Sequence-Dependent Material Properties of Biomolecular Condensates and their Relation to Dilute Phase Conformations

Dinesh Sundaravadivelu Devarajan,[1,*] Jiahui Wang,[1] Arash Nikoubashman,[2,3,4]
Young C. Kim,[5] and Jeetain Mittal[1,6,7,*]

[1]Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, United States

[2]Institute of Physics, Johannes Gutenberg University Mainz, Staudingerweg 7, 55128 Mainz, Germany

[3]Leibniz-Institut für Polymerforschung Dresden e.V., Hohe Straße 6, Dresden 01069, Germany

[4]Institut für Theoretische Physik, Technische Universität Dresden, Dresden 01069, Germany

[5]Center for Materials Physics and Technology, Naval Research Laboratory, Washington, DC 20375, United States

[6]Department of Chemistry, Texas A&M University, College Station, TX 77843, United States

[7]Interdisciplinary Graduate Program in Genetics and Genomics, Texas A&M University, College Station, TX 77843, United States

[*]Corresponding authors email: dineshsd@tamu.edu, jeetain@tamu.edu

Abstract

Material properties of phase-separated biomolecular assemblies, enriched with disordered proteins, dictate their ability to participate in many cellular functions. Despite the significant effort dedicated to understanding how the sequence of the disordered protein drives its phase separation to form condensates, little is known about the sequence determinants of condensate material properties. Here, we computationally decipher these relationships for charged disordered proteins using model sequences comprised of glutamic acid and lysine residues as well as naturally occurring sequences of LAF1's RGG domain and DDX4's N-terminal domain. We do so by delineating how the arrangement of oppositely charged residues within these sequences influences the dynamical, rheological, and interfacial properties of the condensed phase through equilibrium and non-equilibrium molecular simulations. Our computations yield material properties that are quantitatively comparable with experimentally characterized condensate systems. Interestingly, we find that the material properties of both the model and natural proteins respond similarly to the segregation of charges, despite their very different sequence compositions. Condensates of the highly charge-segregated sequences exhibit slower dynamics than the uniformly charge-patterned sequences, because of their comparatively long-lived molecular contacts between oppositely charged residues. Surprisingly, the molecular interactions within the condensate are highly similar to those within a single-chain for all sequences. Consequently, the condensate material properties of charged disordered proteins are strongly correlated with their dense phase contact dynamics and their single-chain structural properties. Our findings demonstrate the potential to harness the sequence characteristics of disordered proteins for predicting and engineering the material properties of functional condensates, with insights from the dilute phase properties.

Introduction

Material properties of biomolecular condensates play a key role in the proper execution of different biological functions, *e.g.*, cell division,[1] selective autophagy,[2, 3] gene regulation,[4, 5] and nuclear-cytoplasmic shuttling.[6] In general, the material state of condensates ranges from reversible liquid-like assemblies with functional relevance to irreversible solid-like assemblies with pathological consequences.[7-10] The viscous nature (liquidity) of the condensates formed *via* phase separation enables them to act as dynamic assemblies, which exchange molecules with the surrounding environment, and dissolve as required, in response to physiological cues.[11, 12] To better understand the spatiotemporal evolution of condensates, researchers have recently started characterizing their mesoscopic material properties such as diffusion coefficient, viscosity, viscoelasticity, and surface tension through experiments.[13-17] However, much remains unknown, especially how the protein sequence dictates these different material properties.

Intrinsically disordered proteins (IDPs) or regions (IDRs) are deemed essential for condensate formation,[18-21] which across proteomes, typically consist of a high fraction of charged residues.[22, 23] Alterations within an IDP sequence, *e.g.*, due to mutation or post-translational modification, can lead to changes in the intra- and inter-molecular interactions, which can impact both the phase behavior and material properties of the condensates formed or can provide the ability to modulate them independently.[24] For example, changes in the electrostatic interactions of the highly disordered full-length FUS protein, either by a specific point mutation to glutamic acid[25] or by phosphorylating serine residues[26] within the prion-like domain, did not significantly influence its phase separation behavior. However, such alterations had a contrasting effect on the material state of its condensates: the mutated variant led to a dynamically arrested condensate,[25] while the phosphorylated variant prevented condensate aggregation.[26] In addition, poly-arginine condensates exhibited significantly reduced dynamics (100-fold higher viscosity) than poly-lysine condensates[27], but replacing arginine with lysine residues in artificial IDPs resulted only in a two-fold decrease in viscosity.[24] Despite the significant progress made regarding the sequence determinants of phase separation,[24, 28-34] identifying how the IDP sequence modulates its condensate material properties remains an important open question.[35] Further, establishing the molecular basis of sequence-encoded material properties will provide insights for condensate biology and for designing synthetic condensates with tunable biophysical characteristics.

Deciphering the relationship between sequence features and material properties of charged IDPs is the major goal of this study. To achieve this, we make use of the patterning of oppositely charged residues within polyampholytic sequences, which has been shown to dictate the conformations and phase behavior of model IDPs and naturally occurring IDRs.[22, 32, 36, 37] Recently, alternating charge blockiness within certain IDRs was found to be crucial for their selective partitioning into condensates for transcriptional regulation,[38] suggesting a role for the condensate interfaces.[34, 39] We perform coarse-grained molecular dynamics (MD) simulations to investigate how changes to the charge patterning within IDP sequences influence their condensate material properties such as diffusion coefficient, viscosity, and surface tension. We use model proteins consisting of negatively charged glutamic acid (E) and positively charged lysine (K) residues as well as charge-rich naturally occurring proteins such as the LAF1's RGG domain (hereafter referred to as LAF1) and the DDX4's N-terminal domain (hereafter referred to as DDX4). Our results reveal that the effect of charge segregation on different material properties of the condensates is highly similar between the model proteins and natural proteins. In general, we find that charge blockiness leads to a slowdown in the condensate dynamics, which originates from pronounced electrostatic attraction between the oppositely charged residues in the condensed phase. Further, we show that the molecular interactions within the condensates closely resemble their single-chain interactions. Consequently, we find the material properties of IDP condensates to be strongly correlated with their dense phase dynamics of contact formation and breakage between the oppositely charged residues and with their single-chain structural properties. Our findings demonstrate a way to utilize the sequence-level features of charge-rich IDPs for modulating their condensate material properties and to infer such properties based on the single-chain conformations, which can be characterized *via* fully atomistic simulations or experiments.

Results

The effect of charge patterning on the thermodynamic phase behavior of disordered proteins with varying sequence characteristics such as the chain length (N), the fraction of charged residues (FCR), and the net charge per residue (NCPR) is a well-studied problem.[32, 40-42] For example, Lin and Chan[40] demonstrated that charge patterning in polyampholytic disordered proteins with fixed N altered their critical temperature T_c for phase separation such that its changes were synonymous with that of their single-chain size. Later, Dignon *et al.*[41] showed that T_c and the single-chain coil-to-globule transition temperature T_θ are strongly correlated for a diverse set of natural disordered proteins, which included sequences with different charge patterning. However, the effect of charge patterning on the condensate material properties (*e.g.*, viscosity and surface tension) of proteins with different sequence lengths and charge content is not well understood. To investigate this aspect, we employed the model E−K variants[37] with $N = 50$ residues, and the naturally occurring LAF1 and DDX4 sequence variants with $N = 168$ residues and $N = 236$ residues, respectively (see **Table S1** for the amino acid code of sequences, and Materials and Methods for model details). All model sequences had FCR $= 1$ and NCPR $= 0$ as they consisted of an equal fraction of oppositely charged E and K residues. In the case of the investigated natural protein variants, all LAF1 sequences had $\mathrm{FCR} = 0.262$ and $\mathrm{NCPR} = 0.024$, whereas all DDX4 sequences had $\mathrm{FCR} = 0.288$ and $\mathrm{NCPR} = -0.017$.

The degree of charge segregation in the sequences was quantified using the sequence charge decoration (SCD) parameter.[43, 44] Smaller values of SCD correspond to a more charge-segregated sequence. To compare different polyampholyte compositions and lengths, we defined a normalized SCD (nSCD) parameter that is scaled by the maximum and minimum SCD values achievable for each of the E−K, LAF1, and DDX4 sequence compositions. The values of this parameter lie in the interval $0 \leq \mathrm{nSCD} \leq 1$, with the lower and upper bounds corresponding to a perfectly alternating sequence and diblock sequence, respectively.[37, 45] The sequences and their nSCD value are shown in **Fig. 1a**. Note that the LAF1 sequence with $\mathrm{nSCD} = 0.010$ and the DDX4 sequence with $\mathrm{nSCD} = 0.021$ are the wild-type (WT) sequences. We characterized the diffusion coefficient, viscosity, and surface tension of the condensates by simulating the protein chains in bulk systems and phase coexistence systems (**Fig. 1b**), respectively (see Materials and Methods for further simulation details).

To investigate the ability of our sequence variants to phase separate, we simulated them in a cubic simulation box at a constant pressure of $P = 0$ atm, which allowed the sequences to reach their preferred dense phase concentration ρ. We found that the E−K variant with $\mathrm{nSCD} = 0$ did not form a stable dense phase, indicating that the intermolecular attraction between oppositely charged residues was too weak due to their perfectly alternating arrangement of charges in the sequence.[40] In addition, the most well-mixed LAF1 ($\mathrm{nSCD} = 0.003$) and DDX4 ($\mathrm{nSCD} = 0.002$) variants also did not for a stable dense phase. However, other E−K variants, and all the LAF1 and DDX4 sequences with $\mathrm{nSCD} \gtrsim 0.01$ formed a dense phase. For these sequences, we found that ρ monotonically increased with increasing nSCD (*i.e.*, increasing charge segregation) in a

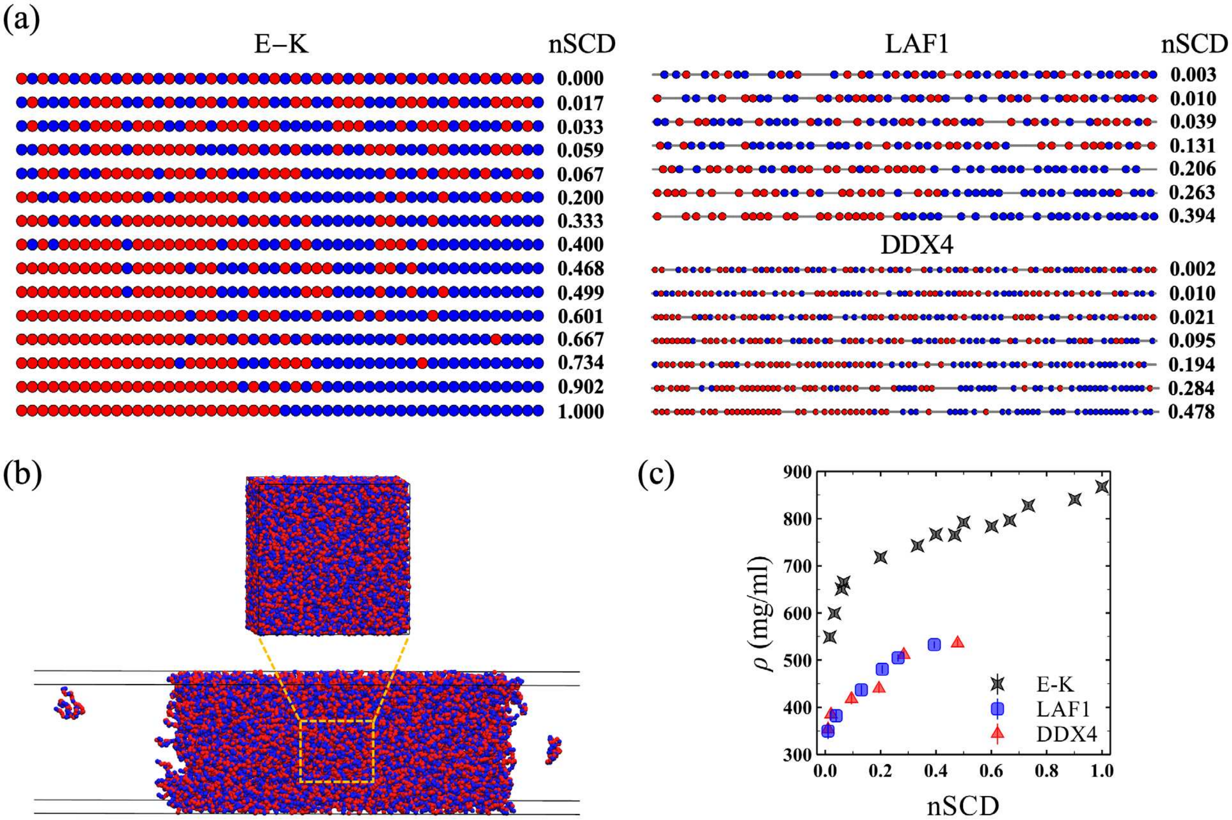

Fig. 1. (a) E−K (chain length, $N = 50$), LAF1 ($N = 168$), and DDX4 ($N = 236$) sequences with their normalized SCD (nSCD) parameter. The LAF1 sequence with nSCD $= 0.010$ and the DDX4 sequence with nSCD $= 0.021$ are the wild-type sequences. The negatively and positively charged residues are shown as red and blue beads, respectively. The uncharged residue positions in LAF1 and DDX4 are shown as gray line segments connecting the charged residues. (b) Simulation snapshots of one of the E−K sequences in a cubic simulation box for characterizing its dense phase dynamical and rheological properties, and in a slab geometry for characterizing its interfacial property between the dense and dilute phases. (c) Dense phase concentration ρ as a function of nSCD for all sequences.

nonlinear fashion: the systems became significantly denser for a small increase in nSCD below 0.2 followed by a slower increase in ρ for larger nSCD (**Fig. 1c**). Further, we found the concentrations of LAF1 and DDX4 variants to be significantly lower compared to the E−K variants. This meant that our sequences facilitated the examination of charge patterning effects over a wide concentration range, which also covers the concentrations of well-studied disordered protein condensates, *e.g.*, low-complexity domains of FUS and hnRNPA2.[20, 21] Note that the sequence variants with nSCD < 0.01 did not form a stable dense phase in the phase coexistence simulations either, while the dense phase concentrations of the other sequences obtained from bulk and phase coexistence simulations were nearly identical (**Fig. S1**). These observations are consistent with the expectation that increasing charge blockiness within the sequence should lead to pronounced electrostatic attraction between the oppositely charged residues, resulting in a denser protein phase.

Charge segregation in IDP sequences leads to monotonic changes in their dense phase material properties. Since our model and natural protein sequences have different sequence compositions, we asked whether the segregation of charges would have a similar effect on the dynamical, rheological, and interfacial properties of their phase-separated condensates. We investigated this aspect by first characterizing the translational motion of the protein chains in the dense phase. In experiments, the chain motion is often estimated from fluorescence recovery after photobleaching (FRAP).[46] In simulations, we can directly measure the chain motion by monitoring the mean square displacement of their residues as a function of time

$$\mathrm{MSD}(t) = \langle [\boldsymbol{r}_i(t) - \boldsymbol{r}_i(0)]^2 \rangle,$$

where $\boldsymbol{r}_i(t)$ is the position of residue i at time t. To avoid chain end effects, we excluded 20 residues on either end of a chain for computing the MSD. We found that the residues of all sequences exhibited the same ballistic motion ($\mathrm{MSD} \propto t^2$) at short times (**Fig. S2**), which was followed by a sub-diffusive motion[47] ($\mathrm{MSD} \propto t^{1/2}$) of residues at intermediate times, before they eventually showed normal diffusion ($\mathrm{MSD} \propto t$) at sufficiently long times. The proportional scaling $\mathrm{MSD} \propto t$ indicated that the condensates of all sequences exhibited largely viscous behavior under the conditions of this study. Most importantly, the MSD showed a strong variation with charge segregation for all sequences in the diffusive regime. To quantify this trend, we computed the diffusion coefficient D by applying the relationship $\mathrm{MSD} = 6Dt$ at long times for all sequences. This way of computing D from the MSD of the residues in a chain accurately reflects its translational motion in the dense interior of the condensates (see Materials and Methods).[48]

To account for the chain length dependence of the diffusion coefficient, we report D^* for the E−K variants obtained after normalizing D by that computed for the reference E−K sequence with $\mathrm{nSCD} = 0.017$, which was the lowest nSCD sequence that phase-separated under the conditions of this study (**Fig. 2a**). Similarly, the D^* values for the LAF1 and DDX4 variants were attained after normalizing D by those computed for their respective WT sequences, which exhibited a well-mixed charge distribution with a similar nSCD value as the reference E−K sequence. We found that D^* monotonically decreased with increasing nSCD for both the model and natural proteins (**Fig. 2a**). Interestingly, the rate of decrease in D^* was also similar, despite their very different sequence compositions. The reduced diffusion coefficient of charge-segregated sequences is consistent with a previous simulation study of similar polyampholyte sequences,[49] and has also been validated experimentally for the LAF1 WT sequence and its charge-shuffled variant.[32] From these observations, we concluded that the segregation of charges leads to a slowdown in the dynamics of condensates formed by charge-rich IDPs.

The material state of the condensates can also be characterized through rheological properties such as the shear viscosity. For this purpose, we used the non-equilibrium MD (NEMD) simulation technique[50-52] (see Materials and Methods for simulation details) for subjecting the dense phase of the protein chains to steady shear at different shear rates $\dot{\gamma}$, and measured the resulting shear stress component τ_{xy}. We then determined their shear viscosity

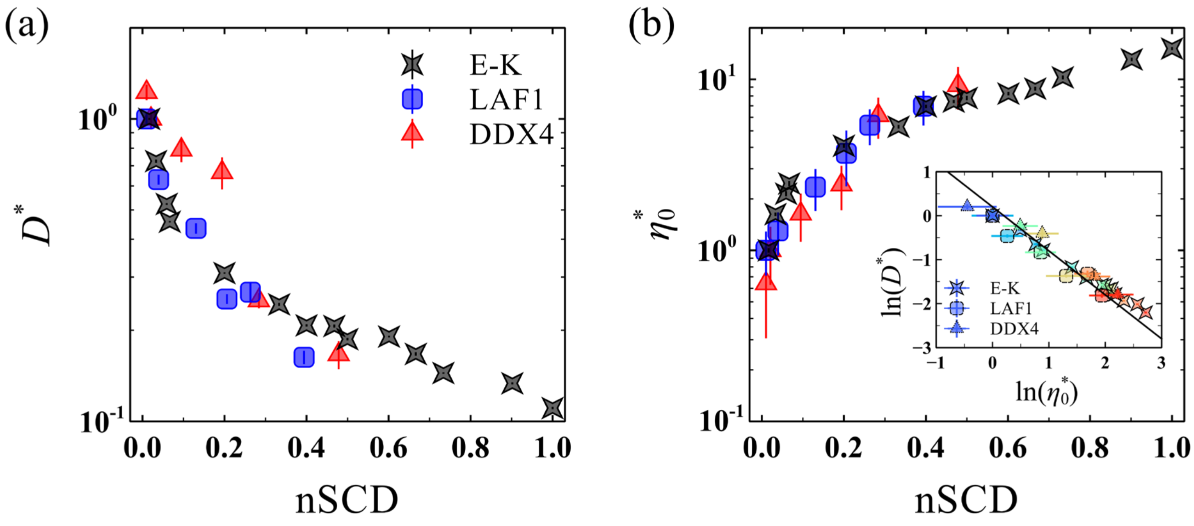

Fig. 2. Normalized material properties of the dense phase as a function of nSCD for the E−K, LAF1, and DDX4 sequences: (a) diffusion coefficient D^* and (b) zero-shear viscosity η_0^*. For the E−K sequences, D^* and η_0^* are obtained after normalizing D and η_0 by those computed for the reference E−K sequence with nSCD = 0.017. For the LAF1 and DDX4 sequences, D^* and η_0^* are obtained after normalizing D and η_0 by those computed for their respective wild-type sequences. The inset shows the inversely proportional correlation between $D^*$and η_0^* for all sequences, with the solid line corresponding to the Stokes-Einstein type relation $D^* = 1/\eta_0^*$. The symbol color, ranging from blue to red, indicates increasing nSCD.

$$\eta = -\frac{\langle \tau_{xy} \rangle}{\dot{\gamma}},$$

and plotted it as a function of $\dot{\gamma}$ (**Fig. S3**). We found near-constant η values at sufficiently low shear rates for all sequences, indicative of their Newtonian liquid-like behavior, similar to that seen for experimentally characterized disordered protein condensates.[18, 24, 32] Similar to the MSD, η depended on the degree of charge segregation within the sequence. To illustrate this dependence, we considered the zero-shear viscosity η_0 of all sequences, which is obtained as an average of the η values from the Newtonian plateau observed at low shear rates (**Fig. S3**). Again, we normalized η_0 by that obtained for the reference sequences (*i.e.*, E−K sequence with nSCD = 0.017 for the E−K variants, and respective WT sequences for the LAF1 and DDX4 variants), which we denote as η_0^* (**Fig. 2b**). We found that η_0^* monotonically increased with increasing nSCD and similar to D^*, η_0^* of all proteins with different sequence characteristics collapsed onto a master curve for the entire range of nSCD. This observation indicated that the interchain electrostatic interactions get stronger as a result of charge blockiness within the sequence of charge-rich IDPs.

The fact that D^* and η_0^* depended on nSCD and that the changes in these quantities was highly similar between the E−K, LAF1, and DDX4 sequence variants, prompts the question: can a simple Stokes-Einstein type relation (*i.e.*, $D^* = 1/\eta_0^*$) capture the changes in these quantities of the charge-patterned sequences? Recent simulation studies have shown that the Stokes-Einstein type relation holds for model associative proteins with varying interaction strength.[53, 54] Indeed, we found that the changes in D^* and η_0^* in response to charge patterning closely followed the relation $D^* = 1/\eta_0^*$ (inset of **Fig. 2b**), indicating the feasibility of predicting these quantities from one another as the charge pattern within the IDPs gets altered. Such a correlation between D^* and η_0^* also highlights that the interchain electrostatic interactions simultaneously dictate both the chain dynamics and rheology of the dense phase of charge-rich IDPs.

Interfacial and dense phase properties of charge-rich IDP condensates obtained from simulations and experiments are quantitatively comparable. Our analysis of the material properties in the dense phase revealed that the IDPs exhibited liquid-like characteristics, though these properties were found to change in a sequence-dependent manner due to changes in the intermolecular interactions within the condensate. Another determinant of condensate liquidity is the surface tension γ at the interface between the dense and dilute phases, which together with the zero-shear viscosity η_0 of the dense phase can dictate the speed of droplet fusion.[15, 16, 55] We determined γ using the Kirkwood-Buff relation[56]

$$\gamma = \frac{L_z}{2}\left\langle P_{zz} - \frac{P_{xx} + P_{yy}}{2}\right\rangle,$$

where L_z is the edge length of the simulation box in the z-direction, P_{ii} is the pressure tensor component in a given direction, and the factor of $1/2$ accounts for the two interfaces present in the slab geometry of the phase coexistence simulations (see **Fig. 1b**). We use a normalized quantity γ^* (**Fig. 3a**), obtained after normalizing γ by the corresponding value of the same reference sequences used in getting D^* and η_0^* of the E−K, LAF1, and DDX4 sequences. We found that the segregation of charges increased γ^*, much like its effect on η_0^*. The values of γ^* and η_0^* for all IDPs were positively correlated (inset of **Fig. 3a**), highlighting that both these quantities can be simultaneously modulated by altering charge patterning of charge-rich IDPs. Taken together, our results revealed that large changes in the material properties of charge-rich IDP condensates can already occur solely through the sequence charge patterning, at constant external conditions such as temperature and salt concentration.

Having assessed the interfacial and dense phase properties of charge-rich IDP condensates using our physics-based coarse-grained (CG) model, we next investigated if they lie in the expected range of experimentally characterized material properties of other naturally occurring protein condensates (**Fig. 3b**). It is known that CG models can significantly underpredict viscosity as compared to experimental measurements,[57] which is primarily due to the smoothening of the free energy landscape[58] and inaccurate solvent frictional effects. Thus, to account for these differences and enable a meaningful comparison between simulations and experiments, it is crucial to establish the connection between their relevant timescales. For this purpose, we derived timescales based on a mesoscopic quantity, specifically the translation diffusive motion of a single protein chain in dilute solution,[59-61] for interpreting the dynamical and rheological properties of

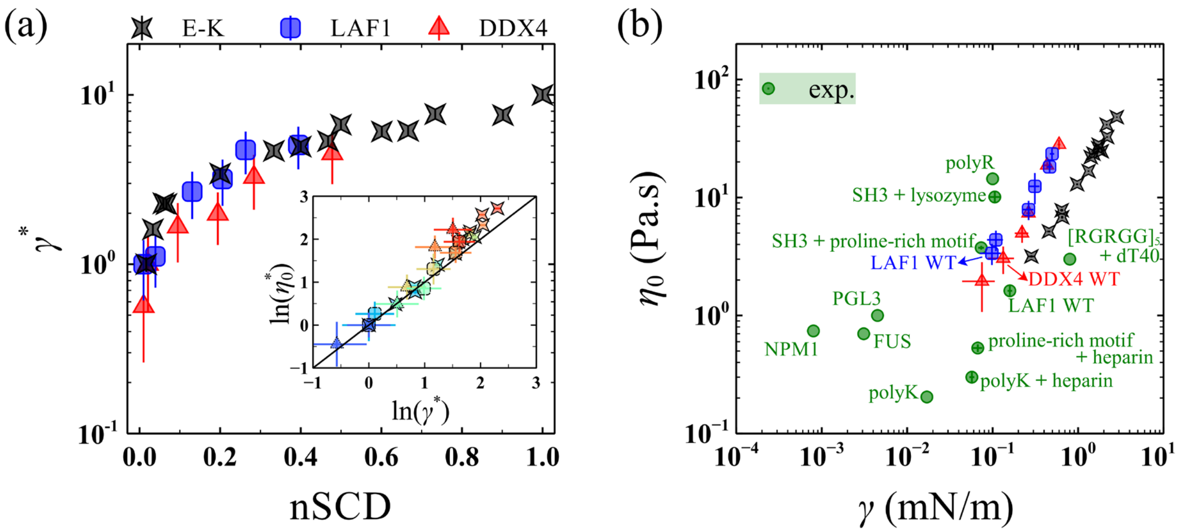

Fig. 3. (a) Normalized surface tension γ^* as a function of nSCD for the E−K, LAF1, and DDX4 sequences. The value of γ^* is obtained after normalizing γ by those computed for the same reference sequences used in the normalization of D and η_0 in **Fig. 2**. The inset shows the correlation between η_0^* and γ^* for all sequences, with the solid lines corresponding to $\eta_0^* = \gamma^*$. The symbol color, ranging from blue to red, indicates increasing nSCD. (b) Comparison of zero-shear viscosity η_0 and surface tension γ obtained for the simulated E−K, LAF1, and DDX4 dense phases with those measured for a wide range of protein condensates through experiments.

the condensed phase obtained from our implicit-solvent CG simulations (see Materials and Methods for details on our timescale mapping strategy). We found that the diffusion coefficients D within LAF1 and DDX4 WT dense phases, based on the timescales mapped at the single-chain level in solution, were in the comparable range to those obtained from FRAP experiments on the corresponding *in vitro* droplets[32, 36] (**Table S2**). Remarkably, the values of zero-shear viscosity η_0 and surface tension γ of the LAF1 WT sequence obtained from simulations and experiments were quantitatively comparable[15] (see Materials and Methods for discussion on the experimental methods used to characterize η_0 and γ). In fact, the range of η_0 and γ for the simulated charge-rich IDP condensates were in close proximity to the η_0 and γ values of other protein condensates that are primarily driven by electrostatic interactions,[16, 27, 62, 63] namely PGL3, polyR, polyK, polyK+heparin, SH3-targeting proline-rich motif+heparin, and $[\mathrm{RGRGG}]_5 + \mathrm{dT40}$ (**Fig. 3b**).

These results demonstrate the ability of our CG model to provide an accurate description of the condensate material properties for a wide range of proteins.

Material properties of IDP condensates have a strong connection to the microscopic dynamics of contact formation and breakage. Our analysis of the material properties of IDP condensates indicated their dependence on the changing nature of molecular interactions as a result of sequence charge patterning. To characterize these changes, we first computed the radial distribution function (RDF) between the oppositely charged residues of each sequence (**Fig. S4**), which we then converted to the potential of mean force (PMF) required for separating them within the dense phase for a given sequence as $\mathrm{PMF} = -k_{\mathrm{B}}T \ln(\mathrm{RDF})$, where k_{B} is the Boltzmann constant (**Fig. S5**). We defined the difference in PMF value at the first minimum (corresponding to the bound state) and $\mathrm{PMF} = 0$ (corresponding to the unbound state) as the free energy change ΔF (**Fig. 4a**). We found $\Delta F > 0$ for the low-nSCD sequences (*i.e.*, $\mathrm{nSCD} \lesssim 0.20$), indicating the absence of an attractive well for these sequences. However, the segregation of charges decreased ΔF for the E−K, LAF1, and DDX4 sequences, highlighting the pronounced electrostatic attraction between the oppositely charged residues of charge-rich IDPs. These results provide a molecular mechanistic description of our finding that the charge-segregated sequences exhibit slower dynamics compared to the uniformly charge-patterned sequences (**Fig. 2**). Further, the implications of sequence charge patterning on the interfacial and dense phase properties (**Fig. 3**) are in line with previous experiments, which showed that the screening of electrostatic interactions at high salt concentrations resulted in the decrease of the surface tension and viscosity of PGL3 protein droplets.[62]

Given that the interaction strength between the oppositely charged residues in the dense phase are highly sensitive to the charge patterning within the IDPs, we next investigated their

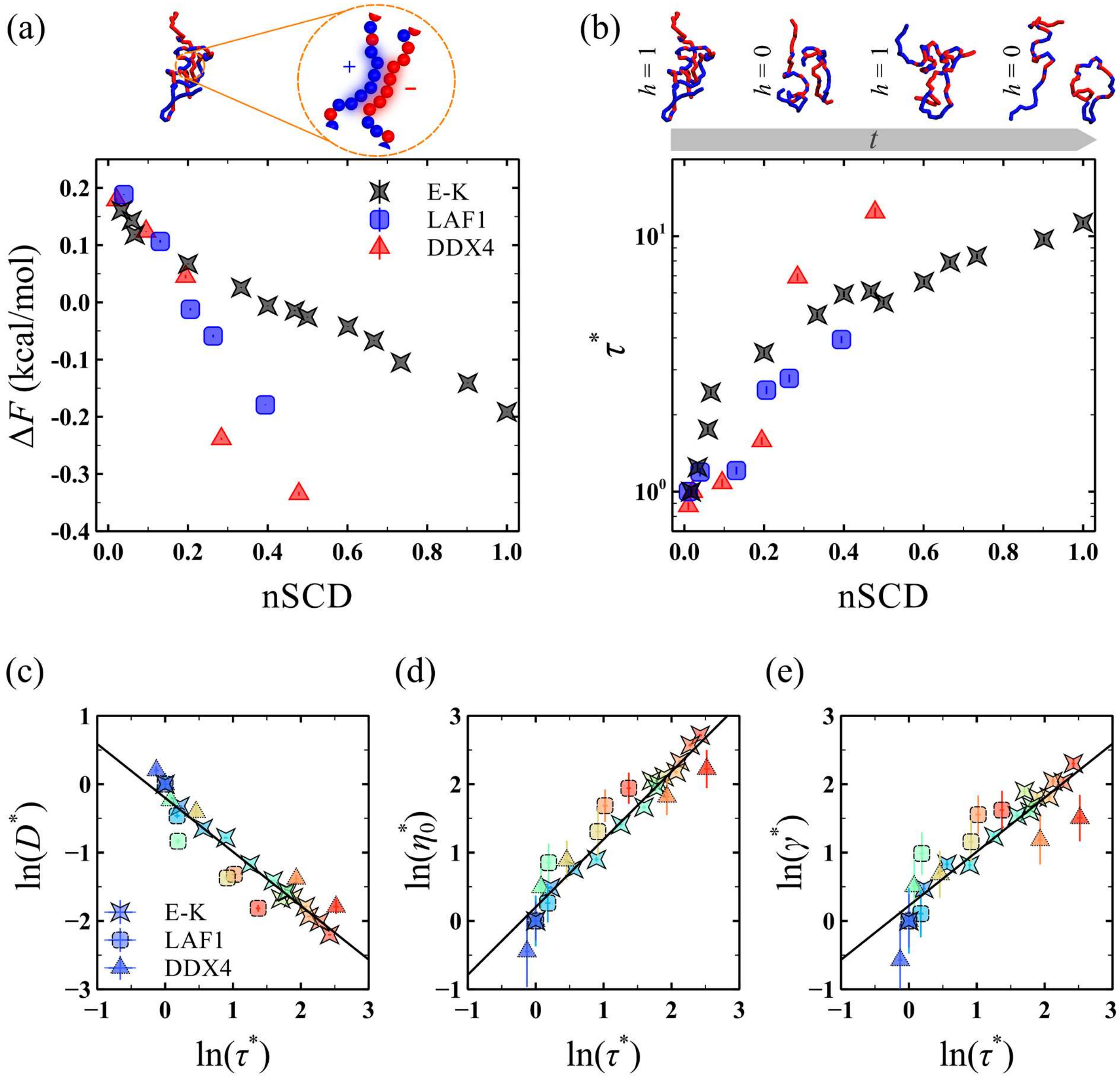

Fig. 4. (a) Free energy change ΔF between the bound and unbound states and (b) normalized intermittent contact lifetime τ^* for the interchain oppositely charged residues as a function of nSCD for the E−K, LAF1, and DDX4 sequences. The values of τ^* are obtained after normalizing τ by those computed for the same reference sequences used in the normalization of the material properties as in **Fig. 2**. In (b), an example for the time-evolution of the intermittent contact formation $(h(t) = 1)$ and breakage $(h(t) = 0)$ between a pair of chains for the E−K variant with nSCD = 0.468 is also shown. Correlation (c) between D^* and τ^*, (d) between η_0^* and τ^*, and (e) between γ^* and τ^* for all sequences. The symbol color, ranging from blue to red, indicates increasing nSCD. The solid lines in (c), (d), and (e) correspond to linear fits $\ln(D^*) = -0.79\ln(\tau^*) - 0.20$ with correlation coefficient $R^2 = 0.90$, $\ln(\eta_0^*) = 0.99\ln(\tau^*) + 0.20$ with $R^2 = 0.93$, and $\ln(\gamma^*) = 0.79\ln(\tau^*) + 0.22$ with $R^2 = 0.83$, respectively.

dynamics of contact formation, which would facilitate a direct comparison with the material properties of IDP condensates. We computed the intermittent contact time autocorrelation[53, 54]

$$c(t) = \frac{\langle h(0)\, h(t) \rangle}{\langle h^2(0) \rangle},$$

with step function $h(t) = 1$ if a pair of oppositely charged residues i and j between two chains were in contact (*i.e.,* if the distance between the pair was less than the cut-off radius of 1.5σ, where σ is the average diameter of residues i and j) at time $t = 0$ and at time t, irrespective of whether they stayed in contact at intermittent times, and $h(t) = 0$ if not (**Fig. 4b**). This definition of autocorrelation accounts for the contacts between residues that get broken and reformed again, thus measuring the duration for which residues from two chains remain in the same vicinity. We found that c (averaged over all interchain pairs of oppositely charged residues) decayed differently with increasing charge segregation for the model and natural proteins (**Fig. S6**). To quantify the variations in c, we computed the intermittent contact lifetime τ (**Fig. 4b**) by integrating c up to the time for which it reached a value of 0.05.[53, 54, 64] Similar to the representation of different material properties, we normalized τ by the corresponding value of the reference sequences, giving rise to τ^*.

Interestingly, τ^* increased with increasing nSCD, in line with the observed decrease in ΔF, for the E−K, LAF1, and DDX4 sequences, further highlighting the role of protein sequence *via* charge distribution. Given that the different material properties of IDPs also depended on the charge distribution along the sequence, we hypothesized that the dynamics of contact formation and breakage between the oppositely charged residues could predominantly contribute to such condensed phase properties. To validate this hypothesis, we tested for a correlation between τ^* and each of the different material properties D^*, η_0^*, and γ^* by plotting a logarithmic version of these quantities against each other (**Fig. 4c-e**). In this representation, the simulation data of each material property fell on a master curve that followed a power-law relation with the contact dynamics, as quantified by the correlation coefficients ranging from $R^2 = 0.83$ to $R^2 = 0.93$. These observations validated our hypothesis that a longer contact lifetime between the oppositely charged residues contributed to a lower diffusion coefficient, a higher viscosity, and a higher surface tension of the charge-rich IDP condensates. Further, the dynamics of contact formation and breakage were highly transient (**Fig. S6**), consistent with the recent finding that the local interactions were exceedingly rapid in the phase-separated condensates formed by two highly charged polyelectrolytic IDPs like histone H1 and prothymosin α.[64] This rapid dynamics of formation and breakage of contacts between oppositely charged residues, in conjunction with its correlation to the material properties, implies that short equilibrium simulations for characterizing the contact dynamics can help to infer the condensed phase properties of charge-rich IDPs.

Condensed phase material properties and single-chain structural properties of charged disordered proteins are strongly correlated. It is well-established that the segregation of charges within the polyampholyte sequences alters their single-chain interactions.[22, 37, 65] Having shown that the interactions in the dense phase are highly sensitive to the charge patterning within the polyampholytic IDPs, we next investigated whether they shared similar features with the corresponding self-interactions at the single-chain level. For this purpose, we computed the probability of a pair of residues i and j to be in contact as $P = \langle n_{ij} \rangle$, with $n_{ij} = 1$ if the distance between the pair was less than the cut-off radius of 1.5σ and $n_{ij} = 0$ if not. We then summed over the contact probabilities $\sum P$ of a residue i with all other oppositely charged residues or like charged residues or uncharged residues in the dense phase and within a single chain for a given sequence. By comparing the $\sum P$ values in the dense phase with those obtained at the single-chain

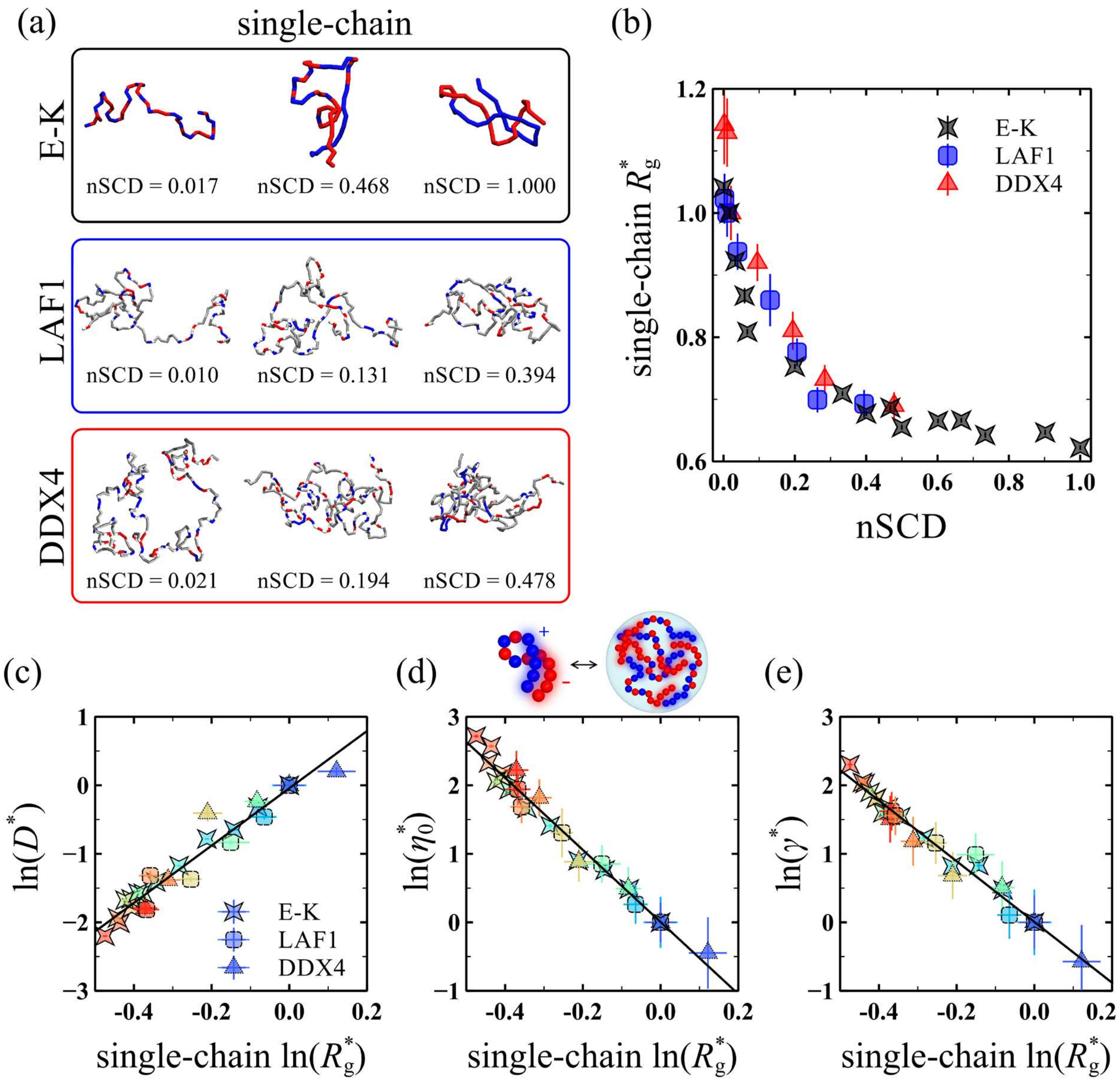

Fig. 5. (a) Representative single-chain conformations chosen close to the mean radius of gyration R_g of a given nSCD shown in **Fig. S10** for select E−K, LAF1, and DDX4 sequences. Negatively charged, positively charged, and neutral residues are shown in red, blue, and gray colors, respectively. (b) Normalized single-chain radius of gyration R_g^*, obtained after normalizing R_g by those computed for the same reference sequences used in the normalization of the material properties as in **Fig. 2**, as a function of nSCD for all sequences. Correlation (c) between D^* and R_g^*, (d) between η_0^* and R_g^*, and (e) between γ^* and R_g^* for all sequences. The symbol color, ranging from blue to red, indicates increasing nSCD. The solid lines in (c), (d), and (e) correspond to the linear fits $\ln(D^*) = 4.18\ln\left(R_g^*\right) - 0.04$ with correlation coefficient $R^2 = 0.94$, $\ln(\eta_0^*) = -5.26\ln\left(R_g^*\right) + 0.005$ with $R^2 = 0.98$, and $\ln(\gamma^*) = -4.43\ln\left(R_g^*\right) + 0.005$ with $R^2 = 0.97$, respectively.

level, we found that they were well-correlated across the entire nSCD range for the E−K, LAF1, and DDX4 sequences (**Figs. S7, S8, and S9**). These observations indicated that similar molecular interactions are at play for a charge-rich IDP in the dense phase as well as in the ultra-dilute phase.[40, 64] This similarity is the basis for using the single-chain conformations (quantified *via* the radius of gyration) as an indicator of the propensity of model IDP solutions and naturally occurring IDP solutions to phase separate into a condensate (quantified *via* the critical temperature or the saturation concentration).[40, 41] In particular, our previous work[41] has established that single-chain simulations of coil-to-globule transitions are powerful tools to interpret the phase behavior of protein sequences. Subsequent studies have shown that the molecular driving forces behind single-chain conformations and condensate formation can be decoupled when the NCPR of the sequence keeps deviating from zero[42] or the hydrophobic residues are clustered in the sequence.[33, 66, 67] However, such deviations were not observed irrespective of the patterning of the oppositely charged residues in IDP sequences.[40, 41]

Based on the above considerations, we next asked whether the single-chain properties of charge-rich IDPs with different NCPR values and with varying distributions of oppositely charged residues could serve as an indicator of the trends in the material properties of the condensates formed by them. We characterized the single-chain conformations of the E−K,[37] LAF1, and DDX4 sequence variants through their radius of gyration $R_\mathrm{g} = \langle G_1 + G_2 + G_3 \rangle^{1/2}$, which we computed from the eigenvalues G_i of the gyration tensor (**Figs. S10 and 5a**).[37] In line with previous computational studies,[22, 37, 49] we found that R_g decreased with increasing nSCD for the model proteins and natural proteins, because of pronounced electrostatic attraction between the oppositely charged residues within the sequence. Further, R_g of the natural proteins were significantly larger than the model proteins because of their longer chain lengths. Thus, to eliminate the chain length dependence of R_g, we normalized it for each sequence variant by the corresponding value of the reference sequences to define R_g^* (**Fig. 5b)**, similar to the normalized representation of the different material properties. Analogous to the trends seen for the different material properties, we found that the rate of decrease in R_g^* in response to the segregation of charges within the sequences of different compositions were nearly identical, thus substantiating the predominant role of electrostatic interactions for their single-chain properties. Next, we tested for the correlation between R_g^* and each of the different material properties D^*, η_0^*, and γ^* of charge-rich IDPs (**Fig. 5c-e**). We found that the simulation data of all sequences collapsed onto a master curve, resulting in a strong power-law correlation between the single-chain size and the different condensed phase properties, with correlation coefficients $R^2 \approx 0.96$. Interestingly, we found that the magnitude of the power-law exponent ~ 5 describing the correlation between the different material properties and single-chain R_g was highly similar to that obtained for the correlation between the critical temperature and single-chain R_g of charge-rich IDPs.[40] Further, we found that the coupling between the single-chain conformations and the condensate material properties persists even when they are computed based on two different hydropathy scales (**Fig. S11**), indicating that the observed correspondence does not depend on the model choice. In general, the higher the single-chain compaction of IDPs induced by sequence charge segregation, the slower the dynamics of the condensates formed by them. These findings highlight that the extent of single-chain compactness can serve as a useful indicator of the material state of IDPs, emergent upon their phase separation to form condensates.

Discussion

Investigating the sequence determinants of the material properties of disordered protein condensates is exciting for two reasons: (1) it can help predict how sequence alterations perturb the intermolecular interactions, thereby affecting the condensate material properties that could modify biological function, and (2) it will enable the development of sequence design principles for engineering condensates with targeted material properties. Using a physics-based transferable CG framework[68, 69] that has been successfully applied to decipher the sequence-phase behavior relationship of IDPs, we performed a molecular-level investigation of how the sequence of the charged disordered proteins dictates their condensate material properties. The fact that the specific arrangement of the charges within an IDP sequence can influence its conformations[22, 37] and phase behavior[32, 40] makes it an attractive feature to decipher the sequence-material property relationship of IDPs. Thus, we employed the charge-patterned variants of a model protein consisting of an equal number of glutamic acid and lysine residues as well as the charge-patterned variants of disordered regions of two naturally occurring proteins, namely LAF1 and DDX4. Our characterization of the condensates formed by these protein variants through the equilibrium and non-equilibrium MD simulations revealed that charge patterning led to monotonic changes in different material properties despite their very different sequence compositions: with increasing charge segregation, the diffusion coefficient of the protein chains within the dense phase decreased, while the dense phase viscosity and the surface tension at the condensate-water interface increased. Moreover, we found that the computed material properties from our simulations were quantitatively comparable with experimental measurements of charge-rich protein condensates. Further, the rate of change of these material properties with changing charge distribution was nearly identical between the model and natural proteins, thus highlighting the interdependence between these properties across a wide range of sequences. These observations highlight that sequence charge patterning can help modulate the material properties within the dense phase as well as at the interface of charge-rich IDP condensates, without changing external conditions such as temperature and salt concentration.

The sensitivity of the material properties to the segregation of charges in the sequence suggested that the interchain electrostatic interactions played a dominant role in influencing them. This hypothesis was supported by our finding that the intermittent lifetime of contacts between the oppositely charged residues increased when the charges were segregated for both the model and natural proteins. This behavior is consistent with the associative polymer theories,[70-72] which imply that stronger interchain interactions would lead to slower condensate dynamics. Interestingly, we found a strong universal correlation between the lifetime of noncovalent bonds formed by the oppositely charged residues and the different material properties of all investigated sequences, despite their different sequence characteristics. However, we found the dynamics of contact formation and breakage to be highly transient, indicative of the local environment within the dense phase showing a dynamic exchange of residue partners. This rapid contact dynamics within the condensates is thought to be the reason behind their ability to facilitate biochemical reactions at a fast rate.[64] Taken together, we concluded that short equilibrium simulations to characterize the contact lifetimes can provide insights into the material properties of charge-rich IDP condensates.

For our model and natural proteins, we found that the interactions in the dense phase were highly similar to that of the interactions for a single-chain. Such a similarity rationalized the observed strong correlation between the different material properties and the single-chain structural properties, as quantified *via* the radius of gyration, for all sequences, a pivotal discovery of this work. Our findings indicate that the sequence-level attributes of IDPs can be harnessed to modulate the material properties emergent upon their condensation, allowing for the rational design of synthetic membraneless organelles. Our predictions regarding the dense phase and interfacial properties of IDP condensates can be tested through routine experimental characterizations: the translational diffusion coefficient can be measured through FRAP experiments[64] or through fluorescence correlation spectroscopy (FCS),[16] while the viscosity of the condensates can be obtained using single particle tracking (SPT) microrheology[27] (see **Tables S2 and S3** for method details). Direct measurements of surface tension can be achieved through the use of dual optical traps (dual-OT) for stretching the condensates[63] (see **Table S3** for method details). Additionally, it is feasible to deduce the condensate material properties by studying the conformations of an isolated protein chain, which are accessible through fully atomistic simulations or experiments, namely *via* single-molecule Förster resonance energy transfer (FRET)[73] and small-angle X-ray scattering (SAXS).[74] We believe that the molecular insights provided in this work will aid in further investigations of diverse sequence features in dictating condensate dynamics, leading to a comprehensive molecular language for the material properties of biomolecular condensates.

Methods

Generation of sequences studied. To investigate the effect of charge patterning on the material properties of charged disordered proteins, we employed model polyampholytic sequences comprised of negatively charged glutamic acid (E) residues and positively charged lysine (K) residues. These sequences, which had zero net charge, were selected from a large set of E−K variants generated in our previous study,[37] such that they spanned the entire range of nSCD (*i.e.*, from a perfectly alternating sequence with nSCD $= 0$ to a diblock sequence with nSCD $= 1$). We also investigated the charge patterning effects on sequences with a nonzero net charge by simulating the disordered regions of two naturally occurring proteins, namely LAF1 and DDX4.[29,

[32] Specifically, we took the wild-type sequences of LAF1's RGG domain and DDX4's N-terminal domain, and shuffled the residues within them to generate respective variants with increasing charge segregation. This procedure allowed us to investigate the LAF1 disordered variants with nSCD ranging from 0.003 to 0.394 and the DDX4 disordered variants with nSCD ranging from 0.002 to 0.478. We note that nSCD is strongly correlated with the parameter κ (quantified based on the local and global charge asymmetries within an IDP)[22] for all investigated sequences (**Fig. S12**), indicating that our interpretations regarding the condensate material properties and single-chain conformations are not influenced by the choice of the patterning parameter.

Model. We used our recently developed CG framework to model the IDP sequences as flexible chains with a single bead per residue representation.[68, 69] Directly bonded residues interacted with each other *via* the harmonic potential

$$U_{\mathrm{b}}(r) = \frac{k_{\mathrm{b}}}{2}(r - r_0)^2,$$

with distance r between residues, spring constant $k_{\mathrm{b}} = 20\,\mathrm{kcal}/(\mathrm{mol}\,\text{Å}^2)$, and equilibrium bond length $r_0 = 3.8\,\text{Å}$. Interactions between nonbonded residues were modeled using a modified Lennard-Jones potential (LJ) that facilitates the attraction between residues i and j to be scaled independently of the short-range repulsion by their average hydropathy $\lambda = (\lambda_i + \lambda_j)/2$:[75, 76]

$$U_{\mathrm{vdW}}(r) = \begin{cases} U_{\mathrm{LJ}}(r) + (1-\lambda)\varepsilon, & r \le 2^{1/6}\sigma \\ \lambda U_{\mathrm{LJ}}(r), & \text{otherwise} \end{cases},$$

where U_{LJ} is the standard LJ potential

$$U_{\mathrm{LJ}}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right].$$

The parameters of U_{vdW} include the average diameter $\sigma = (\sigma_i + \sigma_j)/2$ of residues i and j, and the interaction strength $\varepsilon = 0.2\ \mathrm{kcal/mol}$. For the E−K sequences, the hydropathy λ values based on the Kapcha-Rossky scale were used,[68, 77] while for the LAF1 and DDX4 sequences, the λ values were based on the Urry scale[78] (unless otherwise specified), which captures the changes in the phase behavior of natural proteins upon mutations of arginine to lysine and tyrosine to phenylalanine.[69, 78] We truncated the pair potential U_{vdW} and its forces to zero at a distance of 4 σ. Further, the nonbonded charged residues interacted *via* a Coulombic potential with Debye-Hückel electrostatic screening[79]

$$U_{\mathrm{e}}(r) = \frac{q_i q_j}{4\pi\epsilon_r\epsilon_0 r} e^{-r/\ell},$$

with vacuum permittivity ϵ_0, relative permittivity $\epsilon_r = 80$, and Debye screening length $\ell = 10$ Å. The chosen values of ϵ_r and ℓ correspond to an aqueous solution with a physiological salt concentration of $\sim$100 mM. We truncated the electrostatic potential and its forces to zero at a distance of 35 Å.

Simulation methods. For characterizing the translational motion of IDP chains within the dense phase, we simulated the charge-patterned variants in a cubic simulation box at a constant pressure of $P = 0$ atm for a total duration of 0.5 μs. The sequence variants attained their preferred dense phase concentration ρ at the end of this simulation run, after which we switched to Langevin dynamics (LD) simulations in the canonical ensemble to simulate for a duration of 1 μs. For all the sequences, a damping factor of $t_{\mathrm{damp}} = 1$ ns was used to set the friction coefficient of a residue of mass m_i in the chain to $f_i = m_i/t_{\mathrm{damp}}$. The first 0.2 μs of LD simulations was considered as the equilibration period and the remaining 0.8 μs of the simulation trajectory was used for the computation of the MSD of the residues within a chain in the dense phase, from which the translational diffusivity D was extracted. For sufficiently long times $t \gg \tau_R$, τ_R being the Rouse relaxation time of the entire chain, the MSD of the residues in a chain becomes identical to that of the chain's center of mass,[48] which is the case for all investigated E−K, LAF1, and DDX4 sequences (**Figs. S13a-S15a**). We further corroborated this by computing the MSD of residues relative to the chain's center of mass g_2, which plateaued at long times for all sequences (**Figs. S13b-S15b**), indicating that the reported D values are indeed an accurate representation of the translational motion of the IDP chains within the dense phase of the condensates.

We characterized the shear viscosity η of the dense phase of IDP condensates through the nonequilibrium MD (NEMD) simulation technique.[50-52] Specifically, the charge-patterned variants at their preferred concentration ρ were subjected to a steady shear strain in the x-direction at different shear rates $\dot{\gamma}$ using the SLLOD equations of motion and the sliding-brick periodic boundary conditions. We observed a linear velocity v_x profile in the gradient direction y and found that its slope, which gives the shear rate, is nearly the same as that of the applied value for all IDP sequences (**Figs. S16-S18**). This behavior confirmed that the system is responding as intended at both low and high $\dot{\gamma}$ used in this work. The resulting shear stress τ_{xy} was then measured as a function of time. We used the $\dot{\gamma}$ and τ_{xy} values to determine η of the dense phase of protein sequences. These simulations were carried out for a total duration of 0.8 μs in the case of E−K sequences and 1.2 μs in the case of LAF1 and DDX4 sequences. For both model and natural proteins, we considered the trajectories of the first 0.05 μs duration as part of the equilibration period and hence, discarded them in the computation of η. We also verified that η converged to its mean value at times shorter than the simulation duration (**Fig. S19**).

The surface tension γ of the IDP condensates was characterized by performing the LD simulations of protein chains in a slab geometry. These simulations were carried out using a damping factor of $t_{\mathrm{damp}} = 1$ ns to set the residue friction coefficient f_i in the same way as was done in the simulations of protein sequences in a cubic geometry. The proteins chains were initially placed in a dense slab within the rectangular simulation boxes (150 Å × 150 Å × 1200 Å) and were simulated for a duration of 3 μs. The last 2.5 μs simulation trajectory was used in computing the γ values. Again, we ensured that the total simulation time was much longer than the time it took for γ to converge to the reported mean value for the investigated IDP sequences (**Fig. S20**).

To characterize the effect of charge patterning on the single-chain conformations of the charged IDPs, we previously simulated the E−K variants in a cubic box of edge length 160 Å.[37] We used the same data in this study for establishing its correlations with the material properties of the condensates formed by the E−K variants. In this work, we additionally characterized the single-chain conformations of the LAF1 and DDX4 disordered sequence variants by placing them in a cubic box of edge length 1000 Å. We chose such large box sizes to prevent unphysical self-interactions between the chain and its periodic images. These simulations were carried out for a duration of 1 μs.

All results for the E−K and DDX4 sequence variants were obtained at a fixed temperature of $T = 300$ K. The results of LAF1 sequence variants were obtained at $T = 280$ K as the critical temperature of its WT obtained from experiments[32] and our CG model[69] is $T \approx 297$ K. Further, the physical quantities were averaged over two independent replicas, with each replica divided into five blocks for estimating the standard error of mean. A total of 500 chains were simulated, with periodic boundary conditions applied to all three Cartesian directions and with a timestep of

10 fs. The equilibrium dense phase simulations and the single-chain simulations of the IDPs were carried out using HOOMD-blue (version 2.9.3)[80] with features extended using azplugins (version 0.10.1).[81] The non-equilibrium simulations of the IDPs were carried out using LAMMPS (29 October 2020 version).[82]

Connecting simulation and experimental timescales. Using Å, g/mol, and kcal/mol as our units of length, mass, and energy, respectively, the characteristic simulation time is typically defined as $\tau_{\mathrm{MD}} = \sqrt{(\mathrm{g/mol})\ \mathrm{Å}^2/(\mathrm{kcal/mol})} = 48.89$ fs. However, this choice for τ_{MD} is usually not suitable for interpreting the time-dependent mechanical properties from CG simulations, as it reflects the thermal fluctuations of the monomer beads, which lack the atomic-level representation of the residues.[83] Instead, we systematically derived experimentally-relevant timescales based on the translation diffusive motion of a single protein chain in dilute solution in what follows.

In LD simulations, the Rouse dynamics describes the long-time diffusion coefficient of a single-chain[84]

$$D_{\mathrm{sim}} = \frac{k_{\mathrm{B}}T}{Nf_i},$$

where $k_{\mathrm{B}}T$ is the energy scale and f_i is the friction coefficient of monomer i in a chain. In terms of LJ units, we computed $D_{\mathrm{sim}} = 27.99\sqrt{\varepsilon\sigma^2/m}$, $9.07\sqrt{\varepsilon\sigma^2/m}$, and $6.45\sqrt{\varepsilon\sigma^2/m}$ for single-chain polymers of lengths corresponding to the E−K ($N = 50$), LAF1 ($N = 168$), and DDX4 ($N = 236$) sequences. These D_{sim} values were obtained for the f_i used in our dense phase simulations.

Next, we estimated the experimentally expected diffusion coefficient of the single-chain IDP sequences in water using the Stokes-Einstein relation[84]

$$D_{\mathrm{exp}} = \frac{k_{\mathrm{B}}T}{6\pi\eta_{\mathrm{w}}R_{\mathrm{h}}},$$

where η_{w} is the viscosity of water and R_{h} is the hydrodynamic radius of the chain. Given that the E−K and DDX4 sequences were simulated at $T = 300\ \mathrm{K}$, while the LAF1 sequences were simulated at $T = 280\ \mathrm{K}$, we used $\eta_{\mathrm{w}}(T = 300\ K) = 0.854\ \mathrm{mPa.s}$ and $\eta_{w}(T = 280\ \mathrm{K}) = 1.434\ \mathrm{mPa.s}$, respectively,[85] for computing D_{exp}. Further, we used $R_{\mathrm{h}} = 16.38$ Å (well-mixed E−K), 26.56 Å (LAF1 wild-type), and 29.69 Å (DDX4 wild-type) obtained based on the Kirkwood approximation[86, 87] from our simulations. We matched D_{exp} with D_{sim} to get timescales of 1.782×10^{-9} s, 1.683×10^{-9} s, and 0.744×10^{-9} s for the E−K, LAF1, and DDX4, respectively. We used these timescales, obtained based on the mapping at the single-chain level for which the solvent viscosity dominates, to interpret the dense phase dynamical and rheological properties of our IDP sequences.

Experimental measurements of viscosity and surface tension. We state the different experimental methods used by the literature studies[15, 63] for measuring the viscosity and surface tension of a wide range of protein condensates (**Table S3**) that we have used for comparison with the corresponding simulation measurements of our IDP sequences in **Fig. 3b**. All values of the viscosity of condensates were measured directly through experiments such as micropipette aspiration and single particle tracking microrheology. However, 3 out of 11 experimental measurements of surface tension (*i.e.*, for polyK, polyR, and $[\mathrm{RGRGG}]_5$ + dT40) were indirectly estimated based on the viscocapillary model, which assumes condensates as a purely viscous medium.[16] Recent studies have shown that the condensates display a time-dependent viscoelastic behavior,[14, 63, 88, 89] a finding that questions the use of the viscocapillary model. We and others have also shown that many condensates formed by disordered proteins remain predominantly viscous over the experimentally measured time.[24, 64] Thus, we assumed that the surface tension data obtained based on the viscocapillary model are reliable "estimates" at least for the disordered protein droplets that show predominantly viscous characteristics.

To justify our assumption further, we used the directly measured values of fusion time τ_f, viscosity η, and surface tension γ for four protein condensates formed by oppositely charged binary mixtures (**Table S4**).[63] Using the values of τ_f and η, we estimated the surface tension based on the viscocapillary model $\gamma = 1.97\eta R/\tau_f$ (**Table S4**), where R corresponds to the droplet radius. We used $R = 3$ μm to estimate surface tension, the value for which the fusion times were reported. We found that for the two condensates formed by fully-disordered proteins (polyK+heparin and proline-rich motif+heparin), the estimated surface tension was in good agreement with the directly measured surface tension values (a relative difference of $< 6\%$ when considering the upper bound of the directly measured surface tension value). However, when one of the components consisted of folded domains (SH3 domain+proline-rich motif), the relative difference increased to $\sim 15\%$, but the estimated value is still in the comparable range with the measured value. The viscocapillary model does poorly when involving condensates made of components that all have folded regions (SH3 domain+lysozyme). Thus, the viscocapillary model seems to be a reasonable choice for estimating the surface tension of condensates formed by fully disordered proteins but not necessarily for those involving folded domains.

Data Availability

Data presented in this work are available from the corresponding authors upon reasonable request. Additional data supporting the main findings are available in the Supplementary Information.

Code Availability

Codes to run the CG simulations and perform the analyses discussed in this work are available at https://glotzerlab.engin.umich.edu/hoomd-blue/, https://www.lammps.org/, and https://www.mdanalysis.org/. Additionally, the analysis codes pertaining to some of the physical quantities presented in this work are available via the GitHub repository https://github.com/dsd993/CondensateMaterialPropertyAnalysis.

Acknowledgments

D.S.D. thanks Shiv Rekhi for general discussions on disordered proteins. This research article is based on the work supported by the National Institute of General Medical Science of the National Institutes of Health under the grant R01GM136917 and the Welch Foundation under the grant A-2113-20220331. 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 through the U.S. Naval Research Laboratory base program. The computational resources provided by the Texas A&M High Performance Research Computing (HPRC) are gratefully acknowledged.

Competing Interests

The authors declare no competing interests.

Author Contributions

D.S.D. and J.M. conceived the research presented in this material. D.S.D designed the sequences and conducted the simulations. D.S.D., A.N., Y.C.K., and J.M. designed the analyses. D.S.D. and J.W. analyzed the simulation data. D.S.D., A.N., and J.M. wrote and edited the manuscript with critical input from others.

References

(1) Lasker, K.; Boeynaems, S.; Lam, V.; Scholl, D.; Stainton, E.; Briner, A.; Jacquemyn, M.; Daelemans, D.; Deniz, A.; Villa, E.; et al. The material properties of a bacterial-derived biomolecular condensate tune biological function in natural and synthetic systems. *Nature Communications* **2022**, *13* (1), 5643.

(2) Yamasaki, A.; Alam, J. M.; Noshiro, D.; Hirata, E.; Fujioka, Y.; Suzuki, K.; Ohsumi, Y.; Noda, N. N. Liquidity Is a Critical Determinant for Selective Autophagy of Protein Condensates. *Molecular Cell* **2020**, *77* (6), 1163-1175.

(3) Agudo-Canalejo, J.; Schultz, S. W.; Chino, H.; Migliano, S. M.; Saito, C.; Koyama-Honda, I.; Stenmark, H.; Brech, A.; May, A. I.; Mizushima, N.; et al. Wetting regulates autophagy of phase-separated compartments and the cytosol. *Nature* **2021**, *591* (7848), 142-146.

(4) Brangwynne, C. P.; Mitchison, T. J.; Hyman, A. A. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proceedings of the National Academy of Sciences* **2011**, *108* (11), 4334-4339.

(5) Shin, Y.; Chang, Y.-C.; Lee, D. S. W.; Berry, J.; Sanders, D. W.; Ronceray, P.; Wingreen, N. S.; Haataja, M.; Brangwynne, C. P. Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. *Cell* **2018**, *175* (6), 1481-1491.

(6) Schmidt, H. B.; Görlich, D. Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. *eLife* **2015**, *4*, e04251.

(7) Patel, A.; Lee, Hyun O.; Jawerth, L.; Maharana, S.; Jahnel, M.; Hein, Marco Y.; Stoynov, S.; Mahamid, J.; Saha, S.; Franzmann, Titus M.; et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **2015**, *162* (5), 1066-1077.

(8) Murakami, T.; Qamar, S.; Lin, Julie Q.; Schierle, Gabriele S. K.; Rees, E.; Miyashita, A.; Costa, Ana R.; Dodd, Roger B.; Chan, Fiona T. S.; Michel, Claire H.; et al. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **2015**, *88* (4), 678-690.

(9) Molliex, A.; Temirov, J.; Lee, J.; Coughlin, M.; Kanagaraj, Anderson P.; Kim, Hong J.; Mittag, T.; Taylor, J. P. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **2015**, *163* (1), 123-133.

(10) Wang, Z.; Lou, J.; Zhang, H. Essence determines phenomenon: Assaying the material properties of biological condensates. *Journal of Biological Chemistry* **2022**, *298* (4), 101782.

(11) Li, W.; Hu, J.; Shi, B.; Palomba, F.; Digman, M. A.; Gratton, E.; Jiang, H. Biophysical properties of AKAP95 protein condensates regulate splicing and tumorigenesis. *Nature Cell Biology* **2020**, *22* (8), 960-972.

(12) Mohanty, P.; Kapoor, U.; Sundaravadivelu Devarajan, D.; Phan, T. M.; Rizuan, A.; Mittal, J. Principles Governing the Phase Separation of Multidomain Proteins. *Biochemistry* **2022**, *61* (22), 2443-2455.

(13) 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. *Nature Chemistry* **2017**, *9* (11), 1118-1125.

(14) Jawerth, L.; Fischer-Friedrich, E.; Saha, S.; Wang, J.; Franzmann, T.; Zhang, X.; Sachweh, J.; Ruer, M.; Ijavi, M.; Saha, S.; et al. Protein condensates as aging Maxwell fluids. *Science* **2020**, *370* (6522), 1317-1323.

(15) Wang, H.; Kelley, F. M.; Milovanovic, D.; Schuster, B. S.; Shi, Z. Surface tension and viscosity of protein condensates quantified by micropipette aspiration. *Biophysical Reports* **2021**, *1* (1), 100011.

(16) Alshareedah, I.; Thurston, G. M.; Banerjee, P. R. Quantifying viscosity and surface tension of multicomponent protein-nucleic acid condensates. *Biophysical Journal* **2021**, *120* (7), 1161-1169.

(17) Alshareedah, I.; Moosa, M. M.; Pham, M.; Potoyan, D. A.; Banerjee, P. R. Programmable viscoelasticity in protein-RNA condensates with disordered sticker-spacer polypeptides. *Nature Communications* **2021**, *12* (1), 6620.

(18) 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. *Proceedings of the National Academy of Sciences* **2015**, *112* (23), 7189-7194.

(19) Burke, Kathleen A.; Janke, Abigail M.; Rhine, Christy L.; Fawzi, Nicolas L. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Molecular Cell* **2015**, *60* (2), 231-241.

(20) 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. *Molecular Cell* **2018**, *69* (3), 465-479.

(21) Zheng, W.; Dignon, G. L.; Jovic, N.; Xu, X.; Regy, R. M.; Fawzi, N. L.; Kim, Y. C.; Best, R. B.; Mittal, J. Molecular Details of Protein Condensates Probed by Microsecond Long Atomistic Simulations. *The Journal of Physical Chemistry B* **2020**, *124* (51), 11671-11679.

(22) Das, R. K.; Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proceedings of the National Academy of Sciences* **2013**, *110* (33), 13392-13397.

(23) Bigman, L. S.; Iwahara, J.; Levy, Y. Negatively Charged Disordered Regions are Prevalent and Functionally Important Across Proteomes. *Journal of Molecular Biology* **2022**, *434* (14), 167660.

(24) Rekhi, S.; Garcia, C. G.; Barai, M.; Rizuan, A.; Schuster, B. S.; Kiick, K. L.; Mittal, J. Expanding the molecular language of protein liquid-liquid phase separation. *bioRxiv* **2023**, DOI: 10.1101/2023.03.02.530853.

(25) Rhine, K.; Makurath, M. A.; Liu, J.; Skanchy, S.; Lopez, C.; Catalan, K. F.; Ma, Y.; Fare, C. M.; Shorter, J.; Ha, T.; et al. ALS/FTLD-Linked Mutations in FUS Glycine Residues Cause Accelerated Gelation and Reduced Interactions with Wild-Type FUS. *Molecular Cell* **2020**, *80* (4), 666-681.

(26) Monahan, Z.; Ryan, V. H.; Janke, A. M.; Burke, K. A.; Rhoads, S. N.; Zerze, G. H.; O'Meally, R.; Dignon, G. L.; Conicella, A. E.; Zheng, W.; et al. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *The EMBO Journal* **2017**, *36* (20), 2951-2967.

(27) Fisher, R. S.; Elbaum-Garfinkle, S. Tunable multiphase dynamics of arginine and lysine liquid condensates. *Nature Communications* **2020**, *11* (1), 4628.

(28) Lin, Y.; Currie, S. L.; Rosen, M. K. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *Journal of Biological Chemistry* **2017**, *292* (46), 19110-19120.

(29) Brady, J. P.; Farber, P. J.; Sekhar, A.; Lin, Y.-H.; Huang, R.; Bah, A.; Nott, T. J.; Chan, H. S.; Baldwin, A. J.; Forman-Kay, J. D.; et al. Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proceedings of the National Academy of Sciences* **2017**, *114* (39), E8194-E8203.

(30) Wang, J.; Choi, J.-M.; Holehouse, A. S.; Lee, H. O.; Zhang, X.; Jahnel, M.; Maharana, S.; Lemaitre, R.; Pozniakovsky, A.; Drechsel, D.; et al. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **2018**, *174* (3), 688-699.

(31) 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. *Nature Structural & Molecular Biology* **2019**, *26* (7), 637-648.

(32) Schuster, B. S.; Dignon, G. L.; Tang, W. S.; Kelley, F. M.; Ranganath, A. K.; Jahnke, C. N.; Simpkins, A. G.; Regy, R. M.; Hammer, D. A.; Good, M. C.; et al. Identifying sequence perturbations to an intrinsically disordered protein that determine its phase-separation behavior. *Proceedings of the National Academy of Sciences* **2020**, *117* (21), 11421-11431.

(33) 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* **2020**, *367* (6478), 694-699.

(34) Welles, R. M.; Sojitra, K. A.; Garabedian, M. V.; Xia, B.; Regy, R. M.; Gallagher, E. R.; Mittal, J.; Good, M. C. Determinants of Disordered Protein Co-Assembly Into Discrete Condensed Phases. *bioRxiv* **2023**, DOI: 10.1101/2023.03.10.532134.

(35) Schuster, B. S.; Regy, R. M.; Dolan, E. M.; Kanchi Ranganath, A.; Jovic, N.; Khare, S. D.; Shi, Z.; Mittal, J. Biomolecular Condensates: Sequence Determinants of Phase Separation, Microstructural Organization, Enzymatic Activity, and Material Properties. *The Journal of Physical Chemistry B* **2021**, *125* (14), 3441-3451.

(36) Nott, Timothy J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowietz, A.; Craggs, T. D.; Bazett-Jones, David P.; Pawson, T.; Forman-Kay, Julie D.; et al. Phase Transition of a

Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell* **2015**, *57* (5), 936-947.

(37) 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.

(38) 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.

(39) Zhang, Y.; Pyo, A. G. T.; Jiang, Y.; Brangwynne, C. P.; Stone, H. A.; Wingreen, N. S. Interface resistance of biomolecular condensates. *bioRxiv* **2022**, DOI: 10.1101/2022.03.16.484641.

(40) Lin, Y.-H.; Chan, H. S. Phase Separation and Single-Chain Compactness of Charged Disordered Proteins Are Strongly Correlated. *Biophysical Journal* **2017**, *112* (10), 2043-2046.

(41) 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. *Proceedings of the National Academy of Sciences* **2018**, *115* (40), 9929-9934.

(42) Bremer, A.; Farag, M.; Borcherds, W. M.; Peran, I.; Martin, E. W.; Pappu, R. V.; Mittag, T. Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. *Nature Chemistry* **2022**, *14* (2), 196-207.

(43) Sawle, L.; Ghosh, K. A theoretical method to compute sequence dependent configurational properties in charged polymers and proteins. *The Journal of Chemical Physics* **2015**, *143* (8), 085101.

(44) Ghosh, K.; Huihui, J.; Phillips, M.; Haider, A. Rules of Physical Mathematics Govern Intrinsically Disordered Proteins. *Annual Review of Biophysics* **2022**, *51* (1), 355-376.

(45) Rana, U.; Brangwynne, C. P.; Panagiotopoulos, A. Z. Phase separation vs aggregation behavior for model disordered proteins. *The Journal of Chemical Physics* **2021**, *155* (12), 125101.

(46) Taylor, N. O.; Wei, M.-T.; Stone, H. A.; Brangwynne, C. P. Quantifying Dynamics in Phase-Separated Condensates Using Fluorescence Recovery after Photobleaching. *Biophysical Journal* **2019**, *117* (7), 1285-1300.

(47) Doi, M.; Edwards, S. *The Theory of Polymer Dynamics*. Clarendon Press, Oxford, 1986.

(48) Nikoubashman, A.; Howard, M. P. Equilibrium Dynamics and Shear Rheology of Semiflexible Polymers in Solution. *Macromolecules* **2017**, *50* (20), 8279-8289.

(49) Hazra, M. K.; Levy, Y. Charge pattern affects the structure and dynamics of polyampholyte condensates. *Physical Chemistry Chemical Physics* **2020**, *22* (34), 19368-19375.

(50) Evans, D.; Morriss, G. *Non-Equilibrium Statistical Mechanics of Liquids*. Cambridge University Press, Cambridge, 2008.

(51) Sundaravadivelu Devarajan, D.; Nourian, P.; McKenna, G. B.; Khare, R. Molecular simulation of nanocolloid rheology: Viscosity, viscoelasticity, and time-concentration superposition. *Journal of Rheology* **2020**, *64* (3), 529-543.

(52) Sundaravadivelu Devarajan, D.; Khare, R. Linear viscoelasticity of nanocolloidal suspensions from probe rheology molecular simulations. *Journal of Rheology* **2022**, *66* (5), 837-852.

(53) Ronceray, P.; Zhang, Y.; Liu, X.; Wingreen, N. S. Stoichiometry Controls the Dynamics of Liquid Condensates of Associative Proteins. *Physical Review Letters* **2022**, *128* (3), 038102.

(54) Xiang, Y.-X.; Shan, Y.; Lei, Q.-L.; Ren, C.-L.; Ma, Y.-Q. Dynamics of protein condensates in weak-binding regime. *Physical Review E* **2022**, *106* (4), 044403.

(55) Ghosh, A.; Zhou, H.-X. Determinants for Fusion Speed of Biomolecular Droplets. *Angewandte Chemie International Edition* **2020**, *59* (47), 20837-20840.

(56) Kirkwood, J. G.; Buff, F. P. The Statistical Mechanical Theory of Surface Tension. *The Journal of Chemical Physics* **1949**, *17* (3), 338-343.

(57) Benayad, Z.; von Bülow, S.; Stelzl, L. S.; Hummer, G. Simulation of FUS Protein Condensates with an Adapted Coarse-Grained Model. *Journal of Chemical Theory and Computation* **2021**, *17* (1), 525-537.

(58) Guenza, M. Thermodynamic consistency and other challenges in coarse-graining models. *The European Physical Journal Special Topics* **2015**, *224* (12), 2177-2191.

(59) Faller, R. Automatic coarse graining of polymers. *Polymer* **2004**, *45* (11), 3869-3876.

(60) Milano, G.; Müller-Plathe, F. Mapping Atomistic Simulations to Mesoscopic Models: A Systematic Coarse-Graining Procedure for Vinyl Polymer Chains. *The Journal of Physical Chemistry B* **2005**, *109* (39), 18609-18619.

(61) Nikoubashman, A.; Lee, V. E.; Sosa, C.; Prud'homme, R. K.; Priestley, R. D.; Panagiotopoulos, A. Z. Directed Assembly of Soft Colloids through Rapid Solvent Exchange. *ACS Nano* **2016**, *10* (1), 1425-1433.

(62) Jawerth, L. M.; Ijavi, M.; Ruer, M.; Saha, S.; Jahnel, M.; Hyman, A. A.; Jülicher, F.; Fischer-Friedrich, E. Salt-Dependent Rheology and Surface Tension of Protein Condensates Using Optical Traps. *Physical Review Letters* **2018**, *121* (25), 258101.

(63) Ghosh, A.; Kota, D.; Zhou, H.-X. Shear relaxation governs fusion dynamics of biomolecular condensates. *Nature Communications* **2021**, *12* (1), 5995.

(64) Galvanetto, N.; Ivanović, M. T.; Chowdhury, A.; Sottini, A.; Nüesch, M. F.; Nettels, D.; Best, R. B.; Schuler, B. Extreme dynamics in a biomolecular condensate. *Nature* **2023**, *619* (7971), 876-883.

(65) Huihui, J.; Ghosh, K. An analytical theory to describe sequence-specific inter-residue distance profiles for polyampholytes and intrinsically disordered proteins. *The Journal of Chemical Physics* **2020**, *152* (16), 161102.

(66) Bauer, D. J.; Stelzl, L. S.; Nikoubashman, A. Single-chain and condensed-state behavior of hnRNPA1 from molecular simulations. *The Journal of Chemical Physics* **2022**, *157* (15), 154903.

(67) 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. *The Journal of Physical Chemistry B* **2023**, *127* (17), 3829-3838.

(68) Dignon, G. L.; Zheng, W.; Kim, Y. C.; Best, R. B.; Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLOS Computational Biology* **2018**, *14* (1), e1005941.

(69) Regy, R. M.; Thompson, J.; Kim, Y. C.; Mittal, J. Improved coarse-grained model for studying sequence dependent phase separation of disordered proteins. *Protein Science* **2021**, *30* (7), 1371-1379.

(70) Semenov, A. N.; Rubinstein, M. Thermoreversible Gelation in Solutions of Associative Polymers. 1. Statics. *Macromolecules* **1998**, *31* (4), 1373-1385.

(71) Rubinstein, M.; Semenov, A. N. Thermoreversible Gelation in Solutions of Associating Polymers. 2. Linear Dynamics. *Macromolecules* **1998**, *31* (4), 1386-1397.

(72) Tanaka, F.; Edwards, S. F. Viscoelastic properties of physically crosslinked networks. 1. Transient network theory. *Macromolecules* **1992**, *25* (5), 1516-1523.

(73) Schuler, B.; Hofmann, H. Single-molecule spectroscopy of protein folding dynamics—expanding scope and timescales. *Current Opinion in Structural Biology* **2013**, *23* (1), 36-47.

(74) Bernadó, P.; Svergun, D. I. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Molecular BioSystems* **2012**, *8* (1), 151-167.

(75) Ashbaugh, H. S.; Hatch, H. W. Natively Unfolded Protein Stability as a Coil-to-Globule Transition in Charge/Hydropathy Space. *Journal of the American Chemical Society* **2008**, *130* (29), 9536-9542.

(76) Weeks, J. D.; Chandler, D.; Andersen, H. C. Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids. *The Journal of Chemical Physics* **1971**, *54* (12), 5237-5247.

(77) Kapcha, L. H.; Rossky, P. J. A Simple Atomic-Level Hydrophobicity Scale Reveals Protein Interfacial Structure. *Journal of Molecular Biology* **2014**, *426* (2), 484-498.

(78) Urry, D. W.; Gowda, D. C.; Parker, T. M.; Luan, C.-H.; Reid, M. C.; Harris, C. M.; Pattanaik, A.; Harris, R. D. Hydrophobicity scale for proteins based on inverse temperature transitions. *Biopolymers* **1992**, *32* (9), 1243-1250.

(79) Debye, P.; Hückel, E. De la theorie des electrolytes. I. abaissement du point de congelation et phenomenes associes. *Physikalische Zeitschrift* **1923**, *24* (9), 185-206.

(80) Anderson, J. A.; Glaser, J.; Glotzer, S. C. HOOMD-blue: A Python package for high-performance molecular dynamics and hard particle Monte Carlo simulations. *Computational Materials Science* **2020**, *173*, 109363.

(81) https://github.com/mphowardlab/azplugins. (accessed July 15, 2021).

(82) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **2022**, *271*, 108171.

(83) Padding, J. T.; Louis, A. A. Hydrodynamic interactions and Brownian forces in colloidal suspensions: Coarse-graining over time and length scales. *Physical Review E* **2006**, *74* (3), 031402.

(84) Rubinstein, M.; Colby, R. *Polymer Physics*. Oxford University Press, Oxford, 2003.

(85) Guevara-Carrion, G.; Vrabec, J.; Hasse, H. Prediction of self-diffusion coefficient and shear viscosity of water and its binary mixtures with methanol and ethanol by molecular simulation. *The Journal of Chemical Physics* **2011**, *134* (7).

(86) Kirkwood, J. G.; Riseman, J. The Intrinsic Viscosities and Diffusion Constants of Flexible Macromolecules in Solution. *The Journal of Chemical Physics* **2004**, *16* (6), 565-573.

(87) Aust, C.; Kröger, M.; Hess, S. Structure and Dynamics of Dilute Polymer Solutions under Shear Flow via Nonequilibrium Molecular Dynamics. *Macromolecules* **1999**, *32* (17), 5660-5672.

(88) Kota, D.; Zhou, H.-X. Macromolecular Regulation of the Material Properties of Biomolecular Condensates. *The Journal of Physical Chemistry Letters* **2022**, *13* (23), 5285-5290.

(89) Alshareedah, I.; Borcherds, W. M.; Cohen, S. R.; Farag, M.; Singh, A.; Bremer, A.; Pappu, R. V.; Mittag, T.; Banerjee, P. R. Sequence-encoded grammars determine material properties and physical aging of protein condensates. *bioRxiv* **2023**, 2023.04.06.535902.