

---

This is an electronic reprint of the original article.  
This reprint may differ from the original in pagination and typographic detail.

Batys, Piotr; Fedorov, Dmitrii; Mohammadi, Pezhman; Lemetti, Laura; Linder, Markus B.; Sammalkorpi, Maria

**Self-assembly of silk-like protein into nanoscale bicontinuous networks under phase separation conditions**

*Published in:*  
Biomacromolecules

*DOI:*  
[10.1021/acs.biomac.0c01506](https://doi.org/10.1021/acs.biomac.0c01506)

Published: 08/02/2021

*Document Version*  
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

*Please cite the original version:*  
Batys, P., Fedorov, D., Mohammadi, P., Lemetti, L., Linder, M. B., & Sammalkorpi, M. (2021). Self-assembly of silk-like protein into nanoscale bicontinuous networks under phase separation conditions. *Biomacromolecules*, 22(2), 690-700. <https://doi.org/10.1021/acs.biomac.0c01506>

# Self-assembly of silk-like protein into nanoscale bicontinuous networks under phase-separation conditions

Piotr Batys<sup>1,2\*</sup>, Dmitrii Fedorov<sup>2</sup>, Pezhman Mohammadi<sup>3</sup>, Laura Lemetti<sup>2</sup>, Markus B. Linder<sup>2</sup>, Maria Sammalkorpi<sup>2,4\*</sup>

<sup>1</sup>Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, Niezapominajek 8, PL-30239 Krakow, Poland

<sup>2</sup>Department of Bioproducts and Biosystems, School of Chemical Engineering, Aalto University, FI-00076 Aalto, Finland

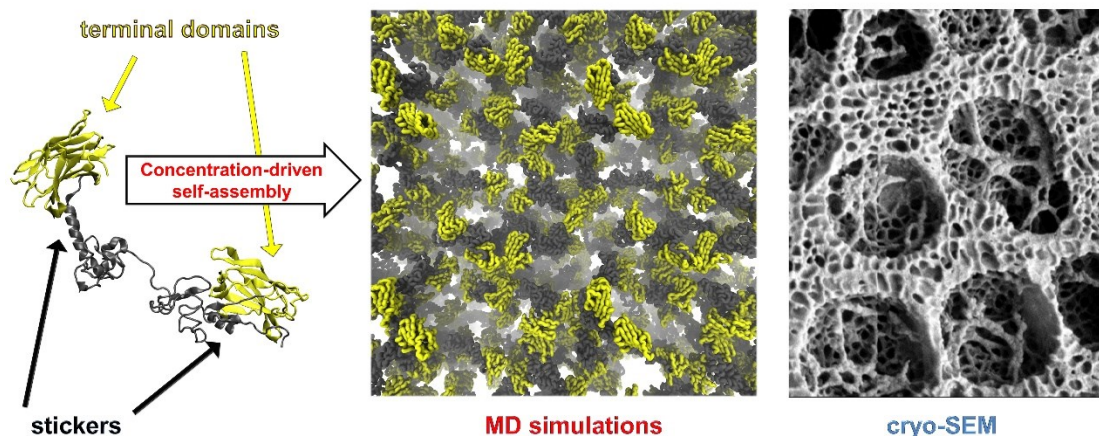
<sup>3</sup>VTT Technical Research Centre of Finland Ltd. VTT, FI-02044 Espoo, Finland

<sup>4</sup>Department of Chemistry and Materials Science, School of Chemical Engineering, Aalto University, FI-00076 Aalto, Finland

\*Corresponding authors:

Piotr Batys ([piotr.batys@ikifp.edu.pl](mailto:piotr.batys@ikifp.edu.pl)), Maria Sammalkorpi ([maria.sammalkorpi@aalto.fi](mailto:maria.sammalkorpi@aalto.fi))

## Graphical abstract



## ABSTRACT

Liquid-liquid phase separation (LLPS) of biomacromolecules is crucial in various inter and extracellular biological functions. This includes formation of condensates to control e.g. biochemical reactions and structural assembly. The same phenomenon is also found to be critically important in protein based high performance biological materials. Here, we use a well-characterized model triblock protein system to demonstrate the molecular level formation mechanism and structure of its condensate. Large-scale molecular modelling supported by analytical ultracentrifuge (AUC) characterization combined with our earlier high magnification precision cryo-SEM microscopy imaging lead to deducing that the condensate has a bicontinuous network structure. The bicontinuous network arises from the proteins having a combination of sites with stronger mutual attraction and multiple weakly attractive regions connected by flexible, multiconfigurational linker regions. These attractive sites and regions behave as stickers of varying adhesion strength. For the examined model triblock protein construct, the  $\beta$ -sheet rich end units are the stronger stickers while additional weaker stickers, contributing to the condensation affinity, arise from spring-like connections in the flexible middle region of the protein. The combination of stronger and weaker sticker-like connections and the flexible regions between the stickers result in a versatile, liquid-like, self-healing structure. This structure also explains the high flexibility, easy deformability and diffusion of the proteins decreasing only 10-100 times in the bicontinuous network formed in the condensate phase in comparison to dilute protein solution. The here demonstrated structure and condensation mechanism of a model triblock protein construct via a combination of the stronger binding regions and the weaker, flexible sacrificial-bond-like network, as well as its generalizability via polymer sticker models, provide means to understand not only intracellular organization, regulation, and cellular function but also identifies direct control factors for and enables engineering improved protein and polymer constructs to enhance control of advanced

fiber materials, smart liquid biointerfaces or self-healing matrices for pharmaceuticals or bioengineering materials.

## INTRODUCTION

In liquid-liquid phase separation (LLPS), a macromolecular solution reorganizes at molecular level to two (or more) coexisting liquid phases differing in their molecular compositions. This spontaneous phase separation, and especially the forming of a condensate phase, is crucially important in biological systems but also emerges as a critical key step to advanced control in many bio and polymer technology applications.<sup>1-4</sup> In cells, LLPS results in, for example, membraneless organelles, such as nucleoli, Cajal bodies and PML nuclear bodies in the nucleus, but also, at a more general level, compartmentalize molecules, control molecular environments and via that command many cellular biochemical reactions.<sup>1,4</sup> The significance of LLPS in biological cells rises from the phase separation giving the cell means to concentrate proteins, nucleic acids, and other important molecular components at discrete cellular sites, and via that provide the cell control over function. Yet, as the condensate phase boundary is a rather permeable composition gradient in a liquid, molecular exchange between the phases occurs readily: the condensates are in constant equilibrium exchange of molecular components with their surroundings.<sup>5</sup> This means also that condensates can readily form from, or dissolve into the dilute phase, according to biochemical signals or conditions which impacts their biological function.<sup>6-8</sup> Examples of this are the reversible stress-induced condensates<sup>9</sup> or function inducing reversible grouping of signaling proteins in biological cells.<sup>10</sup> LLPS has also been demonstrated to have a role in prion formation, i.e. controlling folding of biomacromolecules. Similarly, LLPS may have a role in several neurodegenerative illnesses such as Parkinson's, Alzheimer's and ALS, see Ref. <sup>11</sup> for a recent review.

Further outlining the importance of the LLPS, in biotechnology especially in biological materials engineering, condensates have so far been confirmed to be key steps, e.g. for elastomer protein,<sup>12</sup> amyloid,<sup>13</sup> or silk-like protein construct based materials,<sup>14,15</sup> but also materials rising inspiration from, e.g. mussel feet adhesive proteins,<sup>16–18</sup> squid beak protein components, or nacre protein biomineralization.<sup>19</sup> In polymer technology, the condensate phase is referred to as coacervate phase and it is a key assembly phase to, for example, the superior mechanical characteristics,<sup>20–22</sup> superhydrophobic,<sup>23</sup> or responsive materials<sup>24</sup> rising from hydrated polyelectrolyte assemblies, in particular layer-by-layer polyelectrolyte multilayers (PEMs) and polyelectrolyte complexes (PECs). Hydrated polyelectrolyte assemblies rising from LLPS have a wide set of technological applications in e.g. functional films, energy applications, drug delivery systems and adhesives.<sup>18,24,25</sup> Controlling the coacervate phase via e.g. ionic strength,<sup>26</sup> temperature,<sup>27</sup> pH,<sup>28</sup> but also aqueous solvent,<sup>29</sup> are commonly used means in engineering the characteristics of PEM and PEC based synthetic polymer materials.

Due to this wide basic biology and technological relevance, liquid-liquid phase separation has attained significant interest in the scientific community. Fundamentally, the aqueous, multicomponent macromolecular solutions present in all the above-described biology and bio- and polymer materials science systems phase separate because the lowest achievable total free energy state corresponds to partitioning the aqueous solution molecular composition unevenly,<sup>30,31</sup> i.e. to a macromolecule rich condensate phase and a macromolecule dilute carrier phase. Simplistically thinking, the entropy loss involved in demixing (phase separating) a uniformly mixed molecular solution needs to be countered by enthalpically favorable interactions, i.e. attraction between the condensing species. However, for proteins condensates it is worth noting that an especially prevalent class of proteins present in biomolecular condensates is intrinsically disordered proteins (IDPs).<sup>4,32,33</sup> IDPs do not have a single specific

folded structure, but instead have a large range of flexibility in their conformations, and may change their conformational phase space, or folding state upon surroundings, for example condensation. For example, intracellular phase separation and formation of membrane-less bodies in cells seems to involve proteins with significant conformational heterogeneity and intrinsically disordered regions<sup>34–36</sup>. Altogether, the prevalence of IDPs in protein condensates points toward the conformational flexibility and degree of disorder in the condensing macromolecule having a role in the condensate formation.

In addition to flexibility and disordered regions, condensation of proteins has previously been associated with attractive interactions between parts of the condensing proteins, see e.g. Ref. <sup>1</sup>. Various binding domains, interactions between helical and  $\beta$ -sheet regions,  $\pi$ - $\pi$ - interactions,  $\pi$ -cation, and electrostatics based attraction have been suggested to contribute to separation.<sup>1</sup> Especially segments of oppositely charged residues appear to enhance condensation in many cases, see e.g. Refs. <sup>37–39</sup>. As expected, attraction rising from electrostatics is sensitive to pH and salt.<sup>40</sup> The salt dependency rises mainly because of electrostatic screening and ion adsorption.<sup>41</sup> However, the salt response can also have complex concentration and salt species dependencies.<sup>42,43</sup> Altogether, this points toward interactions with water and water solvation of the ions and macromolecules having a role in the protein condensate formation. Signaling molecules that have an important role in intracellular organization and phase separation contain repeats of weakly binding, associative domains.<sup>34</sup> Additionally, sequences of low-complexity regions<sup>44</sup> and a small subset of aliphatic amino acids, especially alanine, glycine and serine,<sup>45,46</sup> have been associated with phase separation.

Condensation mechanism considerations for multidomain macromolecules, such as complex proteins, have so far focused on strong, specific interactions between regions of the macromolecules forming a network that holds the condensate structurally intact.<sup>33</sup> More recently, the presence of multiple attractive regions (multivalency) has been reported to

promote phase separation<sup>4,47</sup> and also the coexistence of interactions of different magnitudes (stronger and weaker interactions) appear to be condensation prerequisites for a protein system.<sup>48</sup> Additionally, in some other protein systems, structural transitions promoting  $\beta$ -sheet formation seem to promote LLPS.<sup>49</sup> Furthermore, condensate formation examination via various engineered protein systems composed of repeated folded domains connected by flexible linkers have revealed that the number of attractive regions (valency) and their affinity both promote assembly into larger structures, enable phase separation at lower concentrations,<sup>50</sup> and decrease dynamic rearrangements of molecules in the condensed droplets.<sup>48</sup>

Theoretical considerations of condensation mechanisms in macromolecule solutions involve both classical polymer physics-based considerations, see e.g. Refs. <sup>34,36,51</sup> for reviews, and traditional materials science statistical mechanics based phase separation theory approaches, see Ref. <sup>31</sup> for a review. The polymer physics approaches focus on solvation and mixing characteristics, as well as, the effect of associative interactions between polymer segments<sup>52,53</sup> to explain the liquid-liquid phase separation characteristics. Berry et al.<sup>31</sup> argued that the liquid-liquid phase separation can in most cases be understood via equilibrium considerations of the materials mixtures. Linear proteins with multiple attractive interaction regions (multivalent proteins) can be considered as associative polymers<sup>52,53</sup> with specific intra and intermolecular “stickers” of non-covalent interactions between the domains.<sup>54</sup> The stickers in these systems span a range of length scales, can vary in strength and also be directional in nature.<sup>34</sup> Examples include localized, point-like charge interactions or hydrogen bonding but also larger scale hydration-mediated interactions including hydrophobic attraction between both individual protein segments but also larger units of secondary structure such as  $\beta$ -sheet or  $\alpha$ -helical domains. The sticker approach has been relatively popular also in numerical studies. For example, Martin et al.<sup>13</sup> employed such sticker-and-spacer approach to explain the phase

separation behavior of their prion protein system, Zeng et al.<sup>55</sup> a Gaussian cluster theory approach also based on considering the proteins to have sticker sites and Choi et al.<sup>56</sup> employed a lattice model based on the sticker concept to simulate phase transitions of protein systems.

The traditional computational means to model the phase separation is field theory type statistical mechanics-based approaches but as biomolecular condensate formation is often subject to localized interactions, more recently also particle-based simulation methods that cover microscale from atomistic detail to mesoscale have become popular.<sup>57</sup> They can provide predictive frameworks and may aid in interpretation of experimental results by directly probing the underlying physics with molecular detail.<sup>58</sup> Condensate formation in biological systems has been studied via the former approach in Refs.<sup>59,60</sup> while particle-based approaches are presented in Refs.<sup>58,61–63</sup>. While the field-based studies resolve basic premises for LLPS, they suffer from lack of chemical specificity and the microstructural correlations, even at the level of secondary structure. On the other hand, particle-based simulations are limited by their computational cost which results to either limitations in the description resolution or describable system size and time scale as well as their inability to efficiently probe rare events, such as low-probability conformational rearrangements and nucleation.<sup>57</sup> So far, the particle-based simulation were mainly used to correlate single-molecule properties with phase behavior<sup>62–64</sup>. Atomistic detail molecular dynamics to resolve structural features and adhesive sites in condensate forming proteins and their subunits have been reported.<sup>65</sup> Coarse-grained molecular modelling has resolved, e.g., phase transition temperature dependency as the function of protein system properties.<sup>62</sup> Despite efforts in understanding the condensate microstructure, e.g., formed in protein-RNA systems,<sup>66,67</sup> so far, no studies have resolved the formation and the assembly mechanism of silk-like protein condensate at microstructural level. This might indicate that the condensate microstructure is system-specific and also that for many LLPS systems, no specific microstructure emerges.



In our previous experimental work on a model triblock protein construct CBM-AQ12-CBM exhibiting condensate formation<sup>14,15,17</sup>, the LLPS enabled pulling of long, resilient protein fibers out of it,<sup>14</sup> and the condensate turned out to be an efficient adhesive<sup>17</sup>. Motivated by this, we examine here the CBM-AQ12-CBM condensate via large scale molecular modelling, combined with analytical ultracentrifuge (AUC) characterization of the solvated proteins. The role of AUC is to verify the range of protein configurations and folding shapes rising from the molecular simulations of the partially disordered protein. Previously reported cryo-SEM imaging of the condensate<sup>14</sup> provides a direct comparison and verification means to the structures observed in modelling.

We chose this model triblock construct because we already have significant understanding of the condensate macroscopic properties and the conditions over which LLPS occurs. Furthermore, the protein construct is modular, contains intrinsically disordered regions, yet as a model system is simple enough to potentially reveal the condensate formation mechanism and its control means. Additional interest to this particular protein construct choice rises because its condensate has been reported to have a bicontinuous network structure.<sup>14</sup> Hence, examining this particular model protein system could resolve the molecular features giving rise to the bicontinuous network structure, as well as, its potential significance in condensate system properties. Altogether, the performed study is the first that we are aware in which large scale molecular simulations are combined with such high precision experimental characterization methods (both AUC and cryo-SEM) that the modelling can be directly compared and verified against the experimental data. This enables extracting from the modelling data the condensate molecular level structure, its control means and formation mechanism while the overall assembly matches experimental microscopy images.

## COMPUTATIONAL AND EXPERIMENTAL METHODS

## Computational methods

The Gromacs 5.1.4 package<sup>68,69</sup> with the Amber03ws force field<sup>70</sup> was used for all-atom molecular dynamics (MD) modeling of the silk-like proteins. For water, in compliance with the force-field choice, the explicit TIP4P/2005 model was employed.<sup>71</sup> The simulations model the triblock architecture of the CBM-AQ12-CBM, i.e. a repetitive mid-block terminated by CBMs at each end, but instead of the twelve AQ repeats in the experimentally characterized protein, the simulated protein CBM-AQ3-CBM consists of three AQ repeats, see Figure 1. The CBM-AQ3-CBM is used in the modelling because it captures the end unit interactions, effect of linkers, and has a sufficiently long repeat of AQ units to obtain both flexibility and intra- as well as intermolecular contact sites, following the longer repeat unit; the shorter repeat enhances dynamics and improves statistics in the simulation sampling.

Protein bonds were constrained by the LINCS<sup>72</sup> and water by the SETTLE<sup>73</sup> algorithms. The equations of motion were integrated using the leap-frog integration scheme and a 2 fs time step. Configurations were saved for analysis every 10 ps. The protein solutes were coupled to the heat bath separate from the solvent and ions. Temperature was maintained at  $T = 298$  K using the Bussi et al. canonical V-rescale thermostat<sup>74</sup> and pressure at 1 bar by the isotropic Parrinello-Rahman barostat.<sup>75</sup> The coupling constants were  $\tau = 0.1$  ps and  $\tau_p = 2$  ps, respectively. Long-range electrostatic interactions were modelled by the PME method.<sup>76</sup> Van der Waals interactions were described using the Lennard-Jones potential with a 1.0 nm cut-off. Long-range dispersion corrections were applied. Periodic boundary conditions were applied in all directions. All molecular visualizations employ the VMD software package.<sup>77</sup>

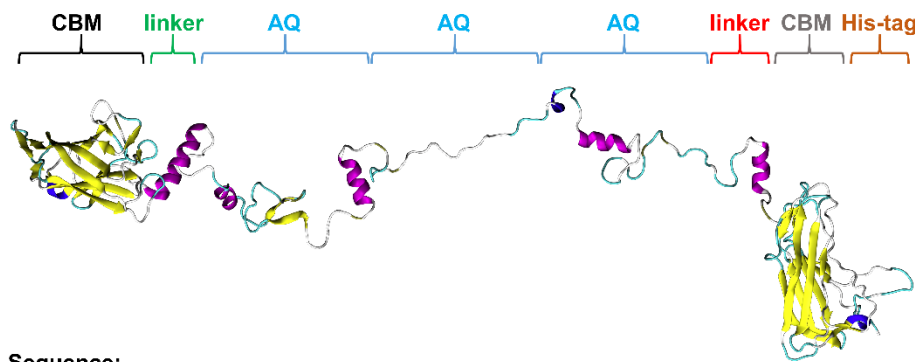
The CBMs forming the terminal domains have the crystal structure of Ref. <sup>78</sup> (Protein Data Bank ID code 4JO5) as their initial configuration. Initial configurations for the linkers, AQ units, and H6-tag were generated using the Avogadro software<sup>79</sup> with initial secondary

structures following Ref. <sup>80</sup> The protein parts were merged using VMD<sup>77</sup> and pdb2gmx tool of Gromacs. To reduce bias resulting from the initial configuration in interactions between different protein regions of the flexible AQ repeat part, the configurations were relatively extended, see Fig. 1. The protonation states of the residues were set to match pH = 7 resulting in a total charge of -5e per each CBM-AQ3-CBM molecule. In all molecular dynamics runs, a corresponding number of Na<sup>+</sup> counter ions to make the simulation systems charge neutral were introduced into the solvent.

Two different types of CBM-AQ3-CBM protein simulations were performed: single protein and 200 mg/ml concentrated solution multiprotein simulations containing either 10 or 28 CBM-AQ3-CBM constructs. For single-protein simulations seven repeat runs were performed. For the concentrated solution, containing 10 proteins, six different initial configurations with varying composition (see Table S1), while for the biggest system, containing 28 proteins, there was a single initial configuration. The single protein simulations targeted at resolving structure and solution characteristics information of the protein construct, and variation of these as the protein contains intrinsically disordered flexible parts, while the concentrated multiprotein simulation represents the conditions at which CBM-AQ12-CBM is known to form a condensate phase in experiments.<sup>14</sup> For the single protein simulations, the single CBM-AQ3-CBM proteins were first solvated by water in a  $29 \times 11.5 \times 9.5$  nm<sup>3</sup> box, periodic boundary conditions in all directions. The AQ repeat part of the protein adopted significantly less extended conformations during the first 10 ns of nanoseconds of simulation: at 100 ns, the protein and a 2 nm thick water shell around it were extracted and solvated again in a  $16 \times 16 \times 16$  nm<sup>3</sup> box as this captured more efficiently the protein shape. The NPT run was continued for 100 ns to a total of 200 ns simulation time. The concentrated, multiprotein solutions were constructed from 10 or 28 final protein conformations of the single protein simulations set in random orientation and placement into a simulation box. In these multiprotein simulations, the

simulation box size for each system corresponded to 200 mg/ml protein concentration and was  $16.5 \times 16.5 \times 16.5 \text{ nm}^3$  (for 10 proteins) or  $22.6 \times 22.6 \times 22.6 \text{ nm}^3$  (for 28 proteins). The multiprotein simulations were run for 300 ns.

Diffusion coefficients  $D$  were calculated using the mean square displacement (MSD) of the protein center of mass. The linear region on the MSD plot was approximated by a straight line according to the Einstein relation. We note that the simulation duration influences the absolute value of  $D$  and the point in presenting the values is a qualitative comparison.



#### Sequence:

```

MGNLKVEFYNSNPSTTNSINPQFKVTNTGSSAIDLSKLTLYYYTVDGQKDQTFWCDHAAIIGSNGSYN
GITSNVKGTfVKMSSSTNNADTYLEISFTGGTLEPGAHVQIQGRFAKNDWSNYTQSNDSYFKSASQFVE
WDQVTAYLNGVLVWGKEPSASASASAGASAAASAGAGAGA(GPYGPGASAAAAAGGYGPGSGQQG
PGQQGPGQQGPGQQGPGQQG)3ASASASAAASAASTVANSSSNLKVEFYNSNPSTTNSINPQFKVTNT
GSSAIDLSKLTLYYYTVDGQKDQTFWCDHAAIIGSNGSYNGITSNVKGTfVKMSSSTNNADTYLEISFT
GGTLEPGAHVQIQGRFAKNDWSNYTQSNDSYFKSASQFVEWDQVTAYLNGVLVWGKELEHHHHHH

```

**Fig. 1. The silk-inspired protein construct CBM-AQ3-CBM used in the molecular dynamics simulations and its amino acid sequence. The visualization shows a sample initial configuration used in the simulations. The initial configurations are extended in comparison to the resulting final simulation configurations for better configurational sampling.**

## Experimental methods

Cloning, expression and purification of the CBM-AQ12-CBM was carried out according to our earlier study.<sup>14</sup> Determination of the anisotropy (deviation from spherical shape) of the protein was conducted by a Beckman Coulter Optima Analytical Ultracentrifuge. Sedimentation velocity experiment were performed at 50000 rpm and at 20 °C. As a buffer we used 0.1 M of NaCl, concentration of the CBM-AQ12-CBM, commonly referred to also as

CBM-eADF3-CBM,<sup>14</sup> protein was 0.5 mg/ml. Sample was measured by UV/Vis absorbance optics at 280 nm.

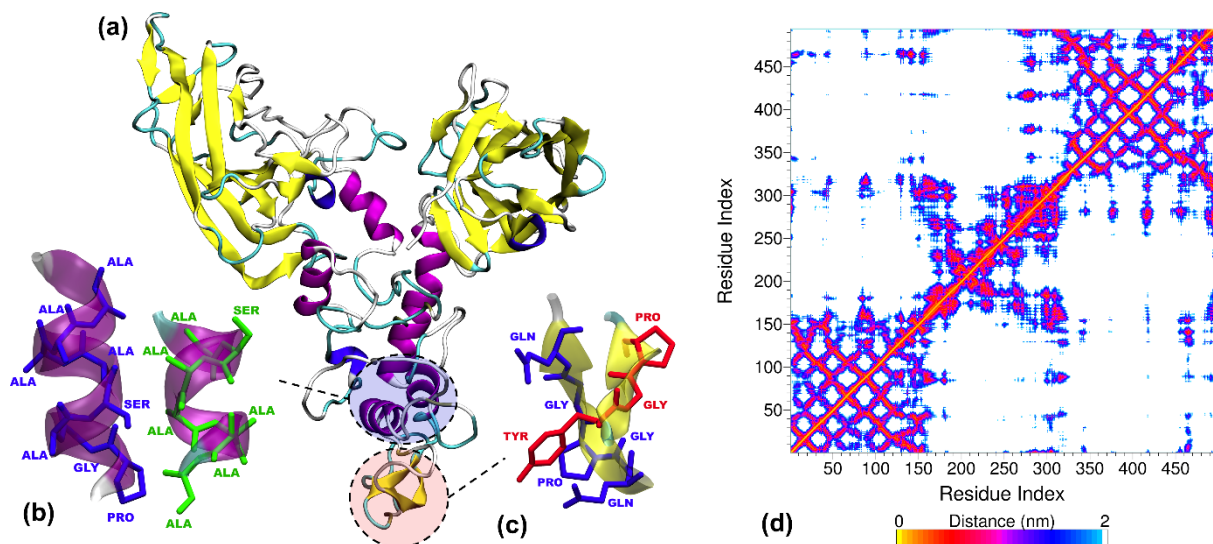
Data analysis were performed by Ultrascan III version 4.0 revision 2528 (<http://www.ultrascan.aucsolutions.com>). Partial specific volume was determined from the sequence and is 0.6994. Noise reduction was conducted by two dimensional spectrum analysis.<sup>81</sup> We used Genetic Algorithm for further refinement and regularization.<sup>82</sup> More detailed information about the experiment setups and data treatment procedures is available in Ref.<sup>83</sup>. In presented data values, the error estimate corresponds to a 95% confidence interval. The frictional ratio of protein constructs were estimated using the UltraScan Solution Modeler (US-SOMO).<sup>84</sup>

## RESULTS AND DISCUSSION

The molecular simulations of single CBM-AQ3-CBM constructs in water (200 ns) indicate the protein adopts a significantly compacted shape in water solution. Figure 2 presents a detailed analysis corresponding to one of the examined seven different single CBM-AQ3-CBM initial configurations, while the data corresponding to the rest of the single protein simulations are provided in SI (Figure S1). A summary of the findings about the conformations adopted by the protein construct is that it has five stable  $\alpha$ -helical Ala-rich regions located one in each linker and AQ unit, see Fig. 2(a-b), as well as, a metastable Gly-Gly  $\beta$ -sheet region, see Fig. 2(c). Stable  $\alpha$ -helical regions observed in MD simulations are in line with the high propensity of Ala-rich regions to form  $\alpha$ -helices that has been reported in literature<sup>85</sup> and confirmed in our previous Fourier transform infrared spectroscopy<sup>14</sup> and circular dichroism<sup>86</sup> experiments for CBM-AQ12-CBM protein. Additionally, the conformations show intramolecular  $\alpha$ -helix interactions and  $\alpha$ -helix bundling, see Fig. 2(b). The structure of CBM is stable in the simulations. The residue contact map corresponding to the sample configuration in Figure 2(a-

c) is presented as Figure 2(d). The contact map reveals the  $\alpha$ -helix bundling as strips directly adjacent to the diagonal line, but the most dominant features are the two large crisscross patterns at the ends of the diagonal of the graph. These are characteristic for parallel and antiparallel  $\beta$ -sheets at the terminal CBM segments.

Analysis of the rest of the single CBM-AQ3-CBM trajectories starting from different initial configurations presents some variation in the degree of compactness of structure and AQ3 region packing conformations but the general structural features described above persist. Additionally, due to the multitude of conformations to which the  $\beta$ -sheet and  $\alpha$ -helical regions can pack such that the hydrophobic segments are shielded from water, the chain extensions show variation and conformational flexibility, as expected for an intrinsically disordered protein. Detailed analysis of the simulations with different initial configurations, their secondary structure evolution, the corresponding residue contact maps, and time evolution of the radius of gyration of the structures showing the flexibility are provided in the Supporting Information (Figs. S1 and S2).



**Fig. 2.** Sample conformation of a single silk-inspired protein CBM-AQ3-CBM after 200 ns simulation and its residue contact map. Subfigure (a) shows the protein, (b) and (c)

**magnifications of  $\alpha$ -helical and  $\beta$ -sheet features in the structure and their residue compositions, and (d) the map of smallest distance between residue pairs of the entire protein construct.**

Let us next compare the single protein conformations present in our simulations (CBM-AQ3-CBM) against the shape of CBM-AQ12-CBM proteins in dilute solution assessed by our AUC experiments. The AUC data analysis gives a value of  $1.81 \pm 0.05$  for the frictional ratio that characterizes the shape of the individual molecules (data provided in SI, Figure S3). The simulations frictional ratio of the CBM-AQ3-CBM constructs, as determined by US-SOMO,<sup>84</sup> is calculated as an average over the seven different final configurations (see SI). These vary between 1.33 and 1.6 in their frictional ratio, with a mean of  $1.45 \pm 0.05$ . For a comparison, we determined the frictional ratio also for sample extended constructs of CBM-AQ3-CBM (see Fig. 1) and CBM-AQ5-CBM. For the sample extended constructs, the frictional ratio was 2.0 and 2.8, respectively. If one calculates a corresponding estimate for an extended CMB-AQ12-CBM protein, the expected value of the frictional ratio is around 5-6. Notably, both the AUC data derived value and the value calculated from the simulations correspond to significantly compacted protein structures. Altogether this means that the examined protein constructs adopt significantly compacted structures in water solution; the difference between experimentally and computationally determined frictional ratios can be directly explained by the higher number of repeating units in the experimental construct. We conclude the compacted conformations adopted by the single proteins represents a similar level of compactness as indicated by the AUC data. Hence, using the final configurations of the single protein simulations to construct the multiprotein initial configurations is a warranted choice.

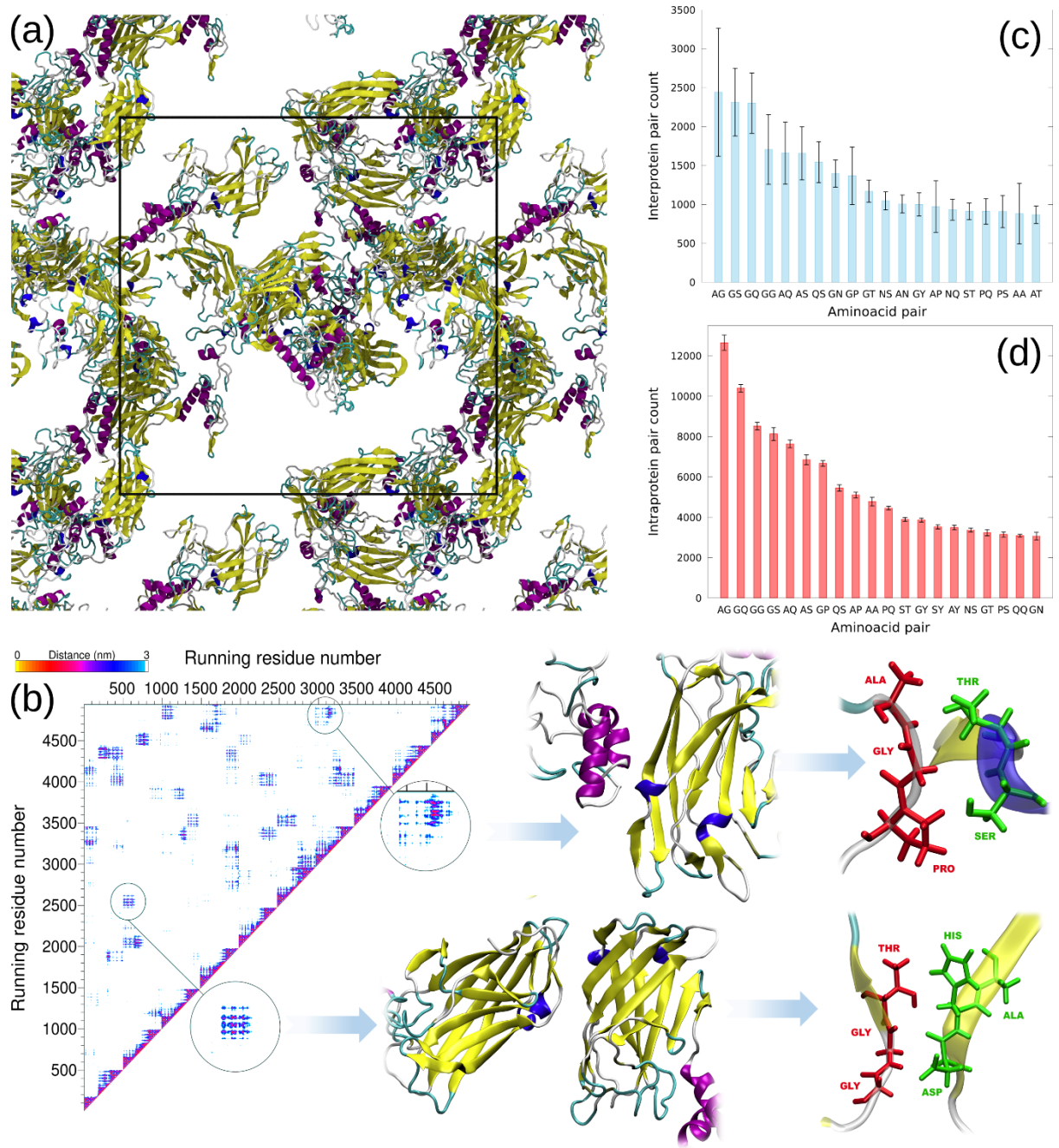
### **Concentrated protein solution**

Next, the concentrated protein solution was modelled by using the conformations rising from the single protein simulations in random order and orientation as initial conformations of the

multiprotein setups. Systems of both 10 and 28 proteins at the same 200 mg/ml concentration were modelled for 300 ns. Notably, the time scale is sufficient for local relaxation, structural changes of the proteins locally, and changing local order but not for full protein chain diffusion.

In all multiprotein simulations corresponding to this elevated protein concentration the proteins reorganize to a non-uniform, sponge-like network arrangement. In this, the individual molecules form a sponge-like frame with the  $\beta$ -sheet rich CBM end units acting as main binding stickers. Examples of the network are presented in Figure 3a (10 protein system) and Figure 4a (28 protein system). The simulations corresponding to the different initial configurations in terms of the single protein conformations, their orientations, and precise placement result in structurally very similar networks as those presented in the figures as samples. This is quantified by the similarity of the amino acid-specific pair counts averaged over all different initial configurations presented in Figs. 3(c) and (d) for the multiple protein system, correspondingly. We have not observed any significant changes in the secondary structure in comparison to the single molecule simulations. The transition of the  $\alpha$ -helical regions to  $\beta$ -sheets is activated by the external forces, i.e. shear forces or pulling.<sup>87</sup>

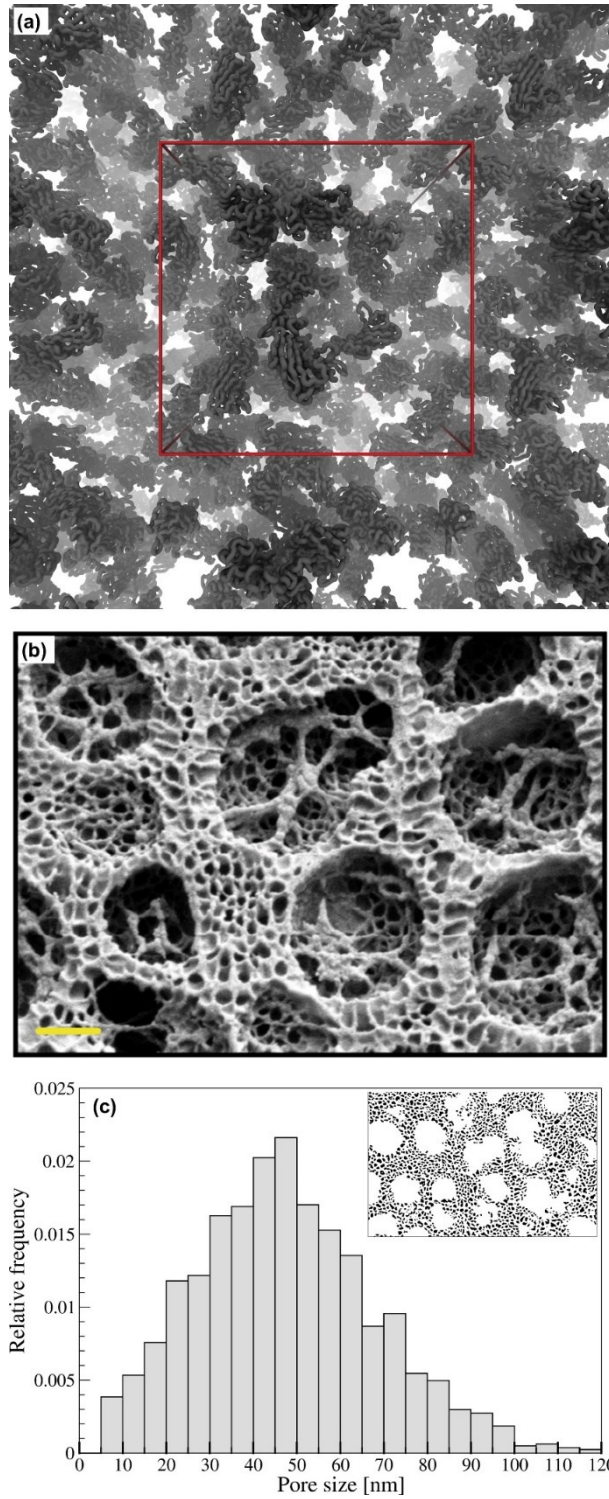




**Fig. 3.** a) A sample protein condensate formed in the molecular dynamics simulations of concentrated, 200 mg/ml protein solution (10 CBM-AQ3-CBM proteins in the simulation box). b) The contact map of the corresponding system. The CBMs are highlighted in yellow. The snapshots on the right show the contacts between proteins corresponding to the different patterns and the amino acids involved. The running residue number  $r = (n-1)R + r_i$ , where  $n$  is the molecule index (protein count index),  $r_i$  the residue number within a protein, and  $R = 493$  the total number of residues within a single protein. (c) Interprotein and (d) intraprotein amino acid-specific pair counts averaged over the 6 different initial configurations.

As visible in Figures 3 and 4a, the network appears to be carried by compacted regions of the  $\beta$ -sheet rich CBMs and bundles of  $\alpha$ -helices, both intramolecular and intermolecular. Figure 3 presents a structural analysis of the network via a contact map and identification of typical binding patterns. Similar to individual proteins (Fig. 2), each pattern in the contact map of Figure 3 can be associated with corresponding intermolecular connection between proteins: the crisscrossing patterns of the  $\beta$ -sheet rich CBMs dominate but also  $\alpha$ -helix –  $\beta$ -sheet connections between the proteins exist. The magnifications in Fig. 3b show examples of these two characteristic connection patterns in this system and the involved protein segments. In both types of connections, attractive interactions between the molecules rise mainly from Ala, Ser, and Gly amino acids, see Fig. 3(c). Furthermore, we assessed the fluctuations of the amino acid pairwise contacts: although difficult to measure quantitatively, the amino acid pairing fluctuations measure a standard deviation of approximately 10%. This means, the pairings are flexible and evolve in time.

Notably, periodic boundary conditions may enhance or even cause the formation of the observed networks and their structural specifics. Both the 10 protein and the significantly larger 28 protein system result systematically in repeat runs with different initial configurations in very similar networks pointing toward the network structure rising from the protein. Contact maps for all the different simulation runs and initial configurations are included in the SI (Figure S4).



**Fig. 4. (a) The protein condensate formed in the molecular dynamics simulations of concentrated 200mg/ml protein solution, with 28 CBM-AQ3-CBM proteins in the  $22.6 \times 22.6 \times 22.6 \text{ nm}^3$  simulation box. (b) A high-magnification SEM image of the internal structure of a condensate droplet shows details of an internal bicontinuous network (scale bar 200 nm), reprinted with permission (CC-BY 4.0 License, <http://creativecommons.org/licenses/by/4.0/>) from Ref. <sup>14</sup> (c) Pore size distribution of**

**condensate bicontinuous structure. The inset shows an image after post-processing, used for the pore size distribution determination.**

In our earlier work, we examined the liquid-liquid phase separation and the condensate structure formation was examined also by Cryo-SEM imaging.<sup>14</sup> Interestingly, Cryo-SEM showed the formation of a bicontinuous network that has a sponge-like structure, see Fig. 4b. Figure 4c presents the pore size characterization while 4a presents the simulation predicted network. Perusal of especially the extended, 28 protein system reveals that the network-like structure observed in simulations bears a remarkable resemblance to the Cryo-SEM imaging of the condensate samples, see Figure 4. To gain understanding to the connection of the experimental microscopy images and the structure adopted by the proteins in the simulations, and via that, the fundamentals of liquid-liquid phase separation and its relation to the formation of protein fibers, we analyzed both the Cryo-SEM images and the simulations for the pore size distribution. For Cryo-SEM images, the average pore diameter is  $47.7 \pm 0.5$  nm (determined by ImageJ software<sup>88</sup>) which is of the same order of magnitude as the size of a single molecule of the protein CBM-AQ12-CMB. In the simulations, even the 28 protein the system is so limited in size that a full assessment of the pore size distribution is unfeasible. However, the pores rising in this system are typically between 10 and 15 nm in diameter, measured using the VMD software.<sup>77</sup> The difference can be directly explained by the shorter AQ repeat between the end domains in the simulations, as the network in the simulations configuration is dominantly carried by the end domain connections. Furthermore, the pore size observed in simulations is also limited by the size of the simulation box. In total, the structure obtained in the simulations is consistent with the Cryo-SEM imaging and provides a molecular level explanation of the formed structure.

As the modelled network seems to be consistent with the Cryo-SEM imaged structure of the condensate, let us next discuss the significance of the simulations findings for understanding the condensate. First, the dominance of CBM connections in the network formed in the simulations suggests that the initial assembly in the protein solution could be driven by attractive interactions between the terminal groups. This is in line with experimental, AUC based observations of isolated CBMs showing dimerization with a weak dissociation constant  $K_D$  of  $90 \pm 30 \mu\text{M}$  even at low concentration in aqueous solutions.<sup>89</sup> Furthermore, the role of terminal domain for condensate formation of AQ12 protein construct is demonstrated by the concentration threshold needed for the phase separation increasing significantly, when SpyCatcher was used as a terminal domain (instead of CBM) in SpyCatcher-AQ12-SpyCatcher proteins as opposed to CBM-AQ12-CBM proteins.<sup>14</sup> Although we are not aware of  $K_D$  values corresponding to SpyCatcher, the finding shows the end unit is important in condensate formation.

In the simulational network, CBM-CBM connections carry the main load in supporting the network structure while  $\alpha$ -helix bundles have a lesser role. It is worth noting that the experiments employ a protein construct with 12 AQ units so the relative concentration ratio of  $\alpha$ -helices from AQ units to CBMs is significantly higher than in the simulations. This might result in a higher number of intermolecular  $\alpha$ -helix- $\alpha$ -helix connections and larger intramolecular  $\alpha$ -helix bundles<sup>14</sup> than what can be observed in the simulations employing proteins with three AQ repeats. This also decreases the CBM concentration in comparison to the AQ repeat unit part in general; a longer AQ repeat unit leads to a larger conformational variety.

Attractive interactions between the terminal domains can be expected to immobilize the proteins. Indeed, for dilute solution (single protein simulations) the mean diffusion coefficient

calculated from simulations was  $2.3 \pm 1.3 \times 10^{-7} \text{ cm}^2\text{s}^{-1}$  and for the concentrated solution,  $0.4 \pm 0.3 \times 10^{-7} \text{ cm}^2\text{s}^{-1}$ , see Tables S2 and S3. The diffusion coefficients, determined using the FRAP method for CBM-AQ12-CBM, in the dilute phase surrounding the condensate are  $1.18 \pm 0.08 \times 10^{-7} \text{ cm}^2\text{s}^{-1}$  and in the condensate phase,  $4.5 \pm 0.5 \times 10^{-9} \text{ cm}^2\text{s}^{-1}$ .<sup>14</sup> This significant 400-fold decrease in the experimentally measured diffusion coefficient in the condensate is related with increased intermolecular interactions between proteins, such as attraction between the terminal domains. Notably, the diffusion coefficients calculated from MD simulations are larger than those in the FRAP measurements, because in the simulations a shorter protein construct is used. Also, the simulation timescale is very likely insufficient to fully capture the translational motion of the protein in the concentrated solution. Nevertheless, the values calculated from MD simulations are in line with the FRAP experiments results, and both indicate clearly the presence of intermolecular attractions in the condensate.

The bicontinuous network observed here in the protein condensate provides an explanation for the exceptionally low surface energy, i.e.,  $1.5 \pm 0.5 \text{ } \mu\text{N/m}$ , observed for the condensate droplets in previous study.<sup>14</sup> In particular, as the network rises from individually relatively weak, multiple level connections, the proteins readily reorganize, enter and exit the condensate resulting in a low surface energy. Additionally, a recent study<sup>90</sup> points toward bicontinuous networks in which the interactions between the water and polyelectrolyte/protein are weak, and thus the water mobility is similar to bulk water, having a very low interfacial tension between the phases. To assess our system against these prior reports, we also determined the diffusion coefficients for water molecules in our protein-rich simulation systems and compared them with the diffusion coefficient calculated from bulk water simulations, using the same water model. The water diffusion coefficient in the highly concentrated protein solution here proved to be only 20% lower than the bulk water diffusion coefficient. The small decrease in the diffusion coefficient results rather from steric effects in the network than strong interactions

between water and proteins. Furthermore, in the simulations, the pore size is underestimated in comparison to the experimental system due to using a shorter protein construct in the simulations: for larger pores, the water diffusion coefficient can be expected to increase as larger water droplets approach bulk water in character. Notably, for example strong water-polyelectrolyte attraction immobilizes water effectively even at relatively high water contents.<sup>91–93</sup>

### **Discussion about the liquid-liquid condensate structure and its implications for protein condensates**

The results indicate that the end units of the protein construct act in the condensed protein phase as mechanical load-bearing, stronger connections that carry the main responsibility of the stability and cohesion of the condensate. A secondary contribution to the condensate structural cohesion rises from the  $\alpha$ -helical and transient  $\beta$ -sheet structures that form a variety of weaker connections. Altogether, this leads to the LLPS having multiple-level hierarchical network structure where the basic characteristics of the condensate, such as low surface energy, fast in/out diffusion of the macromolecules and in general, molecular permeability, rise from coexistence of connections with varying magnitudes of load-bearing strengths.

Notably, the mid-part of the protein has compacted shape also in the concentrated protein solution. The compacted shape rises from a high degree of  $\alpha$ -helical bundling between the different AQ units. This folding can be considered to represent randomly folded, disorganized mechanical springs, i.e. provides flexibility and resilience to the condensate, and decreases condensate viscosity. Furthermore, even small mechanical forces will displace these weaker connections and reorganize the  $\alpha$ -helix bundling in the network.

This intra and intermolecular binding structure with sticker interactions of different magnitudes actually provides a mechanistic explanation for the pulling response of silk-inspired fibers from the condensate, as demonstrated in Ref. <sup>14</sup> During pulling fibers from the condensate, the stronger end unit connections carry most of the mechanical load while the resilient, spring-like weaker bonding gives way and reorganizes. For the compacted, perhaps spring-like, middle part, this signifies unfolding to a more extended configuration in the fiber pulling.<sup>87</sup> Simultaneously, a transition of the  $\alpha$ -helical regions to  $\beta$ -sheets, characteristic for bioelastomeric fibers,<sup>94</sup> has been recently reported for this system.<sup>14</sup>

Our simulation results here show also that any formed  $\beta$ -sheet regions are likely to form attractive pairs with each other, or with the remaining  $\alpha$ -helical regions. The formation of a fiber with aligned, extended proteins can thus be envisioned with newly formed intermolecular connections stabilizing the fiber. Yet, as these connections are relatively weak, and able to form and reform, the fiber remains in fluid state while hydrated, where water acts as molecular lubricant and gives space for relaxation of the macromolecules, see e.g. Refs.<sup>91,95</sup> This also explains previously reported seamless fusing of the protein fibers with other similar fibers.<sup>14</sup>

## CONCLUSIONS

Here, we demonstrated via molecular simulations supported with AUC experiments and our earlier <sup>14</sup> high precision cryo-SEM microscopy work findings the detailed structure of a bicontinuous protein network formed under liquid-liquid phase separation in a recombined triblock protein solution. The results show the bicontinuous, sponge-like structure of the condensate phase rises from attraction between the end units of the protein while secondary contributions and compaction of the proteins in the condensate phase rise from smaller attractive regions, stickers, in the flexible, multiconfigurational middle part; these weaker



middle part interactions easily break and re-form again which provides a high flexibility to the network but also indicate a mechanism for protein chain alignment under directed mechanical strain: indeed, the liquid-liquid phase separation in the same triblock protein solution has earlier been demonstrated to be a prerequisite for pulling silk-like fibers.<sup>14</sup>

As the observed behavior rises from attraction between end units separated by a flexible chain combined with lesser, relatively localized attraction sites in the flexible part, the observations are very likely general for macromolecules with similar architecture, including polymers, proteins and peptide constructs. In total, the findings align well with associative polymer models<sup>52,53</sup> and sticker-site considerations in them. Notably, we emphasize that a hierarchical set of interactions, i.e. weaker and stronger attractive stickers, linked by flexible regions are needed to produce the liquid-liquid phase separation, the bicontinuous network, the low surface tension in it, the diffusion responses, and altogether the response experimentally characterized for this protein system. The findings have significance in understanding cellular function<sup>1,4</sup> but also building artificial cells,<sup>96</sup> and engineering advanced biomaterials solutions.<sup>14,17,18</sup> Moreover, the extracted features, via connection with classical associative polymer theories, provide a handle to design and engineering of liquid-liquid phase separating polymer materials for synthetic fiber, tissue engineering, sensor, energy, and drug carrier applications.

## ASSOCIATED CONTENT

**Supporting Information.** Final conformations and conformation time evolution for the different initial configurations of CBM-AQ3-CBM in the 200 ns molecular dynamics simulations are presented as Fig. S1 Fig. S2. presents the corresponding radius of gyration time evolution. Frictional ratio distribution of the CMB-AQ12-CBM proteins as determined from

the analytical ultracentrifuge (AUC) data is available as Fig. S3. In Fig. S4, the final 200 ns molecular dynamics simulation conformations resulting from the different initial configurations of concentrated CBM-AQ3-CBM solution (200 mg/ml) and their contact maps are presented. Finally, Table S1 lists the number of repeat runs corresponding to each simulation system.

## **AUTHOR INFORMATION**

### **Corresponding Author**

\*Piotr Batys (piotr.batys@ikifp.edu.pl), \*Maria Sammalkorpi (maria.sammalkorpi@aalto.fi)

### **Author Contributions**

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

## **ACKNOWLEDGMENTS**

This work was supported by the Academy of Finland through its Centres of Excellence Programme (2014-2019, HYBER) under Project No. 307474 and Academy of Finland Project No. 309324 (M.S.). The work was supported also by Jenny and Antti Wihuri Foundation (Centre for Young Synbio Scientists) (P.M.). We are grateful for the support by the FinnCERES Materials Bioeconomy Ecosystem and use of the Bioeconomy Infrastructure at Aalto University. Computational resources by CSC IT Centre for Science, Finland, and RAMI – RawMatTERS Finland Infrastructure are also gratefully acknowledged.

## **REFERENCES**

- (1) Boeynaems, S.; Alberti, S.; Fawzi, N. L.; Mittag, T.; Polymenidou, M.; Rousseau, F.; Schymkowitz, J.; Shorter, J.; Wolozin, B.; Van Den Bosch, L.; et al. Protein Phase Separation: A New Phase in Cell Biology. *Trends in Cell Biology*. Elsevier Current Trends June 1, 2018, pp 420–435. <https://doi.org/10.1016/j.tcb.2018.02.004>.
- (2) Perry, S. L. Phase Separation: Bridging Polymer Physics and Biology. *Current Opinion in Colloid and Interface Science*. Elsevier February 1, 2019, pp 86–97. <https://doi.org/10.1016/j.cocis.2019.01.007>.
- (3) Style, R. W.; Sai, T.; Fanelli, N.; Ijavi, M.; Smith-Mannschott, K.; Xu, Q.; Wilen, L. A.; Dufresne, E. R. Liquid-Liquid Phase Separation in an Elastic Network. *Phys. Rev. X* **2018**, 8 (1), 011028. <https://doi.org/10.1103/PhysRevX.8.011028>.
- (4) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular Condensates: Organizers of Cellular Biochemistry. *Nature Reviews Molecular Cell Biology*. Nature Publishing Group May 22, 2017, pp 285–298. <https://doi.org/10.1038/nrm.2017.7>.
- (5) Wei, M. T.; Elbaum-Garfinkle, S.; Holehouse, A. S.; Chen, C. C. H.; Feric, M.; Arnold, C. B.; Priestley, R. D.; Pappu, R. V.; Brangwynne, C. P. Phase Behaviour of Disordered Proteins Underlying Low Density and High Permeability of Liquid Organelles. *Nat. Chem.* **2017**, 9 (11), 1118–1125. <https://doi.org/10.1038/NCHEM.2803>.
- (6) Aumiller, W. M.; Keating, C. D. Phosphorylation-Mediated RNA/Peptide Complex Coacervation as a Model for Intracellular Liquid Organelles. *Nat. Chem.* **2016**, 8 (2), 129–137. <https://doi.org/10.1038/nchem.2414>.
- (7) Pak, C. W.; Kosno, M.; Holehouse, A. S.; Padrick, S. B.; Mittal, A.; Ali, R.; Yunus, A. A.; Liu, D. R.; Pappu, R. V.; Rosen, M. K. Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol. Cell* **2016**,

- 63 (1), 72–85. <https://doi.org/10.1016/j.molcel.2016.05.042>.
- (8) Bratek-Skicki, A.; Pancsa, R.; Meszaros, B.; Van Lindt, J.; Tompa, P. A Guide to Regulation of the Formation of Biomolecular Condensates. *FEBS Journal*. Blackwell Publishing Ltd May 14, 2020, pp 1924–1935. <https://doi.org/10.1111/febs.15254>.
  - (9) Riback, J. A.; Katanski, C. D.; Kear-Scott, J. L.; Pilipenko, E. V.; Rojek, A. E.; Sosnick, T. R.; Drummond, D. A. Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* **2017**, *168* (6), 1028–1040.e19. <https://doi.org/10.1016/j.cell.2017.02.027>.
  - (10) Mitrea, D. M.; Kriwacki, R. W. Phase Separation in Biology; Functional Organization of a Higher Order. *Cell Communication and Signaling*. BioMed Central Ltd. January 5, 2016, pp 1–20. <https://doi.org/10.1186/s12964-015-0125-7>.
  - (11) Elbaum-Garfinkle, S. Matter over Mind: Liquid Phase Separation and Neurodegeneration. *Journal of Biological Chemistry*. 2019, pp 7160–7168. <https://doi.org/10.1074/jbc.REV118.001188>.
  - (12) Reichheld, S. E.; Muiznieks, L. D.; Keeley, F. W.; Sharpe, S. Direct Observation of Structure and Dynamics during Phase Separation of an Elastomeric Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (22), E4408–E4415. <https://doi.org/10.1073/pnas.1701877114>.
  - (13) Martin, E. W.; Holehouse, A. S.; Peran, I.; Farag, M.; Incicco, J. J.; Bremer, A.; Grace, C. R.; Soranno, A.; Pappu, R. V.; Mittag, T. Valence and Patterning of Aromatic Residues Determine the Phase Behavior of Prion-like Domains. *Science (80-. )*. **2020**, *367* (6478), 694–699. <https://doi.org/10.1126/science.aaw8653>.
  - (14) Mohammadi, P.; Aranko, A. S.; Lemetti, L.; Cenev, Z.; Zhou, Q.; Virtanen, S.;

- Landowski, C. P.; Penttilä, M.; Fischer, W. J.; Wagermaier, W.; et al. Phase Transitions as Intermediate Steps in the Formation of Molecularly Engineered Protein Fibers. *Commun. Biol.* **2018**, *1* (1), 86. <https://doi.org/10.1038/s42003-018-0090-y>.
- (15) Lemetti, L.; Hirvonen, S.-P.; Fedorov, D.; Batys, P.; Sammalkorpi, M.; Tenhy, H.; Linder, M. B.; Aranko, A. S. Molecular Crowding Facilitates Assembly of Spidroin-like Proteins through Phase Separation. *Eur. Polym. J.* **2018**. <https://doi.org/10.1016/j.eurpolymj.2018.10.010>.
- (16) Waite, J. H. Mussel Adhesion - Essential Footwork. *Journal of Experimental Biology*. 2017, pp 517–530. <https://doi.org/10.1242/jeb.134056>.
- (17) Mohammadi, P.; Beaune, G.; Stokke, B. T.; Timonen, J. V. I.; Linder, M. B. Self-Coacervation of a Silk-Like Protein and Its Use As an Adhesive for Cellulosic Materials. *ACS Macro Lett.* **2018**, *7* (9), 1120–1125. <https://doi.org/10.1021/acsmacrolett.8b00527>.
- (18) Kim, H. J.; Yang, B.; Park, T. Y.; Lim, S.; Cha, H. J. Complex Coacervates Based on Recombinant Mussel Adhesive Proteins: Their Characterization and Applications. *Soft Matter* **2017**, *13* (42), 7704–7716. <https://doi.org/10.1039/c7sm01735a>.
- (19) Bahn, S. Y.; Jo, B. H.; Choi, Y. S.; Cha, H. J. Control of Nacre Biomineralization by Pif80 in Pearl Oyster. *Sci. Adv.* **2017**, *3* (8), e1700765. <https://doi.org/10.1126/sciadv.1700765>.
- (20) Podsiadlo, P.; Kaushik, A. K.; Arruda, E. M.; Waas, A. M.; Shim, B. S.; Xu, J.; Nandivada, H.; Pumplin, B. G.; Lahann, J.; Ramamoorthy, A.; et al. Ultrastrong and Stiff Layered Polymer Nanocomposites. *Science (80-. )*. **2007**, *318* (5847), 80–83. <https://doi.org/10.1126/science.1143176>.

- (21) Muiznieks, L. D.; Keeley, F. W. Phase Separation and Mechanical Properties of an Elastomeric Biomaterial from Spider Wrapping Silk and Elastin Block Copolymers. *Biopolymers* **2016**, *105* (10), 693–703. <https://doi.org/10.1002/bip.22888>.
- (22) Hermida-Merino, D.; O'Driscoll, B.; Hart, L. R.; Harris, P. J.; Colquhoun, H. M.; Slark, A. T.; Prisacariu, C.; Hamley, I. W.; Hayes, W. Enhancement of Microphase Ordering and Mechanical Properties of Supramolecular Hydrogen-Bonded Polyurethane Networks. *Polym. Chem.* **2018**, *9* (24), 3406–3414. <https://doi.org/10.1039/c8py00604k>.
- (23) Zhai, L.; Cebeci, F. C.; Cohen, R. E.; Rubner, M. F. Stable Superhydrophobic Coatings from Polyelectrolyte Multilayers. *Nano Lett.* **2004**, *4* (7), 1349–1353. <https://doi.org/10.1021/nl049463j>.
- (24) Stuart, M. A. C.; Huck, W. T. S.; Genzer, J.; Müller, M.; Ober, C.; Stamm, M.; Sukhorukov, G. B.; Szleifer, I.; Tsukruk, V. V.; Urban, M.; et al. Emerging Applications of Stimuli-Responsive Polymer Materials. *Nat. Mater.* **2010**, *9* (2), 101–113. <https://doi.org/10.1038/nmat2614>.
- (25) Gucht, J. van der; Spruijt, E.; Lemmers, M.; Cohen Stuart, M. A. Polyelectrolyte Complexes: Bulk Phases and Colloidal Systems. *J. Colloid Interface Sci.* **2011**, *361* (2), 407–422. <https://doi.org/10.1016/J.JCIS.2011.05.080>.
- (26) Zhang, Y.; Yildirim, E.; Antila, H. S.; Valenzuela, L. D.; Sammalkorpi, M.; Lutkenhaus, J. L. The Influence of Ionic Strength and Mixing Ratio on the Colloidal Stability of PDAC/PSS Polyelectrolyte Complexes. *Soft Matter* **2015**, *11* (37), 7392–7401. <https://doi.org/10.1039/C5SM01184A>.
- (27) Tang, J. D.; Caliari, S. R.; Lampe, K. J. Temperature-Dependent Complex Coacervation of Engineered Elastin-like Polypeptide and Hyaluronic Acid

- Polyelectrolytes. *Biomacromolecules* **2018**, *19* (10), 3925–3935.  
<https://doi.org/10.1021/acs.biomac.8b00837>.
- (28) Sui, Z.; Jaber, J. A.; Schlenoff, J. B. Polyelectrolyte Complexes with PH-Tunable Solubility. *Macromolecules* **2006**, *39* (23), 8145–8152.  
<https://doi.org/10.1021/MA061098Q>.
- (29) Meng, S.; Liu, Y.; Yeo, J.; Ting, J. M.; Tirrell, M. V. Effect of Mixed Solvents on Polyelectrolyte Complexes with Salt. *Colloid Polym. Sci.* **2020**, *298* (7), 887–894.  
<https://doi.org/10.1007/s00396-020-04637-0>.
- (30) Hyman, A. A.; Weber, C. A.; Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annual review of cell and developmental biology*. Annual Reviews October 6, 2014, pp 39–58. <https://doi.org/10.1146/annurev-cellbio-100913-013325>.
- (31) Berry, J.; Brangwynne, C. P.; Haataja, M. Physical Principles of Intracellular Organization via Active and Passive Phase Transitions. *Reports on Progress in Physics*. IOP Publishing April 1, 2018, p 046601. <https://doi.org/10.1088/1361-6633/aaa61e>.
- (32) Shin, Y.; Brangwynne, C. P. Liquid Phase Condensation in Cell Physiology and Disease. *Science (80-. )*. **2017**, *357* (6357), eaaf4382.  
<https://doi.org/10.1126/SCIENCE.AAF4382>.
- (33) Peran, I.; Mittag, T. Molecular Structure in Biomolecular Condensates. *Current Opinion in Structural Biology*. Elsevier Ltd February 1, 2020, pp 17–26.  
<https://doi.org/10.1016/j.sbi.2019.09.007>.
- (34) Brangwynne, C. P.; Tompa, P.; Pappu, R. V. Polymer Physics of Intracellular Phase Transitions. *Nat. Phys.* **2015**, *11* (11), 899–904. <https://doi.org/10.1038/nphys3532>.

- (35) Uversky, V. N. Protein Intrinsic Disorder-Based Liquid–liquid Phase Transitions in Biological Systems: Complex Coacervates and Membrane-Less Organelles. *Advances in Colloid and Interface Science*. Elsevier January 1, 2017, pp 97–114.  
<https://doi.org/10.1016/j.cis.2016.05.012>.
- (36) Uversky, V. N. Intrinsically Disordered Proteins in Overcrowded Milieu: Membrane-Less Organelles, Phase Separation, and Intrinsic Disorder. *Current Opinion in Structural Biology*. Elsevier Current Trends June 1, 2017, pp 18–30.  
<https://doi.org/10.1016/j.sbi.2016.10.015>.
- (37) Nott, T. J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowietz, A.; Craggs, T. D.; Bazett-Jones, D. P.; Pawson, T.; Forman-Kay, J. D.; et al. Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Mol. Cell* **2015**, 57 (5), 936–947.  
<https://doi.org/10.1016/j.molcel.2015.01.013>.
- (38) Lai, J.; Koh, C. H.; Tjota, M.; Pieuchot, L.; Raman, V.; Chandrababu, K. B.; Yang, D.; Wong, L.; Jedd, G. Intrinsically Disordered Proteins Aggregate at Fungal Cell-to-Cell Channels and Regulate Intercellular Connectivity. *Proc. Natl. Acad. Sci.* **2012**, 109 (39), 15781–15786. <https://doi.org/10.1073/pnas.1207467109>.
- (39) Das, R. K.; Pappu, R. V. Conformations of Intrinsically Disordered Proteins Are Influenced by Linear Sequence Distributions of Oppositely Charged Residues. *Proc. Natl. Acad. Sci.* **2013**, 110 (33), 13392–13397.  
<https://doi.org/10.1073/pnas.1304749110>.
- (40) Mohammadi, P.; Jonkergouw, C.; Beaune, G.; Engelhardt, P.; Kamada, A.; Timonen, J. V. I.; Knowles, T. P. J.; Penttilä, M.; Linder, M. B. Controllable Coacervation of Recombinantly Produced Spider Silk Protein Using Kosmotropic Salts. *J. Colloid*



- Interface Sci.* **2020**, *560*, 149–160. <https://doi.org/10.1016/j.jcis.2019.10.058>.
- (41) Batys, P.; Nattich-Rak, M.; Adamczyk, Z. Myoglobin Molecule Charging in Electrolyte Solutions. *Phys. Chem. Chem. Phys.* **2020**, *22* (46), 26764–26775. <https://doi.org/10.1039/D0CP03771K>.
- (42) Zhang, Y.; Cremer, P. S. Chemistry of Hofmeister Anions and Osmolytes. *Annu. Rev. Phys. Chem.* **2010**, *61* (1), 63–83. <https://doi.org/10.1146/annurev.physchem.59.032607.093635>.
- (43) Okur, H. I.; Hladílková, J.; Rembert, K. B.; Cho, Y.; Heyda, J.; Dzubiella, J.; Cremer, P. S.; Jungwirth, P. Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions. *J. Phys. Chem. B* **2017**, *121* (9), 1997–2014. <https://doi.org/10.1021/acs.jpcb.6b10797>.
- (44) Martin, E. W.; Mittag, T. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry* **2018**, *57* (17), 2478–2487. <https://doi.org/10.1021/acs.biochem.8b00008>.
- (45) Holehouse, A. S.; Garai, K.; Lyle, N.; Vitalis, A.; Pappu, R. V. Quantitative Assessments of the Distinct Contributions of Polypeptide Backbone Amides versus Side Chain Groups to Chain Expansion via Chemical Denaturation. *J. Am. Chem. Soc.* **2015**, *137* (8), 2984–2995. <https://doi.org/10.1021/ja512062h>.
- (46) Asthagiri, D.; Karandur, D.; Tomar, D. S.; Pettitt, B. M. Intramolecular Interactions Overcome Hydration to Drive the Collapse Transition of Gly15. *J. Phys. Chem. B* **2017**, *121* (34), 8078–8084. <https://doi.org/10.1021/acs.jpcb.7b05469>.
- (47) Mittag, T.; Parker, R. Multiple Modes of Protein–Protein Interactions Promote RNP Granule Assembly. *Journal of Molecular Biology*. Academic Press November 2, 2018,

pp 4636–4649. <https://doi.org/10.1016/j.jmb.2018.08.005>.

- (48) Li, P.; Banjade, S.; Cheng, H. C.; Kim, S.; Chen, B.; Guo, L.; Llaguno, M.; Hollingsworth, J. V.; King, D. S.; Banani, S. F.; et al. Phase Transitions in the Assembly of Multivalent Signalling Proteins. *Nature* **2012**, *483* (7389), 336–340. <https://doi.org/10.1038/nature10879>.
- (49) Hughes, M. P.; Sawaya, M. R.; Boyer, D. R.; Goldschmidt, L.; Rodriguez, J. A.; Cascio, D.; Chong, L.; Gonen, T.; Eisenberg, D. S. Atomic Structures of Low-Complexity Protein Segments Reveal Kinked  $\beta$  Sheets That Assemble Networks. *Science* (80-. ). **2018**, *359* (6376), 698–701. <https://doi.org/10.1126/science.aan6398>.
- (50) Rao, B. S.; Parker, R. Numerous Interactions Act Redundantly to Assemble a Tunable Size of P Bodies in *Saccharomyces Cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (45), E9569–E9578. <https://doi.org/10.1073/pnas.1712396114>.
- (51) Lin, Y. H.; Forman-Kay, J. D.; Chan, H. S. Theories for Sequence-Dependent Phase Behaviors of Biomolecular Condensates. *Biochemistry*. American Chemical Society May 1, 2018, pp 2499–2508. <https://doi.org/10.1021/acs.biochem.8b00058>.
- (52) Rubinstein, M.; V.Dobrynin, A. Solutions of Associative Polymers. *Trends Polym. Sci.* **1997**, *5*, 181–186.
- (53) Dobrynin, A. V. Phase Diagram of Solutions of Associative Polymers. *Macromolecules* **2004**, *37* (10), 3881–3893. <https://doi.org/10.1021/ma0353686>.
- (54) Harmon, T. S.; Holehouse, A. S.; Rosen, M. K.; Pappu, R. V. Intrinsically Disordered Linkers Determine the Interplay between Phase Separation and Gelation in Multivalent Proteins. *Elife* **2017**, *6*, e30294. <https://doi.org/10.7554/eLife.30294>.
- (55) Zeng, X.; Holehouse, A. S.; Chilkoti, A.; Mittag, T.; Pappu, R. V. Connecting Coil-to-

- Globule Transitions to Full Phase Diagrams for Intrinsically Disordered Proteins. *Biophys. J.* **2020**, *119* (2), 402–418. <https://doi.org/10.1016/j.bpj.2020.06.014>.
- (56) Choi, J.-M.; Dar, F.; Pappu, R. V. LASSI: A Lattice Model for Simulating Phase Transitions of Multivalent Proteins. *PLOS Comput. Biol.* **2019**, *15* (10), e1007028. <https://doi.org/10.1371/journal.pcbi.1007028>.
- (57) Falahati, H.; Haji-Akbari, A. Thermodynamically Driven Assemblies and Liquid-Liquid Phase Separations in Biology. *Soft Matter*. Royal Society of Chemistry February 6, 2019, pp 1135–1154. <https://doi.org/10.1039/c8sm02285b>.
- (58) Dignon, G. L.; Zheng, W.; Mittal, J. Simulation Methods for Liquid–liquid Phase Separation of Disordered Proteins. *Current Opinion in Chemical Engineering*. Elsevier Ltd March 1, 2019, pp 92–98. <https://doi.org/10.1016/j.coche.2019.03.004>.
- (59) Shin, Y.; Berry, J.; Pannucci, N.; Haataja, M. P.; Toettcher, J. E.; Brangwynne, C. P. Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated OptoDroplets. *Cell* **2017**, *168* (1–2), 159–171.e14. <https://doi.org/10.1016/j.cell.2016.11.054>.
- (60) Mao, S.; Kuldinow, D.; Haataja, M. P.; Košmrlj, A. Phase Behavior and Morphology of Multicomponent Liquid Mixtures. *Soft Matter* **2019**, *15* (6), 1297–1311. <https://doi.org/10.1039/c8sm02045k>.
- (61) Jacobs, W. M.; Frenkel, D. Phase Transitions in Biological Systems with Many Components. *Biophys. J.* **2017**, *112* (4), 683–691. <https://doi.org/10.1016/j.bpj.2016.10.043>.
- (62) Dignon, G. L.; Zheng, W.; Best, R. B.; Kim, Y. C.; Mittal, J. Relation between Single-Molecule Properties and Phase Behavior of Intrinsically Disordered Proteins. *Proc.*

- Natl. Acad. Sci. U. S. A.* **2018**, *115* (40), 9929–9934.  
<https://doi.org/10.1073/pnas.1804177115>.
- (63) Feric, M.; Vaidya, N.; Harmon, T. S.; Mitrea, D. M.; Zhu, L.; Richardson, T. M.; Kriwacki, R. W.; Pappu, R. V.; Brangwynne, C. P. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **2016**, *165* (7), 1686–1697.  
<https://doi.org/10.1016/j.cell.2016.04.047>.
- (64) Ryan, V. H.; Dignon, G. L.; Zerze, G. H.; Chabata, C. V.; Silva, R.; Conicella, A. E.; Amaya, J.; Burke, K. A.; Mittal, J.; Fawzi, N. L. Mechanistic View of HnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol. Cell* **2018**, *69* (3), 465–479.e7.  
<https://doi.org/10.1016/j.molcel.2017.12.022>.
- (65) Murthy, A. C.; Dignon, G. L.; Kan, Y.; Zerze, G. H.; Parekh, S. H.; Mittal, J.; Fawzi, N. L. Molecular Interactions Underlying Liquid–liquid Phase Separation of the FUS Low-Complexity Domain. *Nat. Struct. Mol. Biol.* **2019**, *26* (7), 637–648.  
<https://doi.org/10.1038/s41594-019-0250-x>.
- (66) Alshareedah, I.; Moosa, M. M.; Raju, M.; Potoyan, D. A.; Banerjee, P. R. Phase Transition of RNA–protein Complexes into Ordered Hollow Condensates. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (27), 15650–15658.  
<https://doi.org/10.1073/pnas.1922365117>.
- (67) Kaur, T.; Raju, M.; Alshareedah, I.; Davis, R. B.; Potoyan, D. A.; Banerjee, P. R. Sequence-Encoded and Composition-Dependent Protein-RNA Interactions Control Multiphasic Condensate Topologies. *bioRxiv* **2020**, 2020.08.30.273748.  
<https://doi.org/10.1101/2020.08.30.273748>.
- (68) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-

- Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1–3), 43–56. [https://doi.org/10.1016/0010-4655\(95\)00042-E](https://doi.org/10.1016/0010-4655(95)00042-E).
- (69) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A Package for Molecular Simulation and Trajectory Analysis. *Journal of Molecular Modeling*. 2001, pp 306–317. <https://doi.org/10.1007/S008940100045>.
- (70) Best, R. B.; Zheng, W.; Mittal, J. Balanced Protein-Water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. *J. Chem. Theory Comput.* **2014**, *10* (11), 5113–5124. <https://doi.org/10.1021/ct500569b>.
- (71) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123* (23), 234505. <https://doi.org/10.1063/1.2121687>.
- (72) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H).
- (73) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13* (8), 952–962. <https://doi.org/10.1002/jcc.540130805>.
- (74) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101. <https://doi.org/10.1063/1.2408420>.
- (75) Parrinello, M. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182. <https://doi.org/10.1063/1.328693>.

- (76) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
- (77) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
- (78) Yaniv, O.; Morag, E.; Borovok, I.; Bayer, E. A.; Lamed, R.; Frolow, F.; Shimon, L. J. W. Structure of a Family 3a Carbohydrate-Binding Module from the Cellulosomal Scaffoldin CipA of *Clostridium Thermocellum* with Flanking Linkers: Implications for Cellulosome Structure. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2013**, *69* (7), 733–737. <https://doi.org/10.1107/S174430911301614X>.
- (79) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* **2012**, *4* (1), 17. <https://doi.org/10.1186/1758-2946-4-17>.
- (80) Slotta, U.; Tammer, M.; Kremer, F.; Koelsch, P.; Scheibel, T. Structural Analysis of Spider Silk Films. *Supramol. Chem.* **2006**, *18* (5), 465–471. <https://doi.org/10.1080/10610270600832042>.
- (81) Brookes, E.; Demeler, B. Genetic Algorithm Optimization for Obtaining Accurate Molecular Weight Distributions from Sedimentation Velocity Experiments. *Prog. Colloid Polym. Sci.* **2006**, *131*, 33–40. [https://doi.org/10.1007/2882\\_004](https://doi.org/10.1007/2882_004).
- (82) Brookes, E. H.; Demeler, B. Parsimonious Regularization Using Genetic Algorithms Applied to the Analysis of Analytical Ultracentrifugation Experiments. *Proc. 9th Annu. Conf. Genet. Evol. Comput. - GECCO '07* **2007**, 361. <https://doi.org/10.1145/1276958.1277035>.

- (83) Fedorov, D.; Batys, P.; Hayes, D. B.; Sammalkorpi, M.; Linder, M. B. Analyzing the Weak Dimerization of a Cellulose Binding Module by Analytical Ultracentrifugation. *Int. J. Biol. Macromol.* **2020**, *163*, 1995–2004.  
<https://doi.org/10.1016/j.ijbiomac.2020.09.054>.
- (84) Brookes, E.; Demeler, B.; Rosano, C.; Rocco, M. The Implementation of SOMO ( SOlution MOdeller ) in the UltraScan Analytical Ultracentrifugation Data Analysis Suite : Enhanced Capabilities Allow the Reliable Hydrodynamic Modeling of Virtually Any Kind of Biomacromolecule. *Eur. Biophys. J.* **2010**, 423–435.  
<https://doi.org/10.1007/s00249-009-0418-0>.
- (85) Pace, C. N.; Scholtz, J. M. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophys. J.* **1998**, *75* (1), 422–427.  
[https://doi.org/10.1016/s0006-3495\(98\)77529-0](https://doi.org/10.1016/s0006-3495(98)77529-0).
- (86) Mohammadi, P.; Sesilja Aranko, A.; Landowski, C. P.; Ikkala, O.; Jaudzems, K.; Wagermaier, W.; Linder, M. B. Biomimetic Composites with Enhanced Toughening Using Silk-Inspired Triblock Proteins and Aligned Nanocellulose Reinforcements. *Sci. Adv.* **2019**, *5* (9), eaaw2541. <https://doi.org/10.1126/sciadv.aaw2541>.
- (87) Giesa, T.; Perry, C. C.; Buehler, M. J. Secondary Structure Transition and Critical Stress for a Model of Spider Silk Assembly. *Biomacromolecules* **2016**, *17* (2), 427–436. <https://doi.org/10.1021/acs.biomac.5b01246>.
- (88) Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W. NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* **2012**, *9* (7), 671–675.
- (89) Fedorov, D.; Batys, P.; Hayes, D. B.; Sammalkorpi, M.; Linder, M. B. Analyzing the Weak Dimerization of a Cellulose Binding Module by Analytical Ultracentrifugation. *Int. J. Biol. Macromol.* **2020**, *163*, 1995–2004.

<https://doi.org/10.1016/j.ijbiomac.2020.09.054>.

- (90) Huang, K.-Y.; Yoo, H. Y.; Jho, Y.; Han, S.; Hwang, D. S. Bicontinuous Fluid Structure with Low Cohesive Energy: Molecular Basis for Exceptionally Low Interfacial Tension of Complex Coacervate Fluids. *ACS Nano* **2016**, *10* (5), 5051–5062. <https://doi.org/10.1021/acsnano.5b07787>.
- (91) Batys, P.; Zhang, Y.; Lutkenhaus, J. L.; Sammalkorpi, M. Hydration and Temperature Response of Water Mobility in Poly(Diallyldimethylammonium)-Poly(Sodium 4-Styrenesulfonate) Complexes. *Macromolecules* **2018**, *51* (20), 8268–8277. <https://doi.org/10.1021/acs.macromol.8b01441>.
- (92) Batys, P.; Kivistö, S.; Lalwani, S. M.; Lutkenhaus, J. L.; Sammalkorpi, M. Comparing Water-Mediated Hydrogen-Bonding in Different Polyelectrolyte Complexes. *Soft Matter* **2019**, *15*, 7823–7831. <https://doi.org/10.1039/c9sm01193e>.
- (93) Eneh, C. I.; Bolen, M. J.; Suarez-Martinez, P. C.; Bachmann, A. L.; Zimudzi, T. J.; Hickner, M. A.; Batys, P.; Sammalkorpi, M.; Lutkenhaus, J. L. Fourier Transform Infrared Spectroscopy Investigation of Water Microenvironments in Polyelectrolyte Multilayers at Varying Temperatures. *Soft Matter* **2020**, *16* (9), 2291–2300. <https://doi.org/10.1039/c9sm02478f>.
- (94) Miserez, A.; Guerette, P. A. Phase Transition-Induced Elasticity of  $\alpha$ -Helical Bioelastomeric Fibres and Networks. *Chem. Soc. Rev.* **2013**, *42* (5), 1973–1995. <https://doi.org/10.1039/C2CS35294J>.
- (95) Zhang, R.; Zhang, Y.; Antila, H. S.; Lutkenhaus, J. L.; Sammalkorpi, M. Role of Salt and Water in the Plasticization of PDAC/PSS Polyelectrolyte Assemblies. *J. Phys. Chem. B* **2017**, *121* (1), 322–333. <https://doi.org/10.1021/acs.jpcc.6b12315>.



- (96) Buddingh', B. C.; Van Hest, J. C. M. Artificial Cells: Synthetic Compartments with Life-like Functionality and Adaptivity. *Acc. Chem. Res.* **2017**, *50* (4), 769–777.  
<https://doi.org/10.1021/acs.accounts.6b00512>.