oriented parallel to the longitudinal axis, including within the fiber core. The latter experiment furthermore demonstrated that the resulting MaSp1 biomimetic fibers were insoluble, and could retain their structural integrity upon immersion in water or in physiological buffer conditions, although a moderate swelling (lateral expansion) could be observed.

Confocal Raman spectroscopy was used to evaluate changes in the secondary structure of MaSp1 during the different steps of biomimetic fiber assembly (Figure 6d). The results clearly revealed the evolution of β -sheet structures within the fiber, which was triggered in response to the mechanical deformation, as demonstrated by the appearance of signature peaks in the amide I (≈ 1664 cm⁻¹) and amide III (≈ 1240 cm⁻¹) regions.^[3c] Conversely, samples that had not been subjected to mechanical perturbation produced peaks consistent with α -helical conforma-

tions (amide I peak at ≈ 1653 cm⁻¹), as expected from the high relative molecular weight of the predominantly α -helical terminal domains in N-R6-C.

3. Discussion

In this study we created a novel recombinant platform for MaSp1 that allowed us to interrogate the self-organization behavior of spider silk under controlled conditions that approximate the physicochemical changes in the spider's native spinning apparatus.

One interesting finding was that MaSp1 exhibited a much greater sensitivity to kosmotropic anions (such as phosphate) than MaSp2, as evidenced by the lower concentration thresholds for LLPS induction. This finding was initially surprising, given that the two recombinant spidroin constructs had identical domain architectures (N-R6-C) and similar molecular weights. However, analysis of the repetitive domain sequences of the two spidroins suggested some clues to explain these results. Based on studies exploring the molecular grammar governing LLPS in various biological systems, it is becoming clear that cation- π interactions, which typically involve arginine and tyrosine residues, respectively, provide potent drivers of biomolecular condensation, especially in systems featuring multivalent, repetitive sequences.^[22] Notably, MaSp1 has one residue each of Tyr and Arg in its consensus repeat sequence, whereas the MaSp2 repeat contains Tyr but no Arg residues. It would be interesting to test this and other sequence-function hypotheses in future studies.

We found that MaSp1 presented a rather complex solubility behavior with respect to ambient NaCl levels, such that outside a certain range of background salt concentrations, MaSp1 readily unmixed from solution in the form of LLPS droplets, in a reversible manner. In contrast, no such NaCl-dependent boundaries were observed with recombinant MaSp2, which apparently could maintain full solubility under similar conditions, including under conditions without salt. This unexpected behavior of MaSp1 recalled the so-called re-entrant phase transitions described in some intrinsically disordered proteins,^[23] whereby changes in a single parameter (in this case, NaCl concentration) could produce multiple phase transitions between single and two-phase states. These solubility requirements could suggest additional factors to consider when designing biomimetic spinning platforms that incorporate MaSp1, especially in the context of multicomponent systems. Additionally, the variability in phase separation behavior between MaSp1 and MaSp2, and perhaps other MaSp subtypes as well, raises the intriguing possibility that within the silk gland the different subtypes could form quasi-discrete, preorganized ensembles or clusters prior to fiber assembly, which could account for the nonhomogeneous distribution of the subtypes in the mature fiber that have been reported.^[24]

Some of the more intriguing results of this study came from the real-time observation of MaSp1 self-assembly at the mesoscale using fluorescence microscopy. To our knowledge, visualization of the complete assembly process from soluble spidroin state into hierarchically organized fibrils has not been reported previously, at least at the short time scales and under the biomimetic conditions explored in this study. In particular,

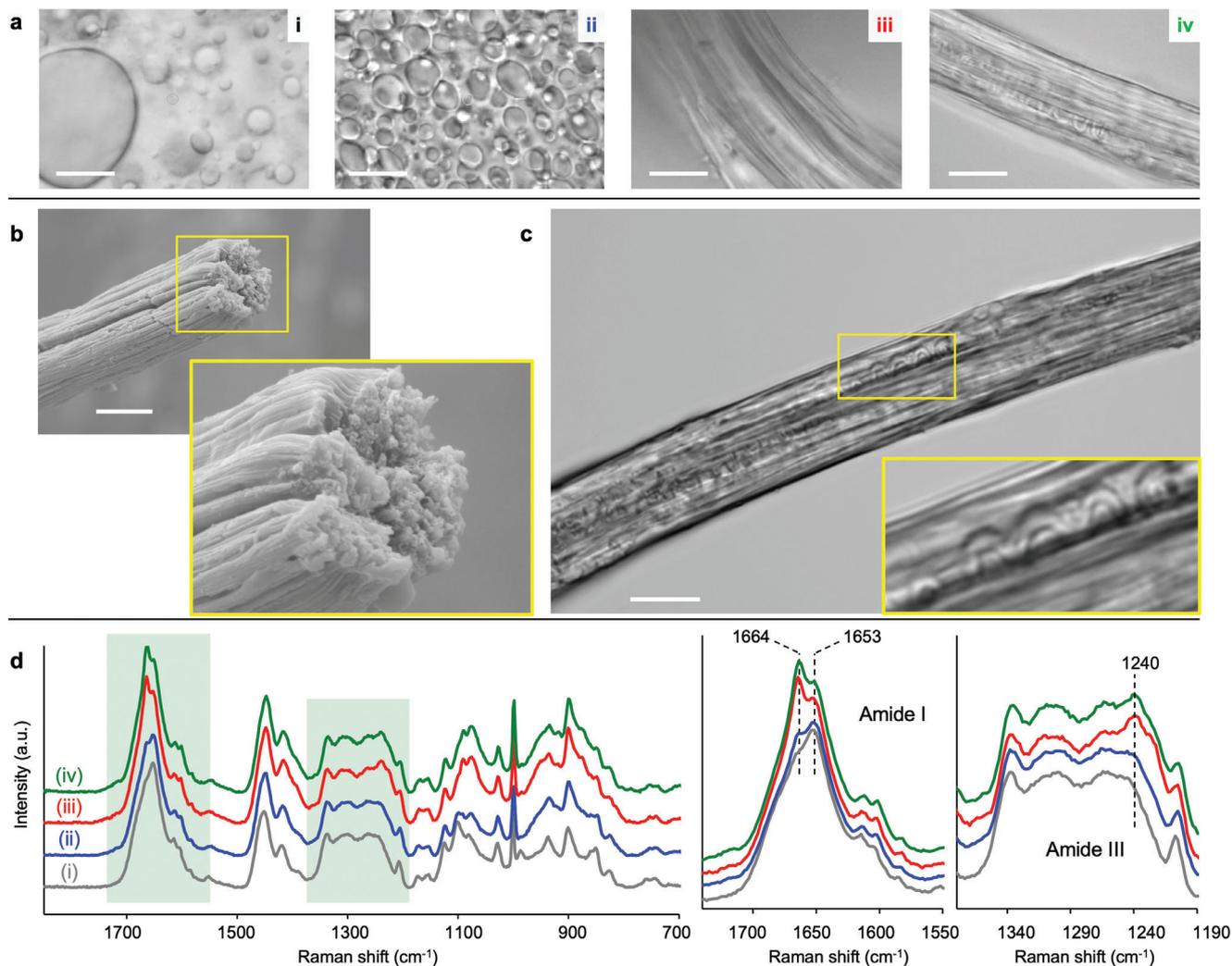


Figure 6. MaSp1 biomimetic fibers exhibit hierarchical structures and deformation-induced crystallization. a) Microscopic analysis of concentrated MaSp1 N-R6-C (120 mg mL^{-1}) at different stages of fiber assembly: (i) in CPB pH 7.0, showing dense droplet formation; (ii) in CPB pH 5.5, showing viscous protein condensate with trapped aqueous droplets; (iii) after subjection to mechanical deformation within the CPB pH 5.5 buffer, showing a dense mass with numerous parallel striations; (iv) upon further uniaxial deformation of the protein condensate in air, and subsequent replacement in CPB. Scale bars = $10 \mu\text{m}$. b) SEM image of MaSp1 N-R6-C fiber fracture surface showing internal submicron fibril bundle arrangement, with the inset showing clearer details. Scale bar = $10 \mu\text{m}$. c) Confocal microscopy image of MaSp1 fiber immersed in 1.0 M CPB reveals an internal organization of extended fibrils oriented along the longitudinal axis of the fiber. The inset highlights a single fibril within the fiber which apparently sprung back from an extended state. Scale bars = $10 \mu\text{m}$. d) Confocal Raman spectroscopy was used to evaluate secondary structural changes of the MaSp1 condensates at the different stages of assembly represented by samples (i)-(iv). The individual spectra are staggered on the y-axis to facilitate viewing. Two highlighted areas representing the amide I and amide III regions are expanded on the right-hand panels, and show the emergence of β -sheet conformations in samples (iii) and (iv) in response to mechanical deformation; these are indicated by clear peaks at ≈ 1664 and 1240 cm^{-1} . Shown are representative spectra from at least five experiments.

the apparent differences in the rates of nucleation of LLPS that were observed relative to pH conditions could help elucidate the self-assembly mechanism that leads to the rapid spinning of hierarchically organized spider silk observed in nature. The development of experimental methods that enable greater spatiotemporal resolutions will likely yield significant advances in this regard.

The study also demonstrated that the recombinant MaSp1 construct could be efficiently converted from disordered protein condensates into solid, water-insoluble macroscopic fibers with an internal hierarchical architecture consisting of aligned nanofibrillar bundles through the application of directional ten-

sion. Crucially, deformation applied to the fibrillar condensate could also trigger the emergence of β -sheet conformations, without the need for prolonged postprocessing, such as reported for the 2-repeat MaSp1 constructs;^[25] this difference could perhaps be attributed to the higher number of tandem repeats in the N-R6-C construct (6 tandem repeats) used in this study. Overall, our results are consistent with the role of deformational stress, as attributed to the pultrusion process, in inducing the emergence of β -sheet nanocrystals within the fiber matrix that are considered to be a prerequisite for achieving the characteristic high tensile strength of the material.^[3c,26]

4. Conclusion

The aim of the present study is to lay the groundwork for facilitating the inducible biomimetic self-assembly of MaSp1 at multiple length scales, from a largely disordered, soluble state to a state of hierarchically organized macroscopic fibers. In future research, an important consideration that deserves further attention concerns the internal geometry of the fiber spinning system, as indeed parameters such as shear and elongational flow are known to exert profound effects on spider silk self-assembly. In this regard, studies on MaSp1 bottom-up assembly can greatly benefit from the use of microfluidic approaches for the controlled delivery of chemical triggers and for the modulation of rheological parameters, as we recently explored in the context of recombinant MaSp2.^[27] Beyond the spider silk field, we hope that our current study will inspire further investigations into the mechanisms that govern the sequence-encoded self-assembly of other complex supramolecular biomaterials.

5. Experimental Section

Plasmid Construction: The *T. clavipes* MaSp1 gene construct N-R6-C was purchased (GenScript) and subcloned into a pET43.1 vector (Novagen) for expression downstream of the NusA protein moiety. Subcloning of the MaSp1 gene was carried out by PCR amplification of the insert from the pET15b construct using KOD-Plus-Neo polymerase and primers capped with SacI and XhoI restriction sites on the upstream and downstream primers, respectively. The PCR product was purified from 1% agarose gel using a Qiagen kit, followed by restriction digestion using SacI and XhoI (Takara) and purification of the digested product (Qiagen). The pET43.1 plasmid was digested with the same restriction enzymes and purified on a 1% agarose gel. DNA ligation was performed using Ligation high ver.2 (Toyobo) under standard conditions, followed by transformation into *DH5 α* . Transformants were screened by colony PCR followed by DNA sequencing. The original vector sequence was further modified to remove the enterokinase cleavage site and thus bring the thrombin cleavage site in proximity with the MaSp1 coding sequence. Plasmid modification was carried out via standard PCR-based site-directed mutagenesis, using KOD-Plus-Neo DNA polymerase (Toyobo), DpnI digestion (Takara), and transformation into *Escherichia coli* strain *DH5 α* . Transformants were screened by colony PCR using GoTaq polymerase (Promega) and T7 primers, and confirmed by DNA sequencing using BigDye Terminator v3.1 cycle sequencing (Thermo Fisher Scientific). The synthetic gene containing isolated MaSp1 4-repeat sequence R4 (ThermoFisher) was subcloned into pET15b vector and transformed into *E. coli* *DH5 α* strain.

Expression of Recombinant MaSp1 Protein: The engineered plasmid was transformed into *E. coli* BLR(DE3) strain and inoculated into 100 mL of LB medium (BD) with 100 $\mu\text{g mL}^{-1}$ ampicillin (Sigma) and incubated overnight at 37 °C with shaking at 180 rpm as a preculture, which was then used to inoculate 2 L of main culture in baffled culture flasks and shaken at 37 °C until OD₆₀₀ was ≈ 0.8 –1. Temperature was lowered to 20 °C and after 30 min additional shaking, isopropyl- β -D-1-thiogalactopyranoside (Wako) was added to 0.4 mM final concentration to initiate expression of the NusA-MaSp1 fusion protein, followed by overnight shaking. Cells were harvested by centrifugation, resuspended in buffer A (20 mM Tris-HCl pH 7.5, 500 mM NaCl and 20 mM imidazole), and stored at –80 °C in 40 mL aliquots (corresponding to 2 L cell culture volume), until purification.

For production of labeled protein for NMR, singly labeled (¹⁵N) MaSp1 N-R6-C and (¹⁵N) CTD were prepared by growing BL21(DE3) in M9 minimal medium containing unlabeled D-glucose (4 g L⁻¹ culture, Wako) and ¹⁵N-ammonium chloride (1 g L⁻¹ culture, Cambridge Isotope Lab). For doubly labeled samples, (¹³C, ¹⁵N) MaSp1 NTD and (¹³C, ¹⁵N) R4 (four tandem repeats), proteins were prepared by growing BL21(DE3) in M9 minimal medium containing ¹³C-glucose (2 g L⁻¹, Cambridge Isotope

Lab) and ¹⁵N-ammonium chloride (1 g L⁻¹, Cambridge Isotope Lab). Cultures were initially grown in 5 mL LB medium with 100 $\mu\text{g mL}^{-1}$ ampicillin, which was used to inoculate a 200 mL preculture of M9 minimal medium containing 100 $\mu\text{g mL}^{-1}$ ampicillin and grown overnight with shaking at 160 rpm, 37 °C, and then transferred into 2 L of main culture in the same medium. The cells were grown until OD₆₀₀ ≈ 1 , upon which expression was induced with 0.4 mM IPTG at 20 °C followed by overnight shaking at 160 rpm.

Protein Purification: For MaSp1 N-R6-C, frozen cell slurry (40 mL) was thawed in a water bath at room temperature, to which was added 40 mg of hen egg lysozyme (Wako), 500 units of TurboNuclease (Accelagen), 4.5 mL of 10% Triton X-100 (to 1% final concentration) and supplemented with cOmplete EDTA-free protease inhibitor cocktail (Roche). Cell lysis was allowed to proceed for 1 h at 4 °C with stirring, followed by centrifugation at 8000 rpm for 30 min at 4 °C, and the supernatant fraction transferred to fresh 50 mL tubes and centrifuged again for 30 min. The cell pellet was further treated with BugBuster cell lysis reagent (Merck) to maximize protein extraction, and the supernatant fractions from the two lysis steps were combined. The clarified supernatant fractions were loaded onto Ni-NTA HisTrap column (2.5 mL columns in tandem; Cytiva) and washed extensively at 4 °C with buffer A via AKTA Explorer (GE Healthcare) until baseline levels of absorbance were reached. Bound protein was eluted using buffer B (20 mM Tris-HCl pH 7.5, 500 mM NaCl, and 250 mM imidazole). The eluted fraction was concentrated and buffer-exchanged using a VivaSpin centrifugal concentrator with 10-kDa cutoff (Merck) at 8000 rpm at 4 °C against 20 mM Tris-HCl pH 7.5, 0.5 M NaCl. The fusion protein was cleaved using bovine thrombin (Sigma) by overnight incubation at 4 °C. MaSp1 was further purified from NusA and other protein contaminants by taking advantage of the propensity of recombinant MaSp1 to undergo reversible phase separation in the presence of potassium phosphate buffer (KP_i) at neutral pH: 1.0 M KP_i pH 7.0 was added dropwise until the protein sample turned irreversibly turbid, indicating phase separation; the sample was centrifuged to produce two phases, a high-density phase (HDP) containing highly concentrated MaSp1 and a low-density phase (LDP) containing the rest of the sample. The LDP was removed and the HDP was reconstituted in a small volume of 20 mM Tris-HCl pH 7.5, 0.5 M NaCl by careful pipetting. High-speed centrifugation was applied to remove any aggregated material, and the concentration of purified MaSp1 was calculated using A280 measurements of diluted sample using NanoDrop (Thermo Fisher). Protein purity was assessed by SDS polyacrylamide gel electrophoresis (SDS-PAGE). Purified MaSp1 was used as-is or otherwise flash-frozen in liquid nitrogen in small batches and stored at –80 °C until further use. In the case of the incomplete protein constructs (isolated NTD, CTD, arepetitive domains), purification was performed essentially according to previous protocols.^[6]

Circular Dichroism: Far-UV CD spectra were measured using a JASCO J820 instrument and a Starna 21Q quartz cuvette. N-R6-C was diluted to 0.1 mg mL⁻¹ in buffer consisting either of 5 mM potassium phosphate, 45 mM NaCl (pH 7.0) or 5 mM sodium acetate, 45 mM NaCl (pH 5.0). Spectra were collected from 260–190 nm at a scan rate of 50 nm min⁻¹ at 1 nm intervals and 4 s response time and 2 accumulations. Spectra were recorded from 20–70 °C at 10-degree intervals, using a temperature ramp of 1 °C min⁻¹.

NMR Experiments: Prior to NMR analysis the (¹⁵N) N-R6-C was prepared in 10 mM phosphate buffer, pH 7, 500 mM NaCl, while the (¹⁵N) CTD, (¹³C, ¹⁵N) NTD, and (¹³C, ¹⁵N) R4 samples were prepared in 10 mM phosphate buffer, pH 7, 300 mM NaCl. All samples were supplemented with 10% D₂O and 0.1 mM 4,4-dimethyl-4-silapentane-1-sulfonate (DSS). 2D ¹H-¹⁵N HSQC spectra of N-R6-C, CTD and NTD were recorded at 35 °C with 144, 4, and 4 scans, respectively, and recycle delay of 1 s. For MaSp1 R4, 2D and 3D NMR experiments were performed with 2 scans and recycle delay of 1 s, at lower temperature (10 °C) to avoid signal loss due to rapid amide proton exchange of unfolded peptide at higher temperature. To assign the 2D (¹H, ¹⁵N) HSQC of R4, 3D NMR experiments, including 3D HNCO/3D HN(CA)CO,^[28] 3D HN(CO)CA/3D HNCA^[29] and 3D CBCA(CO)NH^[30] were performed. All NMR spectra were processed using NMRPipe^[31] and visualized and analyzed using SPARKY.^[32]

NTD Dimerization Assay: NTD dimerization was monitored based on a tryptophan fluorescence shift upon pH-driven dimer formation.^[4a] The assays were prepared in 96-well black plates (Iwaki), with each 100 μL reaction containing 5.0 μM of recombinant MaSp1, background NaCl concentrations of 25, 150, or 300 mM, and a mixed buffer system containing 20 mM each of sodium acetate, MES, and HEPES, with pH adjusted to values ranging from 4.75–7.5 in increments of 0.25 pH units. Fluorescence spectra were measured at 25 $^{\circ}\text{C}$ using Spectramax M3 (Molecular Devices) with excitation at 280 nm and an emission range from 300–400 nm. The response curves were generated by calculating the fluorescence intensity ratios between 339 and 351 nm and plotting the results against pH values.

Phase Separation Map Construction: Different combinations of MaSp1 N-R6-C concentration (ranging from 2.5 to 90 μM) and KP_i pH 7.0 concentration (ranging from 0.1 to 0.8 M), against a constant 0.3 M NaCl background and at 23 $^{\circ}\text{C}$, were evaluated for LLPS formation by optical microscopy. Components were initially mixed in PCR tubes (20 μL final volume), and 5 μL aliquots of each were evaluated for the formation of liquid-like droplets using Olympus BX53 microscope equipped with a DP73 camera. Each data point was tabulated onto a grid representing MaSp1 versus KP_i concentration, from which best-fit curves were constructed by eye. Each sample point was prepared at least twice to ensure reproducibility.

Fluorescence Labeling of MaSp1: MaSp1 N-R6-C was labeled with NHS ester of Dylight-550 (Thermo Fisher Scientific) according to the manufacturer's protocols. No changes in phase separation behavior were observed between the labeled and unlabeled protein samples.

Confocal Laser Scanning Microscopy: Differential interference contrast (DIC) images of unlabeled MaSp1 condensate structures were collected using Zeiss LSM 700 confocal microscope using a Plan-Achromat 40 \times /0.95 Korr M27 lens and pinhole size of 1 AU. Samples were excited at 488 nm and imaged using the T-PMT channel.

For the 3D reconstruction experiments, 20 mg ml⁻¹ MaSp1 N-R6-C (at 1:4 ratio Dylight-500 labeled: unlabeled) was mixed with an equal volume of 1.0 M sodium citrate pH 5.0 in a glass slide chamber and sealed with a coverslip. The resulting nanofibril networks were imaged on a Leica TCS SP8 scanning confocal microscope equipped with a white light laser using the HC PL APO CS2 40 \times /1.30 oil objective and pinhole size of 1 AU. Samples were excited at 549 nm and images collected using PMT detector. Image stacking was performed under z-wide mode and system optimized settings for a range of 30 μm along the z-axis. The final 3D reconstruction was achieved using the Leica LAS X software with slight sharpness adjustment applied using ProcessTools.

Direct Observation of Biomimetic Self-assembly: Two droplets were deposited onto a 24 \times 50 mm coverslip (Matsunami) at a proximity of a few mm, one containing 10 mg mL⁻¹ of Dylight-550 labeled MaSp1 N-R6-C solution, and the other droplet consisting of either 1.0 M KP_i (pH 7.0) or 1.0 M sodium citrate buffer (pH 5.0). A smaller coverslip (18 \times 18 mm), coated with silicone grease on the edges, was placed above the sample, thereby sandwiching the two drops between the two glass planes, without them touching. This setup was placed onto the stage of Zeiss LSM 700 microscope and allowed to equilibrate for 5 min. Continuous imaging was initiated at 555 nm excitation and 573 nm emission using Plan Achromat 20 \times /0.8 M27 objective while focused on the MaSp1 drop. The ion gradient was initiated by applying gentle pressure onto the smaller coverslip using a micropipette tip, to push the ion solution droplet toward the protein sample until the two droplets merged. Time-lapse images were collected continuously at 0.64 or 1.03 frames per second for up to 10 min.

Biomimetic Fiber Formation: To produce fibers, a small volume (typically 0.5 μL) of concentrated recombinant MaSp1 solution (150 mg mL⁻¹) was pipetted onto a glass slide and immediately layered with 4 μL of 1.0 M CPB, pH 5.5, which led to immediate turbidity. By means of forceps, the protein material was pulled carefully from one edge of the condensed mass and drawn unidirectionally in air until the nascent fiber attained a length of \approx 10 cm.

Raman Spectroscopy: Raman spectra were measured using a JASCO NRS-4100 instrument combined with a DU420-OE charge-coupled device detector (Andor). The 532 nm laser was used with the 100 \times /0.90 oil immersion objective, with grating set to 1800 grooves mm⁻¹ and slit size of 10 \times 8000 μm , for a nominal resolution of 0.7 cm⁻¹. Spectra were recorded from 500 to 2000 cm⁻¹ with a beam intensity of 16.9 mw and exposure

time of 30 sec for 8 accumulations. At least five spectra were collected for each sample to ensure reproducible results. The raw spectra were processed by adjusting the background signal using a spline method integrated with the JASCO software followed by intensity normalization via the peak at 1453 cm⁻¹.

Scanning Electron Microscopy: Dried fiber samples were placed on an aluminum stub covered with conductive carbon tape, sputtered with gold using a JEOL Smart Coater, and visualized on a JCM-6000 instrument (JEOL) at 5 kV using secondary electron imaging mode and under high vacuum conditions.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This work was supported by JST-ERATO (grant number JPMJER1602), JST-COI-NEXT, and MEXT Data Creation and Utilization-type MaTerial R&D project. ADM is grateful for support from Japan Society for Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (Kakenhi), and the Mitsubishi Foundation Research Grant in the Natural Sciences. NAO is grateful for support from the RIKEN SPDR Fellowship, a JSPS Grant for Early Career Scientists (Kakenhi), and the Naito Foundation. J.C. is grateful for support from JSPS Grant for early career scientists (Kakenhi).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords

biomolecular condensate, coacervate, fibrous proteins, protein design, spider silk

Received: May 13, 2024
Revised: June 26, 2024
Published online: August 27, 2024

- [1] a) A. Miserez, J. Yu, P. Mohammadi, *Chem. Rev.* **2023**, *123*, 2049; b) A. Rising, M. J. Harrington, *Chem. Rev.* **2023**, *123*, 2155.
- [2] K. Arakawa, N. Kono, A. D. Malay, A. Tateishi, N. Ifuku, H. Masunaga, R. Sato, K. Tsuchiya, R. Ohtoshi, D. Pedrazzoli, A. Shinohara, Y. Ito, H. Nakamura, A. Tanikawa, Y. Suzuki, T. Ichikawa, S. Fujita, M. Fujiwara, M. Tomita, S. J. Blamires, J.-A. Chuah, H. Craig, C. P. Foong, G. Greco, J. Guan, C. Holland, D. L. Kaplan, K. Suresh, B. B. Mandal, Y. Norma-Rashid, et al., *Sci. Adv.* **2022**, *8*, eabo6043.
- [3] a) M. Andersson, G. Chen, M. Otikovs, M. Landreh, K. Nordling, N. Kronqvist, P. Westermarck, H. Jornvall, S. Knight, Y. Ridderstrale, L. Holm, Q. Meng, K. Jaudzems, M. Chesler, J. Johansson, A. Rising, *PLoS Biol.* **2014**, *12*, 1001921; b) D. P. Knight, F. Vollrath, *Naturwissenschaften* **2001**, *88*, 179; c) T. Lefevre, S. Boudreault, C. Cloutier, M.

- Pézolet, *Biomacromolecules* **2008**, *9*, 2399; d) F. Vollrath, D. P. Knight, *Int. J. Biol. Macromol.* **1999**, *24*, 243; e) F. Vollrath, D. P. Knight, X. W. Hu, *Proc. R. Soc. London Series B: Biol. Sci.* **1998**, *265*, 817; f) T. Y. Lin, H. Masunaga, R. Sato, A. D. Malay, K. Toyooka, T. Hikima, K. Numata, *Biomacromolecules* **2017**, *18*, 1350.
- [4] a) G. Askarieh, M. Hedhammar, K. Nordling, A. Saenz, C. Casals, A. Rising, J. Johansson, S. D. Knight, *Nature* **2010**, *465*, 236; b) W. A. Gaines, M. G. Sehorn, W. R. Marcotte, *J. Biol. Chem.* **2010**, *285*, 40745; c) F. Hagn, L. Eisoldt, J. G. Hardy, C. Vendrely, M. Coles, T. Scheibel, H. Kessler, *Nature* **2010**, *465*, 239.
- [5] F. Vollrath, D. P. Knight, *Nature* **2001**, *410*, 541.
- [6] A. D. Malay, T. Suzuki, T. Katashima, N. Kono, K. Arakawa, K. Numata, *Sci. Adv.* **2020**, *6*, eabb6030.
- [7] A. Leppert, G. Chen, D. Lama, C. Sahin, V. Railaite, O. Shilkova, T. Arndt, E. G. Marklund, D. P. Lane, A. Rising, *Nano Lett.* **2023**, *23*, 5836.
- [8] a) T. A. Blackledge, J. Pérez-Rigueiro, G. R. Plaza, B. Perea, A. Navarro, G. V. Guinea, M. Elices, *Sci. Rep.* **2012**, *2*, 782; b) A. D. Malay, K. Arakawa, K. Numata, *PLoS One* **2017**, *12*, 0183397.
- [9] M. Andersson, Q. Jia, A. Abella, X.-Y. Lee, M. Landreh, P. Purhonen, H. Hebert, M. Tenje, C. V. Robinson, Q. Meng, *Nat. Chem. Biol.* **2017**, *13*, 262.
- [10] a) M. Stark, S. Grip, A. Rising, M. Hedhammar, W. Engström, G. Hjälml, J. Johansson, *Biomacromolecules* **2007**, *8*, 1695; b) M. Landreh, M. Andersson, E. G. Marklund, Q. Jia, Q. Meng, J. Johansson, C. V. Robinson, A. Rising, *Chem. Commun.* **2017**, *53*, 3319; c) M. Otikovs, M. Andersson, Q. Jia, K. Nordling, Q. Meng, L. B. Andreas, G. Pintacuda, J. Johansson, A. Rising, K. Jaudzems, *Angew. Chem., Int. Ed.* **2017**, *129*, 12745; d) W. Finnigan, A. D. Roberts, C. Ligorio, N. S. Scrutton, R. Breitling, J. J. Blaker, E. Takano, *Sci. Rep.* **2020**, *10*, 10671; e) D. Onofrei, D. Stengel, D. Jia, H. R. Johnson, S. Trescott, A. Soni, B. Addison, M. Muthukumar, G. P. Holland, *Biomacromolecules* **2021**, *22*, 3377.
- [11] G. Hu, A. Katuwawala, K. Wang, Z. Wu, S. Ghadermarzi, J. Gao, L. Kurgan, *Nat. Commun.* **2021**, *12*, 4438.
- [12] a) G. P. Holland, M. S. Creager, J. E. Jenkins, R. V. Lewis, J. L. Yarger, *J. Am. Chem. Soc.* **2008**, *130*, 9871; b) N. A. Oktaviani, A. Matsugami, A. D. Malay, F. Hayashi, D. L. Kaplan, K. Numata, *Nat. Commun.* **2018**, *9*, 2121; c) J. D. van Beek, S. Hess, F. Vollrath, B. H. Meier, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10266.
- [13] M. Hardenberg, A. Horvath, V. Ambrus, M. Fuxreiter, M. Vendruscolo, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 33254.
- [14] J. Kyte, R. F. Doolittle, *J. Mol. Biol.* **1982**, *157*, 105.
- [15] G. D. Davis, C. Elisee, D. M. Newham, R. G. Harrison, *Biotechnol. Bioeng.* **1999**, *65*, 382.
- [16] J. Bauer, T. Scheibel, *Biomacromolecules* **2017**, *18*, 835.
- [17] S. Ittah, S. Cohen, S. Garty, D. Cohn, U. Gat, *Biomacromolecules* **2006**, *7*, 1790.
- [18] a) N. A. Oktaviani, A. Matsugami, F. Hayashi, K. Numata, *Chem. Commun.* **2019**, *55*, 9761; b) L. Xu, N. Weatherbee-Martin, X. Q. Liu, J. K. Rainey, *Small* **2019**, *15*, 1805294.
- [19] Y. Zhang, P. S. Cremer, *Annu. Rev. Phys. Chem.* **2010**, *61*, 63.
- [20] K. Jaudzems, G. Askarieh, M. Landreh, K. Nordling, M. Hedhammar, H. Jörnvall, A. Rising, S. D. Knight, J. Johansson, *J. Mol. Biol.* **2012**, *422*, 477.
- [21] a) F. Hagn, C. Thamm, T. Scheibel, H. Kessler, *Angew Chem Int Ed Engl* **2011**, *50*, 310; b) N. A. Oktaviani, A. D. Malay, A. Matsugami, F. Hayashi, K. Numata, *Biomacromolecules* **2023**, *24*, 1604.
- [22] a) A. S. Mahadevi, G. N. Sastry, *Chem. Rev.* **2013**, *113*, 2100; b) T. J. Nott, E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowitz, T. D. Craggs, D. P. Bazett-Jones, T. Pawson, J. D. Forman-Kay, *Mol. Cell* **2015**, *57*, 936; c) J. Wang, J.-M. Choi, A. S. Holehouse, H. O. Lee, X. Zhang, M. Jahnel, S. Maharana, R. Lemaitre, A. Pozniakovskiy, D. Drechsel, *Cell* **2018**, *174*, 688.
- [23] a) G. Krainer, T. J. Welsh, J. A. Joseph, J. R. Espinosa, S. Wittmann, E. de Csilléry, A. Sridhar, Z. Toprakcioglu, G. Gudiškytė, M. A. Czekalska, *Nat. Commun.* **2021**, *12*, 1085; b) F. Zhang, F. Roosen-Runge, A. Sauter, M. Wolf, R. M. J. Jacobs, F. Schreiber, *Pure Appl. Chem.* **2014**, *86*, 191.
- [24] A. Spohner, E. Unger, F. Grosse, K. Weisshart, *Nat. Mater.* **2005**, *4*, 772.
- [25] G. Greco, J. Francis, T. Arndt, B. Schmuck, F. G. Bäcklund, A. Barth, J. Johansson, N. M. Pugno, A. Rising, *Molecules* **2020**, *25*, 3248.
- [26] J. Sparkes, C. Holland, *Nat. Commun.* **2017**, *8*, 594.
- [27] J. Chen, A. Tsuchida, A. D. Malay, K. Tsuchiya, H. Masunaga, Y. Tsuji, M. Kuzumoto, K. Urayama, H. Shintaku, K. Numata, *Nat. Commun.* **2024**, *15*, 527.
- [28] a) M. Ikura, L. E. Kay, A. Bax, *Biochemistry* **1990**, *29*, 4659; b) T. Yamazaki, W. Lee, C. H. Arrowsmith, D. R. Muhandiram, L. E. Kay, *J. Am. Chem. Soc.* **1994**, *116*, 11655.
- [29] A. Bax, M. Ikura, *J. Biomol. NMR* **1991**, *1*, 99.
- [30] M. Wittekind, L. Mueller, *J Magn Reson B* **1993**, *101*, 201.
- [31] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. D. Bax, *J. Biomol. NMR* **1995**, *6*, 277.
- [32] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325.