

# Sequence-Dependent Conformational Transitions of Disordered Proteins During Condensation

Jiahui Wang,<sup>1</sup> Dinesh Sundaravadivelu Devarajan,<sup>1</sup> Young C. Kim,<sup>2</sup> Arash Nikoubashman,<sup>3,4,5</sup> and Jeetain Mittal<sup>1,6,7,\*</sup>

<sup>1</sup>Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, United States

<sup>2</sup>Center for Materials Physics and Technology, Naval Research Laboratory, Washington, DC 20375, United States

<sup>3</sup>Leibniz-Institut für Polymerforschung Dresden e.V., Hohe Straße 6, 01069 Dresden, Germany

<sup>4</sup>Institut für Theoretische Physik, Technische Universität Dresden, 01069 Dresden, Germany

<sup>5</sup>Cluster of Excellence Physics of Life, Technische Universität Dresden, 01062 Dresden, Germany

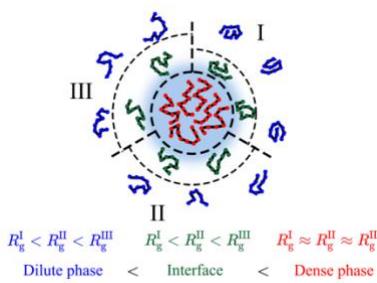
<sup>6</sup>Department of Chemistry, Texas A&M University, College Station, TX 77843, United States

<sup>7</sup>Interdisciplinary Graduate Program in Genetics and Genomics, Texas A&M University, College Station, TX 77843, United States

\*Corresponding author email: [jeetain@tamu.edu](mailto:jeetain@tamu.edu)

**Abstract:** Intrinsically disordered proteins (IDPs) can form biomolecular condensates through phase separation. It is recognized that the conformation of IDPs in the dense and dilute phases as well as at the interfaces of condensates can critically impact the resulting properties associated with their functionality. However, a comprehensive understanding of the conformational transitions of IDPs during condensation remains elusive. In this study, we employ a coarse-grained polyampholyte model, comprising an equal number of oppositely charged residues—glutamic acid and lysine—whereby conformations and phase behavior can be readily tuned by altering the protein sequence. By manipulating the sequence patterns from perfectly alternating to block-like, we obtain chains with ideal-like conformations to semi-compact structures in the dilute phase, while in the dense phase, the chain conformation is approximately that of an ideal chain, irrespective of the protein sequence. By performing simulations at different concentrations, we find that the chains assemble from the dilute phase through small oligomeric clusters to the dense phase, accompanied by a gradual swelling of the individual chains. We further demonstrate that these findings are applicable to several naturally occurring proteins involved in the formation of biological condensates. Concurrently, we delve deeper into the chain conformations within the condensate, revealing that chains at the interface show a strong sequence dependence, but remain more collapsed than those in the bulk-like dense phase. This study addresses critical gaps in our knowledge of IDP conformations within condensates as a function of protein sequence.

## TOC graphics



**Keywords:** Biomolecular condensates, Intrinsically disordered proteins, Protein conformations, Phase separation, Molecular simulations

Liquid-liquid phase separation (LLPS) is crucial for the formation of biomolecular condensates, which in turn play an important role in vital cellular mechanisms such as gene expression, signal transduction, stress response, and the assembly of macromolecular complexes<sup>1–5</sup>. Intrinsically disordered proteins (IDPs) and disordered regions within proteins have been identified as a primary driving force for the formation of the condensed phase in many cases. It is recognized that the conformation of IDPs in biomolecular condensates and at their interface can critically impact the resulting functionality. For example, recent studies highlighted that the conformational changes of Dcp2 within P-bodies influence mRNA decapping<sup>6</sup>, and conformational expansion of Tau in condensates promotes fibrillization<sup>7</sup>. These results emphasize the important role of protein conformations in dictating the function of biomolecular condensates.

Extensive research has been conducted to study protein conformations in the dilute phase<sup>8,9</sup>, elucidating that the conformation of IDPs can be modulated by factors such as sequence composition<sup>10–15</sup>, charge characteristics<sup>16–19</sup>, sequence pattern<sup>20–24</sup> and solvent environment<sup>25–28</sup>. Sequences comprised of charged residues exhibit a globule-to-coil transition with an increase in the net charge per residue<sup>17</sup>. Quantitative analyses of charge patterns have demonstrated that enhanced charge segregation in the sequence typically leads to more compact conformations in the dilute phase<sup>20,21</sup>. Conformational changes are also modulated by varying electrostatic interactions due to the surrounding solvent environment. Recently, Reddy *et al.* illustrated that a pH shift from neutral to acidic prompted Prothymosin- $\alpha$  to shift from a random coil to a partially collapsed state<sup>29</sup>.

In the dense phase, it has been observed experimentally that  $\alpha$ -synuclein shifts towards more “elongated” conformations during LLPS<sup>30</sup>, and Tau K18 exhibits expanded conformations within the droplet phase, in contrast to a compact structural ensemble in the dilute phase<sup>31</sup>. A1-LCD protein and its mutated variants have been shown to adopt more extended conformations in the

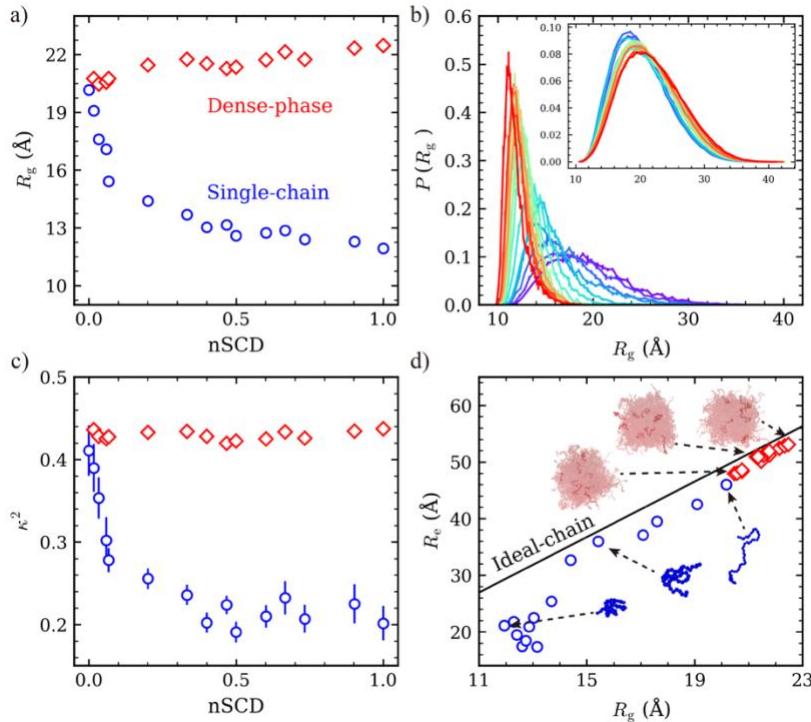
condensates as well through molecular simulations<sup>32</sup>. Despite these excellent prior studies on the conformations of IDPs in the dense phase, there is a general lack of understanding of how the protein conformations change when transitioning from the dilute to the condensed phase; are they always more expanded in condensates than in the dilute phase? What are the polymer scaling properties of IDPs in the dense phase and at the interface, and how do these depend on the protein sequence and dilute phase conformations?

To answer these questions and to decipher the conformational transitions from the dilute to the dense phase during LLPS, we systematically studied a wide range of polyampholyte sequences, which exist in many naturally occurring IDPs<sup>33,34</sup>. Further, these polyampholyte sequences are ideal model systems, since their conformations, as measured by the radius of gyration ( $R_g$ ), can be readily modulated by altering the charge pattern<sup>20,21,35,36</sup>. In particular, we selected 15 sequences composed of glutamic acid (E) and lysine (K) residues (**Figure S1**). Each E-K variant (EKV) consists of 50 residues with an equal number of E and K residues to keep a zero net charge. The degree of charge segregation of EKVs was quantified by the sequence charge decoration (SCD) parameter<sup>21,37</sup>. To allow for a more straightforward comparison with other IDPs, we normalized the SCD values (nSCD) so that nSCD = 0 for the uniformly charge-patterned sequence and nSCD = 1 for the most charge-segregated diblock sequence (normalization method shown in SI text)<sup>38</sup>. The selected 15 EKVs cover diverse conformations in dilute solutions (single chain) at constant temperature ( $T = 300$  K), ranging from an ideal-like chain (EKV1) to a semi-compact chain (EKV15). This selection allowed us to probe the conformational properties within the dense phase starting with different chain configurations in the dilute phase. First, we analyzed the conformations for EKVs in the dense phase and as a single chain (**Figure 1**) by calculating the root mean square radius of gyration ( $R_g$ ) from the average of the trace of the gyration tensor ( $G$ ):

$$G_{ij} = \frac{1}{M} \sum_{n=1}^N m_n \Delta \mathbf{r}_n^i \Delta \mathbf{r}_n^j, \quad i, j = 1, 2, 3 \quad (1)$$

$$R_g = \langle \text{tr} \mathbf{G} \rangle^{1/2}, \quad (2)$$

where  $M$  is the chain's total mass,  $m_n$  is the residue's mass, and  $\Delta \mathbf{r}_n^i$  is the vector pointing from the chain's center of mass to residue  $n$ , while  $i$  and  $j$  are the components in the Cartesian  $x$ ,  $y$ , and  $z$  directions.



**Figure 1.** Conformational properties of EKVs in the dense phase (red diamonds) and as a single chain (blue circles). (a) Radius of gyration ( $R_g$ ) as a function of nSCD. (b) Probability distribution of  $R_g$ ,  $P(R_g)$ , for EKVs as a single chain. The inset shows  $P(R_g)$  for EKVs in the dense phase. Line colors, ranging from purple to red, indicate increasing nSCD. (c) Relative shape anisotropy ( $\kappa^2$ ) as a function of nSCD. (d) The correlation between end-to-end distance ( $R_e$ ) and  $R_g$  for the EKVs compared to an ideal chain (solid line). Snapshots show the conformations in the dense phase (red) and as single chains (blue).

For single chains,  $R_g$  decreased with increasing nSCD (**Figure 1a**), which is consistent with many previous studies that examined this behavior in detail<sup>20,21,35,39</sup>. To study the conformations in the dense phase, we performed bulk simulations maintaining a constant external pressure ( $P =$

0 atm), allowing the systems to adopt their preferred concentration. For EKV1 (nSCD = 0), no dense phase formed at the investigated temperature ( $T = 300$  K) because the alternating distribution of positively and negatively charged residues resulted in weak interchain attractions<sup>39</sup>. For all other EKVs, a dense phase formed, where its concentration increased with increasing nSCD (**Figure S2a**). For these cases, as nSCD increased, we observed a corresponding modest rise in  $R_g$  within the dense phase, which amounted to an approximate 11.4% increase compared to the single chain  $R_g$  of EKV1. Conversely, for a single chain, there was a marked decrease in  $R_g$ , up to 40.8%, from EKV1 to EKV15 (**Figure 1a**). Consequently, the disparity in  $R_g$  between the dense phase and single chains widened as nSCD increased. Specifically, this difference expanded from 8.7% to 88.2% relative to the  $R_g$  of single chains, spanning from EKV2 to EKV15. For single chains, the  $R_g$  probability distribution,  $P(R_g)$ , substantially narrowed with increasing nSCD (**Figure 1b**), which reflects the smaller conformational variety of collapsed blocky EKV sequences. In contrast,  $P(R_g)$  was much broader in the dense phase and slightly broadened with increasing nSCD. These trends indicate that the dense phase of EKVs exhibits a markedly greater diversity of conformations, which depend only weakly on the specific protein sequence, as compared to their single-chain state. These results are consistent with prior experiments of Tau proteins, which also found more expanded conformations and enhanced conformational fluctuations for proteins in droplets<sup>31</sup>.

To characterize the shape of the individual chains, we calculated the average relative shape anisotropy ( $\kappa^2$ ) using the three eigenvalues ( $\lambda_i$ ) of the gyration tensor:

$$\kappa^2 = 1 - 3 \frac{\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1}{(\lambda_1 + \lambda_2 + \lambda_3)^2}, \quad (3)$$

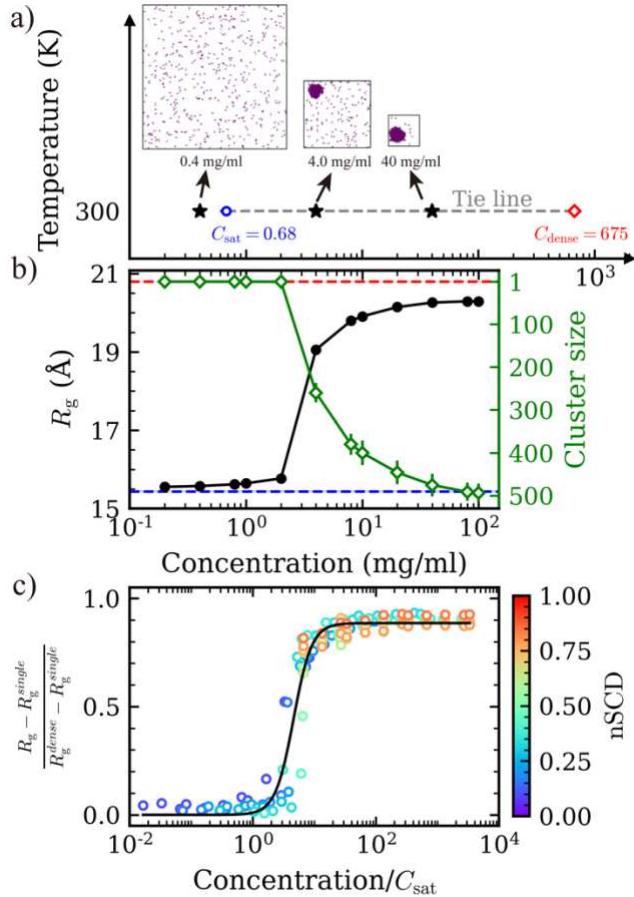
where  $\kappa^2 = 0$  indicates a spherical configuration,  $\kappa^2 \approx 0.39$  an ideal-like chain conformation,

and  $\kappa^2 = 1$  a rod-like structure<sup>40-42</sup>. In the dense phase,  $\kappa^2$  exhibits only small variations with an increase in nSCD (**Figure 1c**), fluctuating slightly above the value expected for ideal-like chains. In contrast, for a single chain,  $\kappa^2$  significantly decreases with increasing nSCD, pointing to a coil-to-globule transition. Like the  $R_g$  distribution, the probability distribution of  $\kappa^2$  for single EKVs narrowed when increasing nSCD (**Figure S3**), which suggests restricted conformational variations for sequences with high charge segregation. Furthermore, we compared these conformations with an ideal chain conformation (**Figure 1d**). Interestingly, for EKVs in the dense phase, the end-to-end distance,  $R_e$ , closely follows the theoretically expected behavior of an ideal chain, *i.e.*,  $R_e = \sqrt{6}R_g$ <sup>43</sup>. However,  $R_e$  of EKVs in single-chain state showed notable deviations from this relationship, particularly pronounced in EKVs with high charge segregation.

These results demonstrate that the conformations observed in the dense phase, as well as those of the sequences with low to moderate degree of charge segregation in the dilute phase, are closely akin to the conformation of an ideal chain. These observed conformational properties can be understood by considering the attractive interactions between monomers. In the dense phase, chains are surrounded by other chains, allowing monomers from a chain to interact with neighboring chains, leading to the observed chain expansion. In contrast, in the dilute phase, monomers can only form intramolecular contacts, thereby leading to collapsed conformation to achieve a state of minimum free energy. Within the dense phase, chain's conformation nearly mirrors the random-walk characteristics of an ideal chain, which maximizes the conformational entropy (and thus minimizes the free energy)<sup>43</sup>. As a result, in the dense phase, conformations of EKVs exhibited a minor sequence-dependent variation. Conversely, the conformation of an isolated chain in a poor solvent is determined by the intrachain interactions that involve an equilibrium between long-range electrostatic repulsion and attraction, which is substantially

modulated by the sequence<sup>20</sup>.

Having established the conformations within the dense and dilute phases, we proceeded to analyze the conformational transitions between these two phases by simulating the EKVs across a series of fixed concentrations, from 0.2 mg/ml to 100 mg/ml. This was achieved by maintaining a constant number of chains ( $N = 500$ ) and modifying the volume of the simulation box (**Figure 2a**). When the concentration exceeds the saturation concentration  $c_{\text{sat}}$ , some chains spontaneously assemble so that the system phase separates into a dilute and a dense phase. To illustrate this transition in more detail, we chose the sequence EKV5 as an example (**Figure 2a,b**), whose  $c_{\text{sat}}$  ( $=0.68$  mg/ml) falls within our investigated concentration range. For concentrations  $c \leq 0.4$  mg/ml  $< c_{\text{sat}}$ , the average cluster size remained at one (see SI for technical details of the cluster analysis), indicating that the chains have do not show signs of condensation yet. Correspondingly, the distributions of  $R_g$  were identical to the ones measured in our single chain simulations (**Figure S4**).



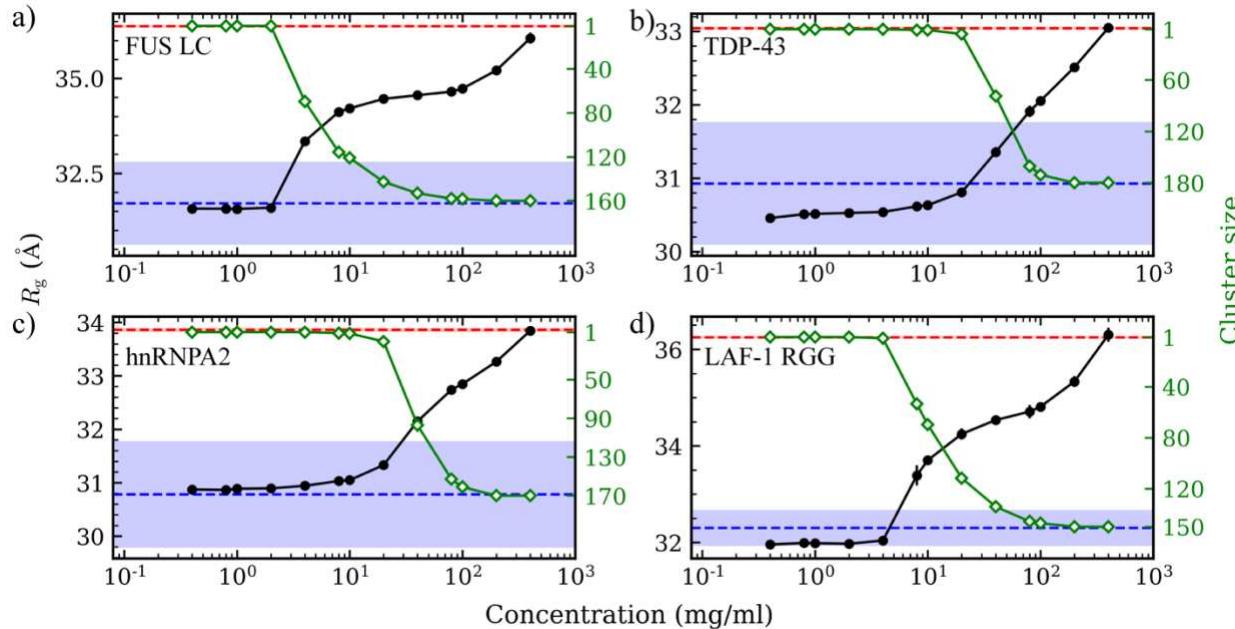
**Figure 2.** Concentration scan for the EKVs. (a) Tie line for EKV5 at  $T = 300\text{ K}$  indicating coexisting dense ( $c_{\text{dense}}$ ) and dilute ( $c_{\text{sat}}$ ) phase concentrations along with snapshots at different concentrations. (b)  $R_g$  (black, left y-axis) and cluster size (green, right y-axis) of EKV5 as functions of concentration. The red and blue dashed horizontal lines represent the  $R_g$  in the bulk dense phase and of a single chain, respectively. (c) Normalized  $R_g$  as a function of normalized concentration for the EKVs. The black solid line is the fitted curve for all simulation data (symbols).

Intriguingly, for  $c_{\text{sat}} < c < 4\text{ mg/ml}$ , the average cluster size remains near one, albeit marginally larger, due to the formation of a small number of clusters ranging in size from 2 to 10 assembled chains (Figure S5a). The local concentration within these clusters is identical to the dense phase concentrations from our bulk simulations (Figure S5b), which suggests that phase separation is indeed occurring above the  $c_{\text{sat}}$ , even if macroscopic phase separation cannot be observed yet due to system size limitations. For concentration around  $c = 4.0\text{ mg/ml}$ , dense phase formation becomes much more evident with the appearance of a droplet (Figure 2a). At these concentrations, the chain  $R_g$ , averaged over the whole system, has a rather wide distribution and

lies between the  $R_g$  in the dense phase and that of a single chain (**Figure 2b**). As the concentration was increased further, the cluster size increased until it saturated at 500, signifying that all chains formed a single condensate. Correspondingly,  $R_g$  continued to increase, approaching the value of the dense bulk phase,  $R_g^{\text{dense}}$ . Note that the small discrepancy between the  $R_g$  measured in our condensate simulations and  $R_g^{\text{dense}}$  likely originates from slightly collapsed chains located at the condensate interface<sup>44</sup>. These findings reveal a gradual increase in average chain size as the system transitions from the dilute phase to the dense phase.

To substantiate the generality of these conclusions, we analyzed all EKVs (**Figures S6, S7**) and normalized the  $R_g$  and concentrations (**Figure 2c**). We normalized the  $R_g$  to span between the value for a single chain,  $R_g^{\text{single}}$ , and the value of a chain in the dense bulk phase,  $R_g^{\text{dense}}$ . We normalized the concentrations by the  $c_{\text{sat}}$  values. We note that the  $c_{\text{sat}}$  values for EKV10 to EKV15 were extrapolated based on the data from sequences with lower nSCD, owing to the scarcity of chains in the dilute phase during slab co-existence simulations (**Figure S2b**). These extrapolated  $c_{\text{sat}}$  values did not affect the overall trend, as all the corresponding  $R_g$  values were similar to  $R_g^{\text{dense}}$ . For EKVs with a low nSCD value (i.e., EKV2 to EKV7),  $R_g$  closely approximated  $R_g^{\text{single}}$  at concentrations below or slightly above the saturation concentration. At higher concentrations, where a distinct condensate formed,  $R_g$  increased until it almost reached  $R_g^{\text{dense}}$ . For the EKVs with higher nSCD values, the smallest concentration that we explored in our simulations is much larger than its saturation concentration, hence we did not observe a  $R_g$  value close to that of a single chain. However, with an increase in concentration, we still noted a rise in  $R_g$  until it plateaued at a value marginally smaller than  $R_g^{\text{dense}}$ . The normalized  $R_g$  values for all EKVs collapse onto a single sigmoidal curve ( $R^2 = 0.95$ ), exhibiting a uniform trend of

increase from the dilute to the dense phase. This convergence demonstrates that the gradual increase in  $R_g$  from the dilute phase to the dense phase is a universal attribute across all EKVs.



**Figure 3.** Concentration scan for the disordered domains of natural proteins: (a) FUS LC, (b) TDP-43, (c) hnRNPA2, and (d) LAF-1 RGG.  $R_g$  (black, left y-axis) and cluster size (green, right y-axis) as functions of concentrations. The red and blue dashed horizontal lines represent the  $R_g$  in the bulk dense phase and of a single chain, respectively.

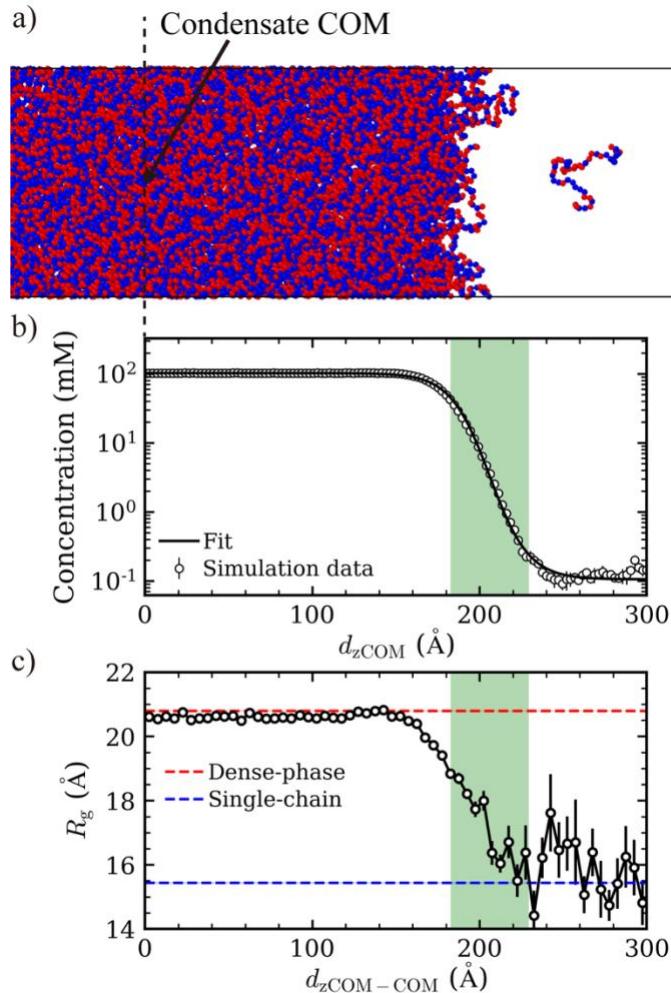
Having established that for EKVs bearing zero net charge with diverse charge patterns,  $R_g$  progressively increases from the dilute phase to the dense phase, we next sought to verify this conclusion for natural IDPs. We selected four proteins, previously demonstrated to undergo LLPS *in vitro*, namely the low-complexity (LC) domain of FUS<sup>45</sup>, the disordered C-terminal domain of TDP-43<sup>46</sup>, the LC domain of hnRNPA2<sup>47</sup> and the N-terminal disordered RGG domain of LAF-1 (LAF-1 RGG)<sup>48</sup>. According to previous research on phase behavior of these IDPs<sup>49</sup>, we conducted simulations at  $T = 300$  K for the first three proteins, and at  $T = 260$  K for LAF-1 RGG. We maintained a constant total monomer count of approximately 25,000 while varying the box volume to explore a range of concentrations from 0.2 mg/ml to 400 mg/ml. For all four sequences, the average  $R_g$  of the entire system initially remained nearly constant when the proteins remained

dispersed in solution, and the average cluster size was close to one. As the dense phase formed (indicated by the growing cluster size shown in **Figure 3**), the average  $R_g$  gradually increased, eventually reaching the  $R_g$  that is characteristic of the bulk dense phase (**Figure 3, Figure S8**). The progression of  $R_g$  with increasing concentration is similar to those observed for the EKVs, with  $R_g$  gradually increasing alongside cluster size as the system evolves from the dilute phase into the dense phase. These findings suggest that during the phase separation process, IDPs undergo a transition from a dilute state to an oligomeric state, and ultimately to a dense state, accompanied by a gradual chain expansion.

Upon condensation, an interfacial region emerges between the dilute and dense phases, which likely plays an important role in the functionality and stability of MLOs<sup>50-52</sup>. Previously, our group elucidated the conformation of homopolymer chains and select IDPs at the interface of condensate droplets, and here we conducted a similar analysis for EKVs<sup>44</sup>. To eliminate the effects of (local) curvature of the condensate interface in a droplet geometry, we performed slab simulations to study conformations at interfaces. Taking EKV5 as an illustrative example (data for the other sequences are presented in the SI), a dense phase was observed, with occasional appearances of several chains in the dilute phase (**Figure 4a**). Interface boundaries were determined by fitting the concentration profile relative to the distance from the condensate center-of-mass to monomer in the  $z$  direction ( $d_{z\text{COM}}$ ) using a hyperbolic tangent function<sup>53</sup>:

$$\begin{aligned} \log_{10}[c(d_{z\text{COM}})] \\ = \frac{1}{2} [\log_{10}(c^{\text{bulk}}) + \log_{10}(c^{\text{dilute}})] \\ - \frac{1}{2} [\log_{10}(c^{\text{bulk}}) - \log_{10}(c^{\text{dilute}})] \tanh\left(\frac{2(d_{z\text{COM}} - d_{z\text{COM}}^{\text{mid}})}{\delta}\right) \end{aligned} \quad (4)$$

where  $c(d_{z\text{COM}})$  is the concentration profile along the  $z$  direction,  $c^{\text{bulk}}$  and  $c^{\text{dilute}}$  are the concentrations of the bulk phase and dilute phase, respectively,  $d_{z\text{COM}}^{\text{mid}}$  is the midpoint of the hyperbolic tangent function and  $\delta$  is the width of the interface. Consequently, the interface boundaries are defined as  $d_{z\text{COM}}^{\text{mid}} \pm \delta/2$ .



**Figure 4.** Analysis of EKV5 conformations at the interface *via* slab simulation. (a) Snapshot depicting one side of the interface of a condensate from the slab simulation. (b) Concentration profiles with respect to the distance from the condensate center-of-mass in the  $z$ -direction,  $d_{z\text{COM}}$ . The solid line is the fitted curve for the simulation data (symbols). (c) Average  $R_g$  with respect to distance from the condensate center-of-mass to chain's center-of-mass in the  $z$  direction,  $d_{z\text{COM}-\text{COM}}$ . The red and blue dashed horizontal lines represent the  $R_g$  in the bulk dense phase and of a single chain, respectively.

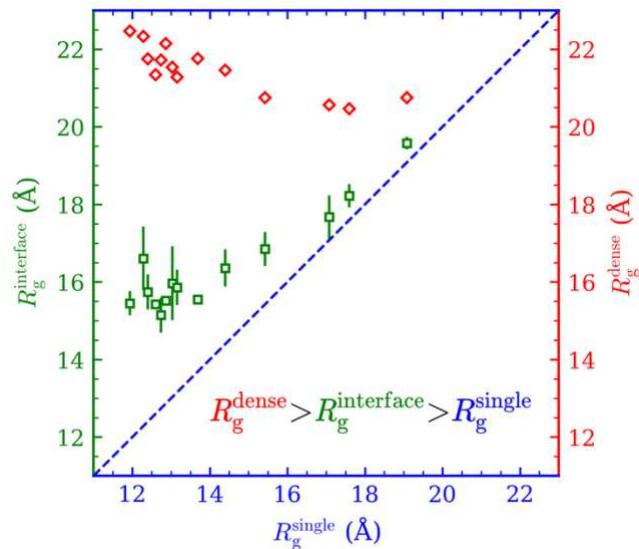
Upon identifying the interfacial region (Figure 4b), we analyzed the local  $R_g$  with respect to the distance along the  $z$ -direction between a chain's center-of-mass and the center-of-mass of the

condensate,  $d_{z\text{COM-COM}}$ . In the dense region of the slab,  $R_g$  overlapped with the value obtained in the bulk dense phase,  $R_g^{\text{dense}}$ . Within the interface region,  $R_g$  decreased and remained smaller than those observed for the chains within the bulk phase. Within the dilute phase,  $R_g$  exhibited pronounced fluctuations around the size of a single chain,  $R_g^{\text{single}}$ . These fluctuations can be attributed to the limited statistical data available due to the low concentration of chains in the coexisting dilute phase. Importantly, all EKVs exhibit similar qualitative behavior, with conformations in the interfacial region being more compact than those in the dense interior and less compact than the dilute phase (**Figures S9, S10**).

In this simulation study, we investigated the conformational properties of proteins during phase separation by simulating model sequences composed of an equal number of oppositely charged monomers arranged in different patterns (referred to as EKVs) and naturally occurring disordered proteins. Despite extensive previous research on phase separation<sup>1,2</sup>, the description of conformational transitions underlying the condensation process remains largely qualitative. Due to the heterogeneous and dynamic nature of protein ensembles, it is difficult to tackle this question experimentally, which makes the simulation approaches used here an attractive avenue for gaining a comprehensive and quantitative understanding of sequence-dependent changes in conformational properties.

We decipher the self-assembly process itself by conducting simulations with increasing protein concentration. As expected, proteins initially stay as individual molecules in the dilute phase at low concentrations but then start to assemble into larger clusters at concentrations only above their saturation concentration. The cluster size grows with increasing protein concentration and eventually a single protein droplet forms with all proteins incorporated in it at very high concentrations. Importantly, ensemble conformations progressively shift from single-chain to bulk

dense-phase and follow a sequence-independent universal behavior as a function of protein concentration normalized by the saturation concentration.



**Figure 5.** Correlation between  $R_g^{\text{single}}$  and  $R_g^{\text{interface}}$  (green squares, left y-axis) or  $R_g^{\text{dense}}$  (red diamonds, right y-axis) for all investigated EKVs.

Even though the EKVs and natural proteins exhibit pronounced sequence-dependent conformational characteristics in the dilute phase (**Figure S11a**)<sup>35</sup>, the conformation within the dense phase demonstrates only a very weak sequence dependence (**Figure 5**). In fact, the relationship between the protein's radius of gyration ( $R_g$ ) and end-to-end distance ( $R_e$ ) within the condensates is consistent with theoretical predictions for an ideal chain ( $R_e = \sqrt{6}R_g$ ) and their intramolecular distance ( $R_{ij}$ ) closely mirrors the scaling expected for an ideal chain (**Figure S11b**). These analyses strongly suggest that the protein conformations within the dense phase are akin to that of an ideal chain, which is characterized by being entropy-driven and not strongly influenced by the protein sequence or dilute-phase conformational properties<sup>43</sup>. We note that our results are consistent with the limited experimental data available in the literature<sup>7,30,3254</sup>.

For conformational characteristics of proteins at the droplet interface, we do observe a

significant dependence on the protein sequence, and hence their dilute-phase properties (**Figure 5**). Importantly, the protein chains at the interface are always more compact than the chains in the dense phase but remain more extended as compared to those in the dilute phase, independent of their sequence.

The results reported here uncover the general conformational characteristics of proteins during the phase separation process and highlight their distinct preferences as they move from the dilute phase to the dense phase including their interfacial behavior. These results should help understand the functional and pathological roles of conformational transitions associated with the formation of biomolecular condensates.

## **Supporting Information**

The details of the model and simulations, method of nSCD calculation, method for cluster analysis, amino acid code for the model and natural proteins, dense phase and saturation concentrations, probability distribution of  $\kappa^2$ , cluster size distributions, effect of concentration scan on probability distribution of  $R_g$ , cluster sizes, and average  $R_g$ , concentration profile and  $R_g$  with respect to the distance from the condensate's center of mass for the EKVs, interresidue distance analysis for the EKVs.

## **Acknowledgments**

This material is based on the research supported by the National Institute of General Medical Science (NIGMS) of the National Institutes of Health under the grant R01GM136917 and the Welch Foundation under the grant A-2113-20220331. The data on TDP-43 disordered domain were generated as part of a project supported by NINDS and NIA grant R01NS116176. A.N. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research

Foundation) through Project 470113688. Y.C.K. is supported by the Office of Naval Research via the U.S. Naval Research base program. We gratefully acknowledge the computational resources provided by the Texas A&M High Performance Research Computing (HPRC) to complete this work.

## References

- (1) Hyman, A. A.; Weber, C. A.; Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annu. Rev. Cell Dev. Biol.* **2014**, *30* (1), 39–58. <https://doi.org/10.1146/annurev-cellbio-100913-013325>.
- (2) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular Condensates: Organizers of Cellular Biochemistry. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (5), 285–298. <https://doi.org/10.1038/nrm.2017.7>.
- (3) Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019**, *176* (3), 419–434. <https://doi.org/10.1016/j.cell.2018.12.035>.
- (4) Dignon, G. L.; Best, R. B.; Mittal, J. Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. *Annu. Rev. Phys. Chem.* **2020**, *71* (1), 53–75. <https://doi.org/10.1146/annurev-physchem-071819-113553>.
- (5) Lyon, A. S.; Peeples, W. B.; Rosen, M. K. A Framework for Understanding the Functions of Biomolecular Condensates across Scales. *Nat. Rev. Mol. Cell Biol.* **2021**, *22* (3), 215–235. <https://doi.org/10.1038/s41580-020-00303-z>.
- (6) Tibble, R. W.; Depaix, A.; Kowalska, J.; Jemielity, J.; Gross, J. D. Biomolecular Condensates Amplify mRNA Decapping by Biasing Enzyme Conformation. *Nat. Chem. Biol.* **2021**, *17* (5), 615–623. <https://doi.org/10.1038/s41589-021-00774-x>.
- (7) Wen, J.; Hong, L.; Krainer, G.; Yao, Q.-Q.; Knowles, T. P. J.; Wu, S.; Perrett, S. Conformational Expansion of Tau in Condensates Promotes Irreversible Aggregation. *J. Am. Chem. Soc.* **2021**, *143* (33), 13056–13064. <https://doi.org/10.1021/jacs.1c03078>.
- (8) Holehouse, A. S.; Kragelund, B. B. The Molecular Basis for Cellular Function of Intrinsically Disordered Protein Regions. *Nat. Rev. Mol. Cell Biol.* **2023**, 1–25. <https://doi.org/10.1038/s41580-023-00673-0>.

(9) Rekhi, S.; Sundaravadivelu Devarajan, D.; Howard, M. P.; Kim, Y. C.; Nikoubashman, A.; Mittal, J. Role of Strong Localized vs Weak Distributed Interactions in Disordered Protein Phase Separation. *J. Phys. Chem. B* **2023**, *127* (17), 3829–3838. <https://doi.org/10.1021/acs.jpcb.3c00830>.

(10) Ashbaugh, H. S.; Hatch, H. W. Natively Unfolded Protein Stability as a Coil-to-Globule Transition in Charge/Hydropathy Space. *J. Am. Chem. Soc.* **2008**, *130* (29), 9536–9542. <https://doi.org/10.1021/ja802124e>.

(11) Vitalis, A.; Pappu, R. V. ABSINTH: A New Continuum Solvation Model for Simulations of Polypeptides in Aqueous Solutions. *J. Comput. Chem.* **2009**, *30* (5), 673–699. <https://doi.org/10.1002/jcc.21005>.

(12) Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. Polymer Scaling Laws of Unfolded and Intrinsically Disordered Proteins Quantified with Single-Molecule Spectroscopy. *Proc. Natl. Acad. Sci.* **2012**, *109* (40), 16155–16160. <https://doi.org/10.1073/pnas.1207719109>.

(13) Miller, C. M.; Kim, Y. C.; Mittal, J. Protein Composition Determines the Effect of Crowding on the Properties of Disordered Proteins. *Biophys. J.* **2016**, *111* (1), 28–37. <https://doi.org/10.1016/j.bpj.2016.05.033>.

(14) Holehouse, A. S.; Pappu, R. V. Collapse Transitions of Proteins and the Interplay Among Backbone, Sidechain, and Solvent Interactions. *Annu. Rev. Biophys.* **2018**, *47* (1), 19–39. <https://doi.org/10.1146/annurev-biophys-070317-032838>.

(15) Baul, U.; Chakraborty, D.; Mugnai, M. L.; Straub, J. E.; Thirumalai, D. Sequence Effects on Size, Shape, and Structural Heterogeneity in Intrinsically Disordered Proteins. *J. Phys. Chem. B* **2019**, *123* (16), 3462–3474. <https://doi.org/10.1021/acs.jpcb.9b02575>.

(16) Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. Charge Interactions Can Dominate the Dimensions of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci.* **2010**, *107* (33), 14609–14614. <https://doi.org/10.1073/pnas.1001743107>.

(17) Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Net Charge per Residue Modulates Conformational Ensembles of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci.* **2010**, *107* (18), 8183–8188. <https://doi.org/10.1073/pnas.0911107107>.

(18) Samanta, H. S.; Chakraborty, D.; Thirumalai, D. Charge Fluctuation Effects on the Shape of Flexible Polyampholytes with Applications to Intrinsically Disordered Proteins. *J. Chem. Phys.* **2018**, *149* (16), 163323. <https://doi.org/10.1063/1.5035428>.

(19) Marsh, J. A.; Forman-Kay, J. D. Sequence Determinants of Compaction in Intrinsically Disordered Proteins. *Biophys. J.* **2010**, *98* (10), 2383–2390. <https://doi.org/10.1016/j.bpj.2010.02.006>.

(20) 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>.

(21) Sawle, L.; Ghosh, K. A Theoretical Method to Compute Sequence Dependent Configurational Properties in Charged Polymers and Proteins. *J. Chem. Phys.* **2015**, *143* (8), 085101. <https://doi.org/10.1063/1.4929391>.

(22) Zheng, W.; Dignon, G.; Brown, M.; Kim, Y. C.; Mittal, J. Hydropathy Patterning Complements Charge Patterning to Describe Conformational Preferences of Disordered Proteins. *J. Phys. Chem. Lett.* **2020**, *11* (9), 3408–3415. <https://doi.org/10.1021/acs.jpclett.0c00288>.

(23) Rumyantsev, A. M.; Johner, A.; de Pablo, J. J. Sequence Blockiness Controls the Structure of Polyampholyte Necklaces. *ACS Macro Lett.* **2021**, *10* (8), 1048–1054. <https://doi.org/10.1021/acsmacrolett.1c00318>.

(24) Kumar Hazra, M.; Levy, Y. Charge Pattern Affects the Structure and Dynamics of Polyampholyte Condensates. *Phys. Chem. Chem. Phys.* **2020**, *22* (34), 19368–19375. <https://doi.org/10.1039/D0CP02764B>.

(25) Tran, H. T.; Mao, A.; Pappu, R. V. Role of Backbone–Solvent Interactions in Determining Conformational Equilibria of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2008**, *130* (23), 7380–7392. <https://doi.org/10.1021/ja710446s>.

(26) Kang, H.; Pincus, P. A.; Hyeon, C.; Thirumalai, D. Effects of Macromolecular Crowding on the Collapse of Biopolymers. *Phys. Rev. Lett.* **2015**, *114* (6), 068303. <https://doi.org/10.1103/PhysRevLett.114.068303>.

(27) Huihui, J.; Firman, T.; Ghosh, K. Modulating Charge Patterning and Ionic Strength as a Strategy to Induce Conformational Changes in Intrinsically Disordered Proteins. *J. Chem. Phys.* **2018**, *149* (8), 085101. <https://doi.org/10.1063/1.5037727>.

(28) Wohl, S.; Jakubowski, M.; Zheng, W. Salt-Dependent Conformational Changes of Intrinsically Disordered Proteins. *J. Phys. Chem. Lett.* **2021**, *12* (28), 6684–6691. <https://doi.org/10.1021/acs.jpclett.1c01607>.

(29) Baidya, L.; Reddy, G. pH Induced Switch in the Conformational Ensemble of Intrinsically Disordered Protein Prothymosin- $\alpha$  and Its Implications for Amyloid Fibril Formation. *J. Phys. Chem. Lett.* **2022**, *13* (41), 9589–9598. <https://doi.org/10.1021/acs.jpclett.2c01972>.

(30) Ubbiali, D.; Fratini, M.; Piersimoni, L.; Ihling, C. H.; Kipping, M.; Heilmann, I.; Iacobucci, C.; Sinz, A. Direct Observation of “Elongated” Conformational States in  $\alpha$ -Synuclein upon Liquid-Liquid Phase Separation. *Angew. Chem.* **2022**, *134* (46), e202205726. <https://doi.org/10.1002/ange.202205726>.

(31) Majumdar, A.; Dogra, P.; Maity, S.; Mukhopadhyay, S. Liquid–Liquid Phase Separation Is Driven by Large-Scale Conformational Unwinding and Fluctuations of Intrinsically Disordered Protein Molecules. *J. Phys. Chem. Lett.* **2019**. <https://doi.org/10.1021/acs.jpclett.9b01731>.

(32) Bauer, D. J.; Stelzl, L. S.; Nikoubashman, A. Single-Chain and Condensed-State Behavior of hnRNPA1 from Molecular Simulations. *J. Chem. Phys.* **2022**, *157* (15), 154903. <https://doi.org/10.1063/5.0105540>.

(33) Madinya, J. J.; Chang, L.-W.; Perry, S. L.; Sing, C. E. Sequence-Dependent Self-Coacervation in High Charge-Density Polyampholytes. *Mol. Syst. Des. Eng.* **2020**, *5* (3), 632–644. <https://doi.org/10.1039/C9ME00074G>.

(34) Lyons, H.; Veettil, R. T.; Pradhan, P.; Fornero, C.; De La Cruz, N.; Ito, K.; Eppert, M.; Roeder, R. G.; Sabari, B. R. Functional Partitioning of Transcriptional Regulators by Patterned Charge Blocks. *Cell* **2023**, *186* (2), 327–345.e28. <https://doi.org/10.1016/j.cell.2022.12.013>.

(35) Sundaravadivelu Devarajan, D.; Rekhi, S.; Nikoubashman, A.; Kim, Y. C.; Howard, M. P.; Mittal, J. Effect of Charge Distribution on the Dynamics of Polyampholytic Disordered Proteins. *Macromolecules* **2022**, *55* (20), 8987–8997. <https://doi.org/10.1021/acs.macromol.2c01390>.

(36) Sundaravadivelu Devarajan, D.; Wang, J.; Nikoubashman, A.; Kim, Y. C.; Mittal, J. Sequence-Dependent Material Properties of Biomolecular Condensates and Their Relation to Dilute Phase Conformations. bioRxiv 2023, DOI: 10.1101/2023.05.09.540038.

(37) Firman, T.; Ghosh, K. Sequence Charge Decoration Dictates Coil-Globule Transition in Intrinsically Disordered Proteins. *J. Chem. Phys.* **2017**, *148* (12), 123305. <https://doi.org/10.1063/1.5005821>.

(38) Rana, U.; Brangwynne, C. P.; Panagiotopoulos, A. Z. Phase Separation vs Aggregation Behavior for Model Disordered Proteins. *J. Chem. Phys.* **2021**, *155* (12), 125101. <https://doi.org/10.1063/5.0060046>.

(39) Lin, Y.-H.; Chan, H. S. Phase Separation and Single-Chain Compactness of Charged Disordered Proteins Are Strongly Correlated. *Biophys. J.* **2017**, *112* (10), 2043–2046. <https://doi.org/10.1016/j.bpj.2017.04.021>.

(40) Rudnick, J.; Gaspari, G. The Shapes of Random Walks. *Science* **1987**, *237* (4813), 384–389. <https://doi.org/10.1126/science.237.4813.384>.

(41) Jagodzinski, O.; Eisenriegler, E.; Kremer, K. Universal Shape Properties of Open and Closed Polymer Chains: Renormalization Group Analysis and Monte Carlo Experiments. *J. Phys. I* **1992**, *2* (12), 2243–2279. <https://doi.org/10.1051/jp1:1992279>.

(42) Dima, R. I.; Thirumalai, D. Asymmetry in the Shapes of Folded and Denatured States of Proteins. *J. Phys. Chem. B* **2004**, *108* (21), 6564–6570. <https://doi.org/10.1021/jp037128y>.

(43) Rubinstein, M; Colby, R. H. *Polymer Physics*; Oxford University Press: New York, 2003.

(44) Wang, J.; Sundaravadivelu Devarajan, D.; Nikoubashman, A.; Mittal, J. Conformational Properties of Polymers at Droplet Interfaces as Model Systems for Disordered Proteins. *ACS Macro Lett.* **2023**, *12* (11), 1472–1478. <https://doi.org/10.1021/acsmacrolett.3c00456>.

(45) Patel, A.; Lee, H. O.; Jawerth, L.; Maharana, S.; Jahnel, M.; Hein, M. Y.; Stoynov, S.; Mahamid, J.; Saha, S.; Franzmann, T. M.; Pozniakovski, A.; Poser, I.; Maghelli, N.; Royer, L. A.; Weigert, M.; Myers, E. W.; Grill, S.; Drechsel, D.; Hyman, A. A.; Alberti, S. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **2015**, *162* (5), 1066–1077. <https://doi.org/10.1016/j.cell.2015.07.047>.

(46) Conicella, A. E.; Zerze, G. H.; Mittal, J.; Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by  $\alpha$ -Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **2016**, *24* (9), 1537–1549. <https://doi.org/10.1016/j.str.2016.07.007>.

(47) 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>.

(48) Elbaum-Garfinkle, S.; Kim, Y.; Szczepaniak, K.; Chen, C. C.-H.; Eckmann, C. R.; Myong, S.; Brangwynne, C. P. The Disordered P Granule Protein LAF-1 Drives Phase Separation into Droplets with Tunable Viscosity and Dynamics. *Proc. Natl. Acad. Sci.* **2015**, *112* (23), 7189–7194. <https://doi.org/10.1073/pnas.1504822112>.

(49) 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.* **2018**, *115* (40), 9929–9934. <https://doi.org/10.1073/pnas.1804177115>.

(50) Gouveia, B.; Kim, Y.; Shaevitz, J. W.; Petry, S.; Stone, H. A.; Brangwynne, C. P. Capillary Forces Generated by Biomolecular Condensates. *Nature* **2022**, *609* (7926), 255–264. <https://doi.org/10.1038/s41586-022-05138-6>.

(51) Böddeker, T. J.; Rosowski, K. A.; Berchtold, D.; Emmanouilidis, L.; Han, Y.; Allain, F. H. T.; Style, R. W.; Pelkmans, L.; Dufresne, E. R. Non-Specific Adhesive Forces between Filaments and Membraneless Organelles. *Nat. Phys.* **2022**, *18* (5), 571–578. <https://doi.org/10.1038/s41567-022-01537-8>.

(52) Lipiński, W. P.; Visser, B. S.; Robu, I.; Fakhree, M. A. A.; Lindhoud, S.; Claessens, M. M. A. E.; Spruijt, E. Biomolecular Condensates Can Both Accelerate and Suppress Aggregation of  $\alpha$ -Synuclein. *Sci. Adv.* **2022**, *8* (48), eabq6495. <https://doi.org/10.1126/sciadv.abq6495>.

(53) Kuo, I.-F. W.; Mundy, C. J.; Eggimann, B. L.; McGrath, M. J.; Siepmann, J. I.; Chen, B.; Vieceli, J.; Tobias, D. J. Structure and Dynamics of the Aqueous Liquid–Vapor Interface: A Comprehensive Particle-Based Simulation Study. *J. Phys. Chem. B* **2006**, *110* (8), 3738–3746. <https://doi.org/10.1021/jp056330t>.

(54) Joshi, A.; Walimbe, A.; Avni, A.; Rai, S. K.; Arora, L.; Sarkar, S.; Mukhopadhyay, S. Single-Molecule FRET Unmasks Structural Subpopulations and Crucial Molecular Events during FUS Low-Complexity Domain Phase Separation. *Nat. Commun.* **2023**, *14* (1), 7331. <https://doi.org/10.1038/s41467-023-43225-y>.