Biophysical Journal
Article

Biophysical Society

'RNA modulation of transport properties and stability in phase-separated condensates

Andrés R. Tejedor,[1] Adiran Garaizar,[2] Jorge Ramírez,[1,*] and Jorge R. Espinosa[2,*]
[1]Department of Chemical Engineering, Universidad Politécnica de Madrid, Madrid, Spain and [2]Cavendish Laboratory, Maxwell Centre, Department of Physics, University of Cambridge, Cambridge, United Kingdom

ABSTRACT One of the key mechanisms employed by cells to control their spatiotemporal organization is the formation and dissolution of phase-separated condensates. The balance between condensate assembly and disassembly can be critically regulated by the presence of RNA. In this work, we use a chemically-accurate sequence-dependent coarse-grained model for proteins and RNA to unravel the impact of RNA in modulating the transport properties and stability of biomolecular condensates. We explore the phase behavior of several RNA-binding proteins such as FUS, hnRNPA1, and TDP-43 proteins along with that of their corresponding prion-like domains and RNA recognition motifs from absence to moderately high RNA concentration. By characterizing the phase diagram, key molecular interactions, surface tension, and transport properties of the condensates, we report a dual RNA-induced behavior: on the one hand, RNA enhances phase separation at low concentration as long as the RNA radius of gyration is comparable to that of the proteins, whereas at high concentration, it inhibits the ability of proteins to self-assemble independently of its length. On the other hand, along with the stability modulation, the viscosity of the condensates can be considerably reduced at high RNA concentration as long as the length of the RNA chains is shorter than that of the proteins. Conversely, long RNA strands increase viscosity even at high concentration, but barely modify protein self-diffusion which mainly depends on RNA concentration and on the effect RNA has on droplet density. On the whole, our work rationalizes the different routes by which RNA can regulate phase separation and condensate dynamics, as well as the subsequent aberrant rigidification implicated in the emergence of various neuropathologies and age-related diseases.

SIGNIFICANCE Biomolecular condensates are liquid-like membraneless compartments that contribute to the spatiotemporal organization of the cellular material. Remarkably, the assembly of these condensates can be critically regulated by RNA. Here, by means of a residue-resolution coarse-grained model, we investigate the role of RNA concentration and length in the stability and kinetics of FUS, hnRNPA1, and TDP-43 protein condensates. Besides observing the "so-called" RNA-induced reentrant behavior with concentration, we find that RNA length plays a major role in condensate dynamics. We identify a minimal critical RNA length at which phase separation is enhanced at low RNA concentration while condensate viscosity can be still reduced at moderate RNA concentration. Overall, our simulations provide a detailed characterization of the underlying molecular driving forces behind RNA-binding-protein aggregation.

INTRODUCTION

Liquid-liquid phase separation (LLPS) is one of the key processes employed by cells to control the spatiotemporal organization of their many components (1–6). This phenomenon—displayed by a large variety of biomolecules such as multivalent proteins and nucleic acids (7–11)—is involved in wide-ranging aspects of the cell function such as membraneless compartmentalization (6,12–16), signaling (2,17), genome silencing (18–20), formation of superenhancers (21), helping cells to sense and react to environmental changes (22), and buffering cellular noise (23), among many others (24–27). The spontaneous demixing of the cell components into different coexisting liquid compartments occurs both inside the cytoplasm (e.g., P granules (1) and RNA granules/bodies (28,29)) and in the cell nucleus (e.g., Cajal bodies (30), nucleoli (31), nuclear speckles (32,33), and heterochromatin domains (19,20)) and enables the coordinated control of thousands of simultaneous chemical reactions that are required to maintain biological activity (34). Beyond these diverse functionalities, membraneless organelles have also been observed to exert mechanical

Submitted April 30, 2021, and accepted for publication October 3, 2021.
*Correspondence: jorge.ramirez@upm.es or jr752@cam.ac.uk
Editor: James Shorter.
https://doi.org/10.1016/j.bpj.2021.11.003

forces to induce chromatin reorganization (35,36) or to act as molecular sensors of intracellular and extracellular exchanges (22). Still, novel biological roles, such as the propensity of condensates to buffer protein concentrations against gene expression noise, continue to be discovered (23,37).

The biomolecular building blocks behind LLPS are usually proteins with intrinsically disordered regions (IDRs) or proteins with globular domains connected by flexible linkers that can establish multiple homotypic or heterotypic interactions with cognate biomolecules (e.g., a different IDR, RNA, or DNA) over their interactions with the solvent (9,11). Several DNA and RNA-binding proteins such as FUS (38–40), hnRNPA1 (15,16), TDP-43 (41,42), LAF-1 (43), G3BP1 (44–46), or HP1 (19,20) have been observed to undergo phase separation both in vivo and in vitro. These proteins, besides their intrinsically disordered regions, frequently present additional specific domains with high physicochemical affinity for RNA (termed RNA recognition motifs (RRMs)) (47) or DNA (48). In particular, the intermolecular binding between IDRs and RNA (either via specific RNA-RRM interactions or nonselective electrostatic and π-π interactions) have been found to be critical in regulating LLPS (43,49–54).

In vitro experimental evidence shows how protein aggregation can be enhanced upon addition of RNA at low concentration but inhibited at high concentration (50,55,56). Such reentrant behavior is in agreement with the hypothesis that solid-like aggregates are more readily formed in the cytoplasm than in the cell nucleus, where the abundance of RNA is higher (57). Moreover, besides modulating the stability of the condensates, RNA can affect their kinetic properties. A viscosity reduction of LAF-1 droplets (a key protein in P granule formation) after the addition of short RNA strands has been observed without significantly affecting droplet stability (43). On the contrary, the inclusion of long RNA chains inside the condensates can also notably enhance their viscosity at certain given concentrations (49,58). Such RNA-induced modulation of droplet viscoelasticity (and recently observed by DNA (59)) is crucial in the regulation and dysregulation of the liquid-like behavior of RNA-binding proteins (RBPs) such as FUS (38–40), hnRNPA1 (15,16), TDP-43 (41,42,60), TAF-15 (57,61), and EWSR1, among many others (51,57,61). The resulting rigidification of these condensates can lead to the formation of pathological solid aggregates, which are behind the onset of several neurodegenerative diseases such as amyotrophic lateral sclerosis, frontotemporal dementia, and Alzheimer (15,62–66). Because of that, a huge effort is being devoted to understanding the underlying molecular factors involved in RNA-induced regulation of condensate stability and viscoelasticity (8,12,53,54,67,68).

Recent experimental advances in single-molecule Förster resonance energy transfer have enabled the direct observation of the structural and dynamic protein behavior in diluted conditions (69–71); however, the thermodynamic and kinetic aspects inside the condensates are still hardly accessible (72,73). Notably, particle tracking microrheology techniques have been successfully used to provide data about the mean-square displacement (MSD) of marked beads inside droplets, and then, via that MSD, condensate viscosity has been estimated (43,49,58,74). Nevertheless, other fundamental magnitudes such as the protein mean-square displacement, end-to-end distance relaxation times, protein radius of gyration, and droplet surface tensions are extremely challenging to obtain (53). Moreover, direct measurements of the molecular contacts that promote phase separation are of great relevance, and rarely, this information can be unequivocally extracted (39,61,75). The mutation and/or phosphorylation of specific residues along sequences can help in deciphering which contacts are key in sustaining LLPS (76,77), but a higher level of mechanistic and molecular resolution is still needed.

In that respect, computer simulations emerge as a great tool to enlighten such a blind spot (78–80). The most recent advances in computational science have allowed to carry out impressive calculations mimicking in vivo conditions (81). Atomistic molecular dynamics (MD) simulations have also been successfully proved in characterizing the conformational ensemble of single proteins and protein complexes (80,82,83), pinpointing the link between chemical modifications and the modulation of protein-protein and protein-DNA interactions (84,85) or guiding the development of chemically accurate coarse-grained models for LLPS (86–89). Simultaneously, a huge effort is being devoted to developing different levels of coarse-grained (CG) potentials, including mean field models (90–93), lattice-based simulations (94–97), minimal models (98–102), and sequence-dependent models (86,103,104). By retaining the specific physicochemical features of proteins, DNA, and RNA while averaging out others for computational efficiency, CG models have been widely used to elucidate key factors behind LLPS and their dependency on protein length (105,106), amino acid sequence (86,103,107,108), multivalency (94,109–114), conformational flexibility (115,116), and multicomponent composition (117–120). Nevertheless, further work is needed regarding the role of RNA in LLPS (121). On the one hand, atomistic MD simulations have provided binding free energies of specific protein-RNA complexes but are limited to very few protein replicas (122,123). On the other hand, coarse-grained models have been recently proposed to elucidate the effect of RNA on phase separation of small prion-like domains such as those of FUS (124), protamine (125), and LAF-1 (119). Remarkably, the work by Regy et al. (119) presents a detailed parameterization of a CG model for RNA within the framework of the hydrophobicity scale (HPS) protein force field (86), opening up new possibilities to link the molecular mechanisms of RNA-RBP condensates to their macroscopic phase behavior.

This work aims to narrow down this problem by shedding light on the RNA modulation of transport properties and stability of RBP condensates. By employing a high-resolution CG model for RNA and intrinsically disordered proteins (IDPs) (86,119,126), we explore the phase behavior of different RNA-binding proteins that undergo LLPS such as FUS, hnRNPA1, and TDP-43 as well as their corresponding prion-like domains and RNA recognition motifs in the absence versus presence of poly-uridine (poly-U) RNA. After validating the model against experimental saturation concentration trends of these pure proteins at physiological salt concentration, we characterize how RNA regulates the coexistence line of these condensates as a function of RNA concentration for a constant poly-U length, as well as for different strand lengths at a constant poly-U/protein concentration. Beyond evidencing RNA-induced reentrant phase separation (50,55–57), we find a critical minimal length below which RNA cannot promote LLPS even at low concentration. Moreover, we characterize the transport properties (i.e., protein mobility and viscosity) of the condensates as a function of RNA saturation and length. Although protein diffusion is predominantly controlled by RNA concentration rather than by strand length, the viscosity of the droplets is critically regulated by both factors, being RNA length a key element in LLPS. Taken together, our work provides a framework to rationalize from a molecular and thermodynamic perspective the ubiquitous dual effect of RNA in the stability and kinetics of RNA-RBP condensates.

MATERIALS AND METHODS

A detailed explanation of our model and methods can be found in the Supporting materials and methods.

RESULTS

Sequence-dependent model validation

Biomolecular condensates are stabilized by chemically diverse weak protein-protein interactions, which are determined by the specific nature (e.g., hydrophobicity, aromaticity, and charge) of the encoded protein amino acids (61,111). Here, to capture such sequence specificity, we employ a novel, reparameterization (126) of the high-resolution HPS model from the Mittal group (103) that accounts for sequence-dependent hydrophobic and cation-π interactions by means of short-range pairwise potentials and for electrostatic interactions through Yukawa long-range potentials (see Supporting materials and methods, Section SI). Bonded interactions between subsequent amino acids are restrained by a harmonic potential (Eq. S2), and nonbonded hydrophobic interactions are modeled via an Ashbaugh-Hatch potential (Eq. S4). Additionally, cation-π and electrostatic interactions are described by Lennard-Jones (Eq. S5) and Yukawa/Debye-Hückel potential terms (Eq. S3), respectively. The low salt physiological concentration regime ($\sim$150 mM) of the implicit solvent is controlled by the screening length of the Yukawa/Debye-Hückel potential. Given that the original HPS model (103) has been shown to underestimate LLPS-stabilizing cation-π interactions (85), we employ the recently proposed reparameterization by Das et al. (126). Additionally, to account for the "buried" amino acids contained in the protein globular domains, we scale down those interactions with respect to the original set of HPS parameters by 30% as proposed in (85,127). All the details regarding the model parameters and simulation setups are provided in the Supporting materials and methods.

To validate the model (103,126), we evaluate the relative ability to phase separate of several archetypal RNA- and DNA-binding proteins that are known to undergo LLPS both in vivo and in vitro. These proteins are fused in sarcoma (FUS) (38–40), heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) (15,16), and the TAR DNA-binding protein 43 (TDP-43) (41,42) (Fig. 1 *A*). We evaluate the phase diagram for the full protein sequences as well as for some of their specific protein domains such as the RRMs or the prion-like domains (PLDs). More precisely, we focus on the following sequences: FUS (full sequence), FUS-PLD, hnRNPA1 (isoform A1-B, hereafter named as hnRNPA1), hnRNPA1-PLD, hnRNPA1-RRM, TDP-43 (full sequence), and TDP-43-PLD (sequences are provided in the Supporting materials and methods). For TDP-43, we also distinguish between two different variants, one including the α-helix structured domain in the C-tail intrinsically disordered region (h-TDP-43) and another in which the whole PLD region remains fully disordered (wild-type wt-TDP-43). Despite h-TDP-43 and wt-TDP-43 only differing by less than 10% of their sequence structural conformation (133,134), the presence of the α-helical structured domain has been shown to moderately affect the protein's ability to phase separate (135). We also study the low-complexity domain (LCD) of the isoform A1-A of hnRNPA1 (15) (termed as hnRNPA1-A-LCD), as it has been shown to be a key part of the hnRNPA1 sequence in promoting LLPS in absence of RNA (15).

By means of direct coexistence (DC) simulations (136–138) in combination with the laws of rectilinear diameters and critical exponents (128), we compute the phase diagram (Fig. 1 *B*) of all the aforementioned proteins (hnRNPA1-PLD and hnRNPA1-RRM are shown in Fig. S3; see Supporting materials and methods, Section SIII, or (101) for details on how to extract coexisting densities from DC simulations). The predicted renormalized critical points (T/T^*_c, where T^*_c refers to the highest critical temperature of the set) against the experimental saturation concentration of the proteins to undergo LLPS for FUS (61,85), FUS-PLD (61), hnRNPA1 (15), hnRNPA1-A-LCD (15), wt-TDP-43 (42,85), h-TDP-43 (42,85), and TDP-43-PLD (132) are

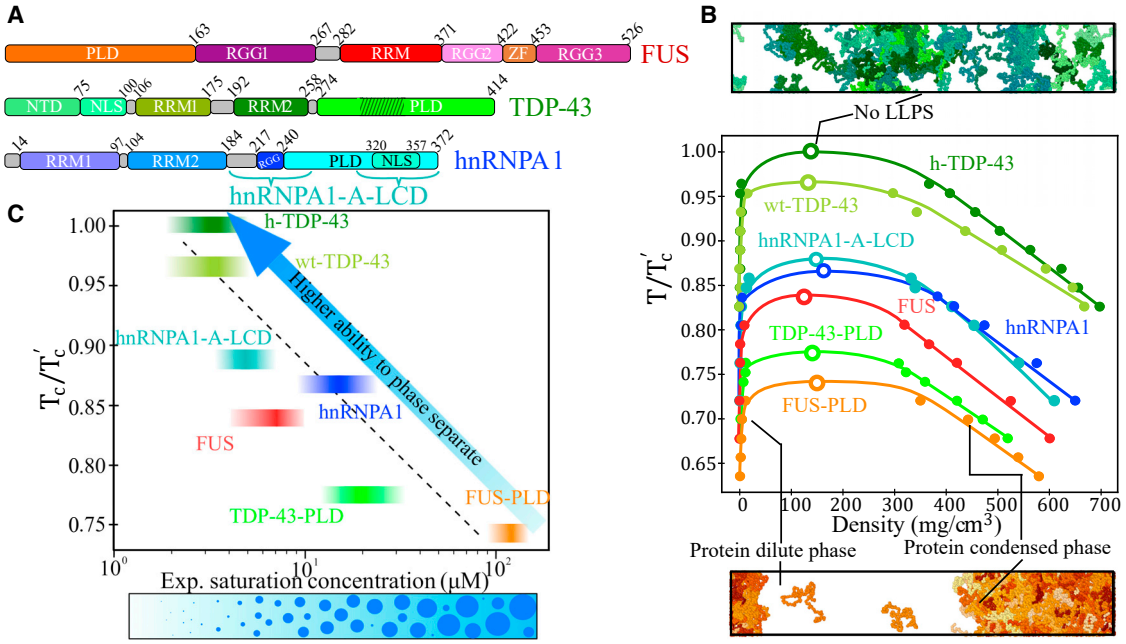


FIGURE 1 Experimental validation of the sequence-dependent coarse-grained model. (*A*) Different sequence domains of the three studied proteins, FUS, TDP-43, and hnRNPA1: prion-like domain (PLD), arginine-glycine-rich regions (RGG), RNA recognition motifs (RRMs), zinc finger (ZF), N-tail domain (NTD), and nuclear localization sequence (NLS). Dashed lines in TDP-43 PLD indicate the position of the α-helical domain. Braces in hnRNPA1 sequence indicate the two LCD segments comprising the sequence of the isoform A1-A of hnRNPA1 (hnRNPA1-A-LCD), corresponding to the residues 186–251 and 304–372 (see Supporting materials and methods, Section SII for the sequences). (*B*) Phase diagram in the (T/T^*_c)-density plane for FUS (*red*), FUS-PLD (*orange*), hnRNPA1-A-LCD (*turquoise*), hnRNPA1 (*blue*), h-TDP-43 (*dark green*), wt-TDP-43 (*lime green*) and TDP-43-PLD (*light green*). Solid circles indicate the coexistence densities obtained through DC simulations and open circles the estimated critical points via the critical exponent and rectilinear diameter laws (128) (Eqs. S6 and S7). T^*_c accounts for the highest critical temperature of the protein set (h-TDP-43), which is $T^*_c = 472$ K. Top: DC simulation of h-TDP-43 above the critical point at which LLPS is no longer observed. Bottom: direct coexistence simulation of FUS-PLD at $T/T^*_c = 0.70$ exhibiting two coexisting phases. In both snapshots, different protein replicas are represented by distinct color tones. (*C*) Experimental saturation concentration of the proteins to undergo phase separation versus the renormalized critical temperatures shown in (*B*). The experimental saturation concentration (129) at physiological salt conditions for FUS (50,61,85) (including FUS-PLD), hnRNPA1 (15), hnRNPA1-A-LCD (130,131), TDP-43 (42,85), and TDP-43-PLD (132) are depicted by intervals to consider concentration uncertainty. The height of the intervals accounts for the computational uncertainty in T/T^*_c. The dashed black line is a linear fit to the displayed data (considering the mean concentration and critical temperature of the interval), and the blue arrow indicates higher ability to phase separate. At the bottom, a schematic cartoon summarizing the expected phase behavior while increasing protein concentration is included. Note that temperatures in this model are unrealistic and only describe the relative ability of the different proteins to phase separate; thus, temperature is only meaningful when is renormalized. To see this figure in color, go online.

plotted in Fig. 1 *C* (please note that the experimental saturation concentration reported by Molliex et al. (15) corresponds to the isoform A1-A, but the difference in the critical concentration between the two isoforms is assumed to be minor). We find a positive correlation between the predicted critical point in our simulations and the experimental protein saturation concentration at physiological salt concentration. The uncertainties in the determination of both the critical temperature and the experimental saturation concentration in Fig. 1 *C* are depicted by the height and the width of the colored bands, respectively. Such impressive qualitative agreement (coarse-grained models with implicit solvent are not expected in principle to quantitatively capture the actual T_c) demonstrates that the cation-π reparameterization proposed by Das et al. (126) on top of the Mittal group's model (103) is able to describe the relative ability of these proteins to self-assemble into phase-separated condensates with better agreement than the original HPS model (103) (Fig. S2). Furthermore, we observe a non-negligible difference between the phase diagram of the α-helical-structured TDP-43 and that of the wt-TDP-43, with the latter showing a moderately lower critical temperature, as reported in (135). Notably, both prion-like domains of FUS and TDP-43 exhibit a significant lower ability to phase separate than their full counterparts as experimentally found (61). On the contrary, hnRNPA1-A-LCD exhibits a similar critical temperature as that of the hnRNPA1 full sequence (15). To rationalize these observations, in the following section we perform a detailed molecular and structural characterization of the condensates.

Structural and interfacial properties of the condensates without RNA

The specific composition and patterning of the amino acids along the sequence has a huge impact on the protein

macroscopic phase behavior (86,103,134). Moreover, beyond sequence, the protein conformational ensemble plays a crucial role not only in their ability to phase separate (116) but also in the condensate structure (134,135,139,140). A close example of this is TDP-43, in which a subtle conformational difference on its C-terminal intrinsically disordered domain produces a moderate change on its phase diagram (Fig. 1 *B*). To further characterize the molecular, structural, and interfacial properties of the previous protein condensates, we now perform a comprehensive full analysis of their surface tension, LLPS-stabilizing most frequent contacts, protein conformational ensembles in and out of the droplet, and condensate structure.

In Fig. 2 *A*, we plot the surface tension (γ) between the condensate (protein-rich) and protein-poor liquid phases as a function of temperature (renormalized by the highest critical temperature of the protein set, T^*_c, of h-TDP-43). An advantage of computer simulations is that γ between two coexisting fluid phases (or between a fluid and a vapor one) can be easily computed, as explained in the Supporting materials and methods, Section SIV, compared with more challenging approaches (i.e., based on the tie-line width of the phase diagrams) as required in experimental setups (53,141). We find that the conformational difference in the 40-residue helical region of the TDP-43-PLD terminal domain has significant consequences on the droplet surface tension of TDP-43. For the whole range of studied temperatures, wt-TDP-43 shows smaller γ than h-TDP-43. At the same temperature, the presence of the helical structure in h-TDP-43 promotes a more compact assembly of proteins in the condensed phase, increasing the surface tension. Additionally, TDP-43-PLD droplets present much smaller γ than those of any of its two full-sequence variants at moderate temperatures, explaining why TDP-43-PLD domains are markedly exposed toward the interface in wt-TDP-43 condensates (Fig. 2 *B*). Similarly, the surface tension of FUS-PLD droplets is lower than that of FUS (full sequence). However, interestingly, γ for hnRNPA1 and hnRNPA1-A-LCD droplets are remarkably similar (as their phase diagrams show; see Figs. 1 *B* and S3), confirming the significant importance of the hnRNPA1-A-LCD sequence in contributing to phase separation (15). Our results clearly evidence a direct correlation between droplet surface tension and condensate stability except for wt-TDP-43 and h-TDP-43 condensates, for which their characteristic heterogeneous arrangement contributes to decrease γ (Fig. S5 *B*). Proteins with higher γ can typically phase separate until higher temperatures or at lower protein concentration.

Next, we focus on the structural organization of the different protein condensates. A significant contrasting behavior between both FUS and hnRNPA1 droplets and those of TDP-43 (both variants) is observed. Although both FUS and hnRNPA1 exhibit homogeneous density droplet distribution with their PLDs indistinctly located along the condensate (although more clustered in hnRNPA1 condensates), TDP-43 shows a highly marked heterogeneous distribution exposing its prion-like domains toward the droplet boundaries (Fig. 2 *B*), evidencing that their PLD interactions barely favor aggregation (134,135). This condensate arrangement allows the minimization of the droplet surface tension and the simultaneous maximization of its enthalpic gain (in absolute value) through a higher density of LLPS-stabilizing contacts at the droplet core (142). In the case of wt-TDP-43, such structural heterogeneity is so pronounced that condensates split into smaller nearly interacting liquid droplets, as shown in Fig. 2 *B* (*center*). Conversely, the α-helix structure of h-TDP-43 notably favors the interaction between helical domains and hence between the rest of the intrinsically disordered neighbor regions, significantly enhancing the PLD connectivity and thus reducing droplet heterogeneity as experimentally suggested (134). Moreover, our simulations show that the structured α-helical domain considerably reduces the local density fluctuations of the droplet and further stabilizes the condensate (Fig. 1 *B*).

To rationalize the molecular driving forces behind these structural differences, we compute 1) the amino acid contact map frequency of the proteins within the condensates (Figs. 2 *C* and S8) and 2) the most persistent residue-residue pair interactions along the aggregated proteins (Figs. S9–S11). We develop a smart cutoff analysis of each specific residue-residue interaction (adapted to the range of the HPS potential ((103); see Supporting materials and methods, Section SVI for further details) to elucidate the key molecular interactions promoting LLPS according to our model (103,126).

In FUS condensates, the most repeated contacts are G-G, R-Y, and G-Y (Fig. S9 *A*), highlighting how hydrophobic, cation-π, and more modestly electrostatic interactions contribute to stabilize the droplets. Because glycine (G) represents nearly 30% of the residues along FUS sequence, the frequency of G-G is the highest despite not being one of the strongest pair of amino acid interactions (85). However, when normalizing the computed number of contacts by the amino acid abundance, we find that the cation-π interaction R-Y becomes the most relevant one inducing LLPS (76,77) according to this force field (see Fig. S9 *B*). Furthermore, when analyzing the FUS contact map (Fig. 2 *C*), we observe that its prion-like domain, despite showing a much lower ability to phase separate on its own than the full protein, markedly interacts with the three RGG domains. The top contacts of the PLD alone are very different from those of the full-sequence FUS (Fig. S9 *A*), resulting in much worse phase-separation capabilities for the PLD than for the full-FUS sequence (Fig. 1 *B*) as experimentally observed (50,61). We also find moderate LLPS-stabilizing interactions among different RNA recognition motifs in FUS (Fig. 2 *C*).

Although in FUS condensates, the PLD plays a vital role in LLPS (39,103), the aggregation of TDP-43 (wild-type) is

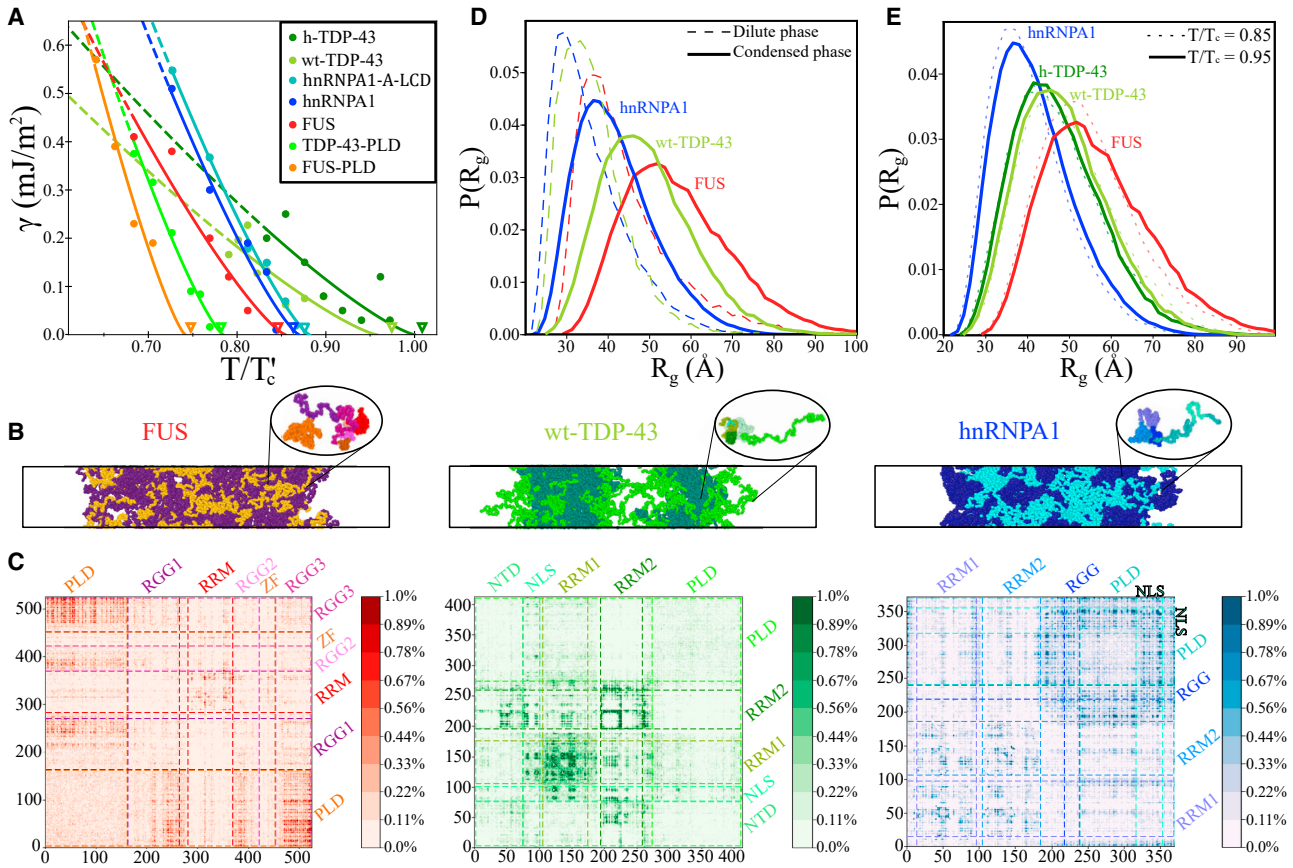


FIGURE 2 Molecular, structural, and interfacial properties of different RNA-binding protein condensates in absence of RNA. (*A*) Condensate surface tension (γ) of FUS, FUS-PLD, hnRNPA1, hnRNPA1-A-LCD, wt-TDP-43, h-TDP-43, and TDP-43-PLD as a function of temperature (renormalized by the highest critical temperature of the protein set, $T'_c = 472$ K for h-TPD-43). Solid circles indicate the obtained γ from DC simulations (see Supporting materials and methods, Section SIV for further details on the calculations) and solid curves the $\gamma \propto (T_c - T)^{1.26}$ fit to our data (128) (*dashed curves* depict the predicted surface tension at low T extrapolated from the fit). Open down triangles represent the obtained (renormalized) critical temperatures of each sequence using the law of rectilinear diameters and critical exponents as in Fig. 1 *B*. (*B*) Snapshots of direct coexistence simulations of the three full sequences at $T = 0.9T_c$ (T_c meaning the critical temperature of each protein): FUS (*left*, $T = 360$ K), wt-TDP-43 (*center*, $T = 410$ K), and hnRNPA1 (*right*, $T = 390$ K). FUS, wt-TDP-43, and hnRNPA1 prion-like domains are highlighted in orange, bright green, and cyan, respectively, and the rest of their sequences in purple, dark green, and dark blue, respectively. The structure of the condensates clearly shows the contrast between homogeneously distributed PLD domains as in FUS, clustered PLD domains as in hnRNPA1, and interfacially exposed PLD domains as in wt-TDP-43 condensates. (*C*) Frequency amino acid contact maps of FUS (*left*), wt-TDP-43 (*center*), and hnRNPA1 (*right*) droplets at $T = 0.9T_c$. Scale bars indicate the averaged percentage of amino acid contact pairs per protein (see Supporting materials and methods, Section SVI for further details on these calculations). Dashed lines depict the limits of the different protein domains as indicated in Fig. 1 *A*. (*D*) Protein radius of gyration distribution function of the three sequences at $T/T_c = 0.9$ and at the bulk equilibrium coexisting density of the diluted (*dashed curves*) and the condensed phase (*continuous curves*). (*E*) Protein radius of gyration distribution function within the condensates at moderate ($T/T_c = 0.85$) and high temperature ($T/T_c = 0.95$). To see this figure in color, go online.

mainly sustained by contacts between RRMs, either with themselves or with other protein regions such as the N-tail domain or the nuclear localization sequence, but mostly dominated by RRM1-RRM1 and RRM2-RRM2 interactions. (Fig. 2 *C*). Nonetheless, the wt-TDP-43 PLD region is still the second protein domain establishing more contacts in total after the RRM1 segment, but mostly because of its length. The three most predominant contacts in wt-TDP-43 (according to our model (103,126)) are K-F, K-E, and K-D (Fig. S10 *A*), clearly denoting the key role of cation-π and electrostatic interactions in driving condensation. However, when the structured helical region is present (h-TDP-43), R-F contacts sensibly increase, becoming the third most dominant interaction. Interestingly, the renormalization of contacts by amino acid abundance in TDP-43 barely modifies the list of the most frequent interactions, probably because of the very homogeneous distribution of amino acids along its sequence (Fig. S10 *C*) when compared with that of FUS. However, similarly to FUS, TDP-43-PLD shows a completely different list of the most repeated interactions compared with the full protein (Fig. S8 *A*), which is likely contributing to reducing its critical temperature (Fig. 1 *B*) and surface tension (Fig. 2 *A*).

In hnRNPA1, the most frequent contacts are G-G, G-S, and G-R (Fig. S11 *A*), but because glycine is the most abundant amino acid (~25%), followed by serine (~15%), the

normalized contacts by amino acid abundance show that R-Y, R-F, and K-Y are dominant interactions, again highlighting the importance of cation-π interactions in hnRNPA1 LLPS. The list of top interactions of hnRNPA1-PLD, even after normalization, is very similar to that of hnRNPA1 (Fig. S11, *A* and *B*), which explains why the phase diagrams of both sequences are hardly distinguishable (Fig. S3 *A*). Surprisingly, the list of the most frequent interactions of hnRNPA1-A-LCD is also remarkably similar to that of the hnRNPA1 full sequence (Fig. S11 *A*). In fact, the detailed contact map of hnRNPA1-A-LCD corresponds to the region of hnRNPA1 that contains more LLPS-stabilizing interactions (*dashed lines* in Fig. S8). Thus, the ability of hnRNPA1 to phase separate alone can be mainly captured by these protein interactions in hnRNPA1-A-LCD (see Fig. S3 *A*).

Finally, we investigate the protein conformational ensemble within the condensates and the diluted phase by computing the radius of gyration distribution function of the proteins $P(R_g)$. Our simulations reveal that in all cases, when proteins transition from the diluted to the condensed phase, their conformations adopt larger radii of gyration (Fig. 2 *D*). Also, the width of $P(R_g)$ considerably increases, indicating the more versatile conformations that proteins can exhibit within the condensate. This structural behavior allows proteins to maximize their number of intermolecular contacts and thus the droplet connectivity, as recently shown in (116). Phase-separation-driven expansion of proteins undergoing homotypic LLPS has been observed for Tau-IDP (143) using steady-state fluorescence measurements of pyrene and fluorescein-labeled Tau-K18 proteins, a protein associated with Alzheimer disease (62). Even if modest, phase-separation-induced expansion enables IDRs to establish a surplus of enthalpy-maximizing (more energetically favorable) interprotein contacts in the condensed phase compared to those that they would adopt if they remained unchanged or underwent collapse. On the other hand, very recently, NMR and electron paramagnetic resonance (EPR) spectroscopies have shown that the N-terminal domain of FUS is compacted when entering in the condensed phase under agarose hydrogel conditions (144). However, because of the employed different experimental matrix composition, our model predictions cannot be directly related to these striking observations. Now, when regarding the protein conformational ensemble within the condensates along temperature, we note a mild change in the hnRNPA1, FUS, and TDP-43 conformations as we approach the critical T (Fig. 2 *E*), in contrast to those measured in the diluted phase as a function of T (Fig. S6 and (86)). Moreover, when comparing both TDP-43 variant $P(R_g)$ distributions, we find almost identical protein ensembles, exhibiting the wild-type variant slightly more open conformations. Such a small surplus of extended conformations shown by wt-TDP-43, which can enable a higher number of intermolecular contacts (116), is not enough to enhance LLPS as through the α-α helical interactions present in the h-TDP-43 (134).

RNA-induced reentrant behavior in phase separation

RNA has been recently shown to critically regulate both the phase behavior of different RNA-binding proteins (43,50,52,56,57,145) and, most importantly, the emergence of aberrant liquid-to-solid pathological phase transitions (51,62). In this section, we explore the impact of poly-U RNA in LLPS of RBPs from a molecular and a physicochemical perspective. By means of the novel coarse-grained model of RNA recently proposed by Regy et al. (119) and direct coexistence simulations (136–138), we characterize the condensate stability of different RNA-binding proteins (and domains) from low to moderately high poly-U concentration regimes. We choose poly-U RNA for simplicity (85), and to follow previous landmark works on RNA-RBP phase separation (43,56).

First, we mix poly-U RNA strands of 250 nucleotides (nt) with the proteins studied above. Remarkably, not all proteins were able to favorably interact with poly-U in our simulations. We find that FUS-PLD and TDP-43 (including both variants) do not associate with poly-U even at very low RNA concentration (i.e., $\sim$0.05 mg poly-U/mg protein). We further test the affinity of wt-TDP-43 with poly-U strands by performing a separate analysis of each of its major protein sequence domains (PLD, RRM1, and RRM2). None of these domains exhibited a conclusive interaction with poly-U at temperatures moderately below the critical one. That is not entirely surprising because 1) several experimental studies have shown that TDP-43-RRM1 only presents a strong affinity for RNA strands enriched in UG nucleotides (146–148) and 2) TDP-43-RNA heterotypic interactions are mainly driven by the RRM1, whereas the RRM2 plays a supporting role (146). Furthermore, in the employed model, the interactions between poly-U and TDP-43 are mainly electrostatic, and therefore, other factors such as RNA secondary and tertiary structures that might sensibly promote RRM binding to specific RNA sequences are not explicitly considered (149). On the contrary, the noninteracting behavior between FUS-PLD and poly-U strands was completely expected because the FUS-PLD sequence does not present either RNA-binding domains or positively charged domains, thus, precluding their association.

We now evaluate the phase diagram of all proteins (or protein domains) that favorably interact with poly-U; these are FUS, hnRNPA1, hnRNPA1-PLD, hnRNPA1-A-LCD, and hnRNPA1-RRMs. In all these systems except for hnRNPA1-PLD, the resulting phase behavior is similar to that shown in Fig. 3, *A* and *B* for FUS (note that poly-U/hnRNPA1-PLD condensates show a very mild LLPS enhancement at low poly-U concentration (Fig. S4 and Table S3), so hereafter, the results are just discussed

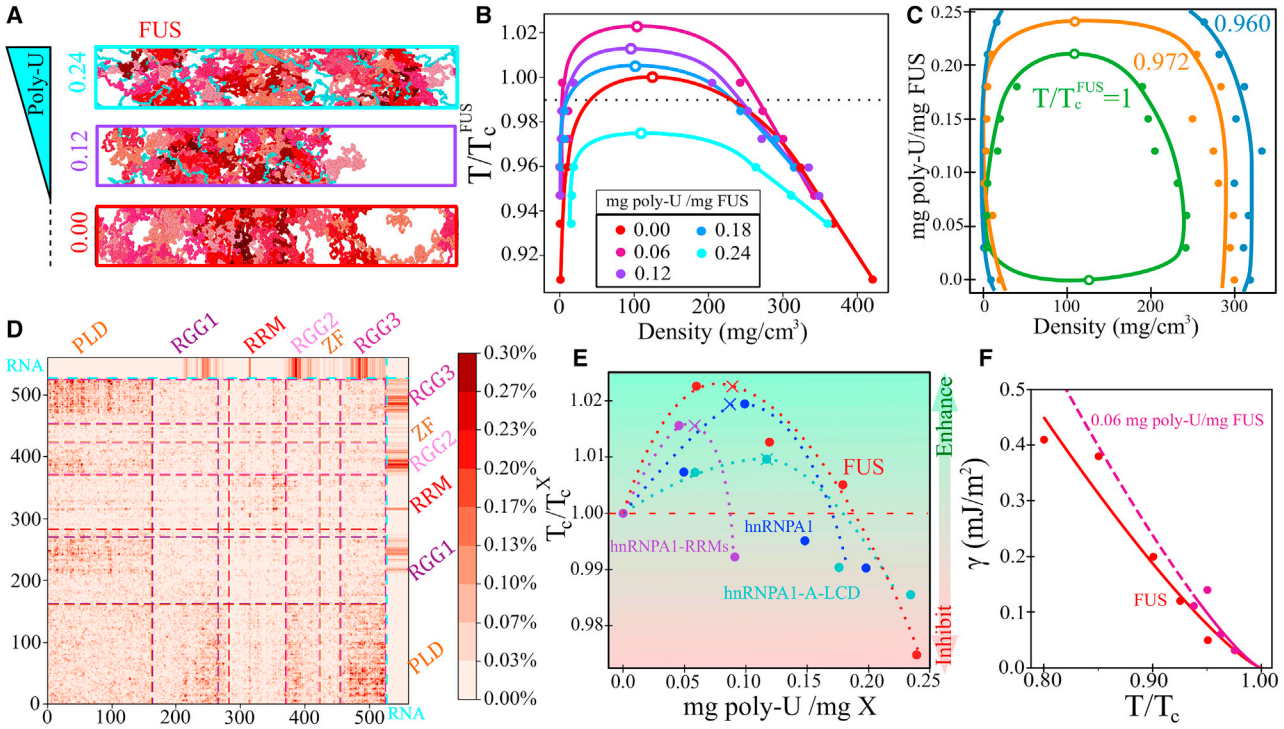


FIGURE 3 RNA-induced reentrant behavior in RBP phase separation. (*A*) Snapshots of direct coexistence simulations of FUS (*red*) and poly-U (*cyan*) at temperature ($T/T_c^{FUS} \approx 0.99$, where T_c^{FUS} refers to the critical temperature of FUS in absence of poly-U) with increasing poly-U/FUS mass ratios as indicated at the left side of the simulation boxes. (*B*) Phase diagrams in the temperature-density plane for five different poly-U/FUS mass ratios as indicated in the legend. Open circles represent the estimated critical point and solid circles the obtained coexisting densities from DC simulations. The horizontal dotted line depicts the temperature at which the DC snapshots shown in (*A*) were taken. (*C*) Phase diagram in the (poly-U/FUS mass ratio)-density plane for three different temperatures: $T/T_c^{FUS} = 1$, green; $T/T_c^{FUS} = 0.972$, orange; and $T/T_c^{FUS} = 0.960$, blue. Open circles depict the estimated critical points and solid circles the obtained coexisting densities from DC simulations. (*D*) Map of intermolecular contacts per protein replica for FUS in presence of poly-U (250) at $T/T_c = 0.95$ and at the coexisting droplet equilibrium density at that temperature. The mass between poly-U (250 nt)/FUS was 0.119. The intermolecular contacts between poly-U RNA and FUS are included in the upper and right-side edges of the map. Distinct domains of FUS have been labeled as in Fig. 2 *C* (*left*). (*E*) Reentrant behavior of several RNA-binding proteins and domains as a function of the poly-U/protein mass fraction. Solid circles depict the critical temperature (renormalized by that of each pure protein in absence of poly-U) of the different protein mixtures. Cross symbols indicate the poly-U/protein mass fraction at which mixtures possess neutral electrostatic charge, and the horizontal red dashed line shows the limit above which phase separation is enhanced by poly-U (T_c^X refers to the critical temperature of each pure protein/domain). (*F*) FUS droplet surface tension (γ) as a function of temperature (renormalized by T_c of each system) with (*purple*) and without (*red*) poly-U as indicated in the legend. Solid circles account for the obtained γ values from DC simulations, whereas solid lines account for the fit given by the following expression (128) $\gamma \propto (T - T_c)^{1.26}$, which can be conveniently extrapolated to moderate lower temperatures (*dashed curve*). To see this figure in color, go online.

for hnRNPA1-A-LCD, FUS, hnRNPA1, and hnRNPA1-RRMs). At low poly-U/protein ratios, the stability of the condensates moderately increases (~2% higher critical temperature), whereas at high concentration, the critical point decreases below the critical temperature without RNA (Fig. 3 *E*). This reentrant behavior has been experimentally observed for synthetic peptides such as RP_3 and SR_8 in poly-U mixtures (56) and for RNA-binding proteins such as FUS (50,55–57), Whi3 (58), G3BP1 (44), and LAF-1 (43). In fact, for FUS and hnRNPA1 proteins, it has been reported that at RNA/protein mass ratios close to 0.9, phase separation can be inhibited (57), which is in qualitative agreement with our observations (~0.3 mg RNA/mg protein). The higher RNA reentrant concentration measured in vitro may come from the fact that it refers to the total solution concentration rather than within the phase-separated condensates, as in our simulations, which is very likely to be lower than in the diluted phase. From our simulations, we also note that although a 2% shift in the critical temperature might seem insignificant, the actual increment in temperature according to the force field (103,119,126) may be as large as 10 K, which represents a huge temperature rise when referred to the physiological cell environment. We also plot the phase diagram for FUS with poly-U in the RNA/protein mass ratio-density plane for different temperatures close to the pure FUS critical one (Fig. 3 *C*). At the pure FUS critical temperature, we observe a closed-loop diagram (*green curve*), and for slightly lower temperatures, reentrant phase behavior is also recovered in agreement with experimental findings (50,55–57). To microscopically rationalize this behavior, we compute the protein-protein, protein-RNA, RNA-RNA, and total number of contacts as

a function of poly-U concentration (Fig. S14), which clearly shows that at low RNA concentration, the total number of contacts per protein within the condensates is higher (∼30) than at high poly-U concentration (∼20) (just before phase separation is no longer possible). Moreover, a maximum in FUS-poly-U contacts can be seen at moderate concentration (0.17 mg poly-U/mg FUS), whereas RNA-RNA contacts are almost negligible at any concentration. We note that to accurately determine the specific RNA-induced temperature raise, atomistic simulations would be needed (150), although that is far beyond current computational capability. Nevertheless, just the fact that a coarse-grained model successfully captures the experimental reentrant behavior observed in some RBP-RNA condensates is outstanding (119). For the studied proteins, FUS (*red*) exhibits the highest variation in critical temperature at either low or high RNA concentration (Fig. 3 *D*). Interestingly, hnRNPA1 (*blue*) shows an intermediate behavior between that of its A-LCD (*cyan*) and RRM (*purple*) domains. The maximal critical temperature in hnRNPA1-RRM is reached at the lowest RNA concentration of the set, and it sharply decays after the maximum. Contrarily, hnRNPA1-A-LCD suffers only a moderate increment of the critical temperature, but its reentrant behavior is smoother and appears at much greater concentration (two times higher) than that of hnRNPA1-RRM. Overall, hnRNPA1 condensates present higher RNA-induced stabilization in the low RNA regime than that of their PLD and RRMs separately. Moreover, it is worth mentioning that in all sequences, the larger enhancement of LLPS is reached at a poly-U concentration close to the electroneutrality point (depicted by *crosses* in Fig. 3 *D*), which emphasizes the major importance of electrostatic nucleotide-amino acid interactions in RNA-RBPs phase separation (56,119).

To characterize the RNA-RBP condensates from a microscopic perspective, we analyze the key molecular contacts enabling phase separation (Fig. 3 *D*; Figs. S12 and S13). We find that near the optimal poly-U/protein concentration promoting LLPS, the most frequent contacts in poly-U/FUS condensates are now R-U and G-U (Fig. S13). This demonstrates how poly-U (even at low fraction) plays a major role in sustaining the condensates, given that the two most frequent contacts are now shifted from G-G and R-Y to the electrostatic cation-anion R-U interaction; and G-U interactions (Fig. S13). In terms of the FUS sequence domains, the RGG regions and the RNA recognition motif are those presenting more contacts with poly-U strands, explaining why G-U becomes one of the most dominant molecular contacts by proximity (Fig. S12). On the other hand, the PLD region presents the least favorable interaction with poly-U. The fact that poly-U strands are not specifically recognized by the zinc finger domain needs to be further tested to check whether this may be caused by model deficiencies (lack of secondary or tertiary structured driven interactions) and/or due to the fact that poly-U strands are not specifically recognized by zinc finger domains (122,151). We also analyze the protein and RNA conformational ensemble as a function of poly-U concentration by computing the radius of gyration histograms for FUS and poly-U (125 nt) (Fig. S7 *A*). We strikingly find that despite varying the stability and density of the droplets with RNA concentration (Fig. 3, *A*, *B*, *C*, and *E*), the structural conformation of the proteins and RNA does not significantly change (at least by analyzing the R_g). Regarding poly-U/hnRNPA1 droplets, our simulations reveal that G-G contacts remain as the dominant amino acid pair interaction (although it substantially decreases by a factor of two) and R-U and G-U become the next two most frequent contacts (further details in Supporting materials and methods, Section SVI and Fig. S13). However, the behavior of poly-U/hnRNPA1-A-LCD condensates is radically different; despite its phase diagram being altered by poly-U addition, the most frequent contacts remain similar to those in absence of RNA but include a very modest excess contribution in R-U interactions (Fig. S13). On the contrary, when just considering the RRM1-RRM2 hnRNPA1 domains (*purple curve* in Fig. 3 *E*), even at the lowest RNA/protein ratio at which the droplet stability attains its maximal value, R-U and K-U emerge as some of the most frequent contacts despite the very modest poly-U concentration (Fig. S13). Finally, if we examine the contact map between poly-U and different hnRNPA1 (full-sequence) domains, we strikingly observe that the PLD comprises the highest number of interactions with poly-U strands. However, such observation can be explained through the longer length of the PLD with respect to the two RNA recognition motifs. Yet, the strongest electrostatic interactions (mainly R-U and K-U) between hnRNPA1 and poly-U are those held through the two RRM domains (Fig. S12).

We also determine the surface tension (γ) of the condensates with the dilute phase in presence of poly-U as a function of temperature (Fig. 3 *F*). Both for FUS (Fig. 3 *F*) and hnRNPA1-A-LCD (Fig. S5) condensates, we observe that poly-U at low concentration significantly increases the droplet surface tension in addition to further stabilizing the droplets as shown in Fig. 3 *D*. Our simulations suggest that the molecular origin behind such surface tension increase comes from the reallocation of the positively charged residues (R, H, and K) within the bulk condensate to maximize the molecular connectivity with poly-U, rather than remaining more exposed to the interface as in the pure component, and therefore, contributing to minimize the droplet surface tension because of their higher hydrophilicity. On the contrary, at moderately high poly-U ratios, the surface tension seems to decrease, although the scattering of our data does not allow us to conclude whether a non-monotonic behavior in γ may also exist (Fig. S5).

To further elucidate the role of RNA-regulated RBP condensate stability, we now focus on the effect of poly-U length in LLPS. A landmark study by Maharana et al. (57)

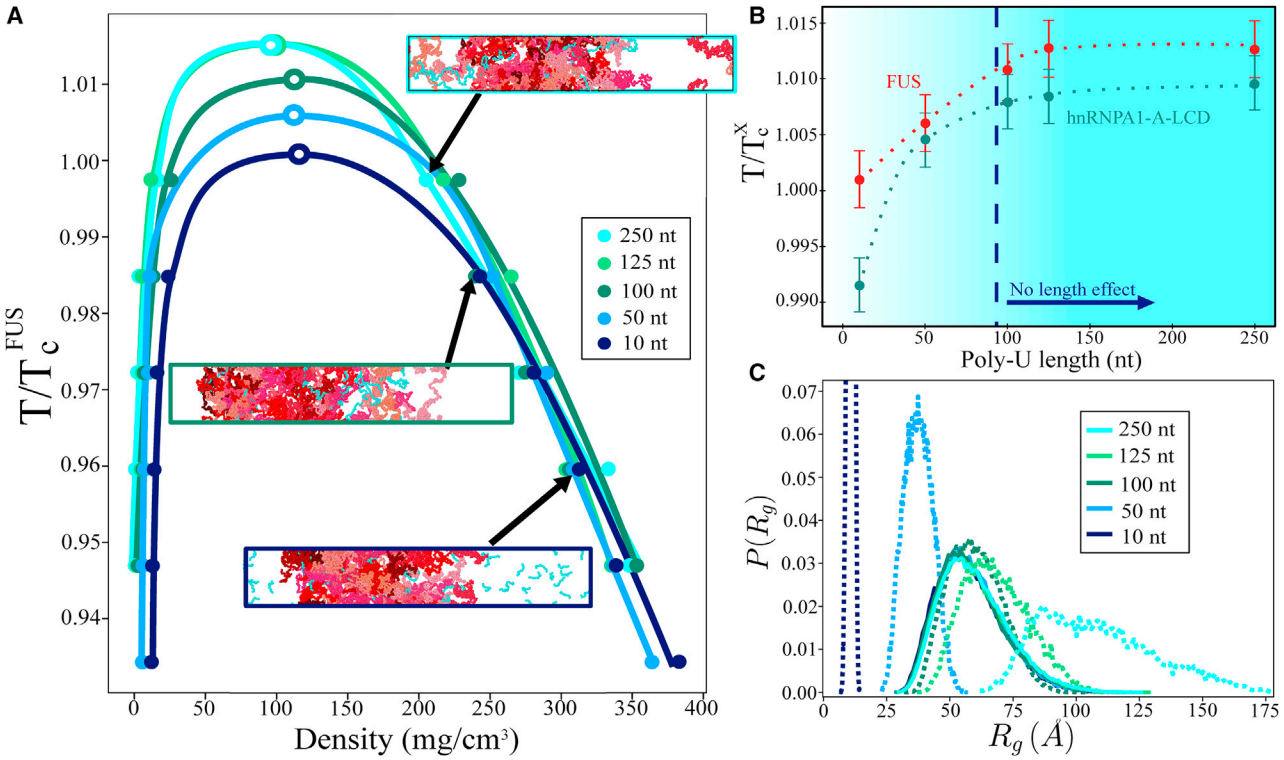


FIGURE 4 Condensate stability dependence on poly-U RNA length. (*A*) Phase diagrams in the temperature-density plane for poly-U/FUS mixtures of different poly-U strand lengths (as indicated in the legend) at a constant concentration of 0.119 mg poly-U/mg FUS. Temperature is normalized by the critical one of FUS (T_c^{FUS}) without poly-U RNA. DC snapshots of three representative cases of poly-U/FUS mixtures at the temperature indicated by the arrow and with poly-U lengths as depicted by the box side color (see legend) are also included. (*B*) Renormalized critical temperature of poly-U/FUS (*red*) and poly-U/hnRNPA1-A-LCD (*green*) condensates as a function of poly-U length for a constant concentration of 0.119 mg poly-U/mg FUS and 0.117 mg poly-U/mg hnRNPA1-A-LCD, respectively. Temperature is normalized by the corresponding critical temperature (T_c^X) of each protein in absence of poly-U. The vertical dashed line indicates the minimal RNA length required to maximize droplet stability at this given concentration. Error bars depict the uncertainty in the estimation of the critical temperature. (*C*) Radius of gyration histograms of FUS (*continuous curves*) and poly-U (*dotted curves*) extracted from the condensed phase of the DC simulations shown in (*A*) for different strand lengths as indicated in the legend. Please note that all histograms have been normalized. To see this figure in color, go online.

showed that smaller RNAs were more potent than larger ones in solubilizing FUS condensates. On the other hand, Zacco et al. (60) found that longer RNA repeats presented weaker dissociation constants with N-RRM1-2 domains of TDP-43 than threefold shorter RNA strands. Given the critical role that RNA performs on the behavior of many different RBP organelles (15,43,58), we investigate the role of RNA length by introducing poly-U strands of different lengths (i.e., 10, 50, 100, 125, and 250 nucleotides) at a constant poly-U/protein mass ratio that maximizes droplet stability (∼0.12 mg RNA/mg protein) for FUS and hnRNPA1-A-LCD sequences (Fig. 3 *E*). Our simulations reveal that very short poly-U strands (∼10 nt) do not enhance phase separation in FUS and hnRNPA1-A-LCD droplets (Fig. 4, *A* and *B*). In fact, 10 nt poly-U strands in hnRNPA1-A-LCD droplets inhibit LLPS even at low concentration. On the other hand, we observe that RNA strands longer than ∼100 uridines (hereafter called minimal critical length) promote a similar droplet stabilization independently of their length (Fig. 4 *B*). To further investigate the molecular insights behind these observations, we analyze the FUS-RNA conformational ensemble within phase-separated droplets with distinct RNA lengths by computing their radius of gyration histograms. As can be seen in Fig. 4 *C*, RNA strands with radii of gyration comparable or longer than those of the proteins (i.e., above the minimal critical length, Fig. 4 *B*) promote maximal condensate stabilization, whereas RNA poly-U strands with shorter R_g than those of FUS proteins (i.e., below the critical length) cannot achieve the same degree of droplet stabilization (and density) for the same RNA/FUS concentration. We note that the observed minimal critical RNA length in FUS and hnRNPA1-A-LCD droplets may be also modulated by some protein- or RNA-specific features and modifications such as RNA sequence, secondary structure interactions, protein charge distribution, post-translational modifications, and RRM patterning effects (43,51). Moreover, if we compute the number of protein contacts within the condensate when adding short and long RNA chains (Fig. S14 *B*), we find that RNA strands longer than the minimal critical length promote a higher number of protein intermolecular interactions, whereas short RNA chains (i.e., 10 nt) considerably

hinder the liquid-network connectivity of the proteins within the droplets (117), hence, RNA behaving as a ligand or client instead of a co-scaffold, as when RNA is longer than 100 nucleotides. An extensive characterization (and rationalization) of the critical aspects controlling RBP-RNA aggregation, such as the RNA length dependence studied here, may provide highly valuable insights for designing therapeutic RNA strategies to combat neurodegenerative disorders whose development is deeply linked to aberrant accumulation and solidification of RBP condensates (63,131).

RNA modulates the transport properties of RBP condensates

Besides controlling condensate stability, RNA has been proved to play a critical role in regulating the dynamics of many membraneless organelles (15,51,57). A seminal study of Zhang et al. (58) showed that the RNA-binding protein Whi3 phase separates into liquid-like droplets whose biophysical properties can be subtly tuned by changing the concentration of the mRNA binding partner, showing that larger RNA content increases Whi3 droplet viscosity. On the other hand, RNA has been also observed to provoke the opposite effect in LAF-1 condensates when short strands (50 nt) were introduced (43). Nonetheless, when long RNAs were used (up to 3000 nt), LAF-1 condensates presented significantly higher viscosity (49). Moreover, beyond length, RNA sequence can be also an important factor in modulating droplet dynamics (152). However, a full understanding of the precise effect of RNA in different RBP condensates still requires further work (54). Here, we aim to provide new, molecular insights on this discussion by measuring via computer simulations the protein diffusion and viscosity of several RBP condensates as a function of poly-U concentration and length.

In vitro, viscosity (η) is usually obtained by bead tracking within droplets using microrheology techniques (29,43,153,154) so that the trajectory can be registered and the MSD of the beads calculated and thus their diffusion coefficient. Then, the droplet viscosity is inferred from the diffusion coefficient by using the Stokes-Einstein relation (155). However, in computer simulations we can measure both observables independently. The linear viscoelasticity of a material can be straightforwardly computed by integrating in time the relaxation modulus $G(t)$ of the system (156,157) (see Supporting materials and methods, Section SVII), whereas the diffusion coefficient can be extracted from the MSD of the proteins. The direct calculation of $G(t)$ provides useful information about the underlying relaxation mechanisms of the proteins (see Fig. 5 *A* for FUS condensates with and without poly-U), either at short times (*white region*) at which the relaxation modes mostly depend on short-range and intramolecular interactions (i.e., internal protein conformational fluctuations such as bond or angle relaxation modes) or at long timescales (*beige region*) at which $G(t)$ is dominated by intermolecular forces, long-range conformational relaxation, and protein diffusion within crowded liquid-like environments. Moreover, in Fig. 5 *A*, the fact that $G(t)$ presents a faster decay when condensates contain RNA (*purple circles*) suggests that their viscosity will be lower than those of pure FUS droplets (*red circles*).

We characterize the condensate dynamics of FUS, hnRNPA1, and hnRNPA1-A-LCD as a function of poly-U concentration at constant temperature (just below the critical T of each protein in absence of poly-U, $T/T_c \sim 0.98$) and at the corresponding bulk droplet equilibrium density corresponding to each poly-U concentration at that temperature. First, we introduce poly-U strands of 125 nucleotides. As shown in Fig. 4 *B*, the phase diagram for a given concentration is not expected to change either by using strands of 125 or 250 nucleotides. For both FUS and hnRNPA1-A-LCD droplets, we observe a mild nonmonotonic behavior with a maximum in viscosity at low poly-U ratios (*solid circles* in Fig. 5 *B*), which might be directly related to the maximum in droplet stability shown in Fig. 3 *D* or due to a coincidental scattering of our measurements. Nevertheless, at moderate poly-U mass ratios (i.e., >0.20 mg poly-U/mg protein), the viscosity of the condensates (using 125 nt RNA strands) is about 30% lower than that without poly-U. On the other hand, a monotonic decreasing trend in viscosity was detected for hnRNPA1 condensates, for which almost a $\sim$50% drop in η is found at high poly-U mass fractions (0.24 mg poly-U/mg hnRNPA1). Even though the observed maximum in viscosity could be easily related to the reentrant behavior depicted in Fig. 3 *E*, further work needs to be devoted to clarifying whether this is a real feature of the model and ultimately of these RBP-RNA condensates. Furthermore, we investigate how poly-U strands of 250 nucleotides can regulate droplet viscosity at the same concentrations. Although poly-U 125 nt strands significantly reduce viscosity at high mass ratios, poly-U 250 nt strands barely varies the condensate viscosity at the same concentrations except for FUS, for which a moderate viscosity increase was detected (*open symbols* in Fig. 5 *B*). These observations are in full agreement with those reported for LAF-1 condensates in presence of short (43) and long (49) RNA strands. Long RNA chains, even at low to moderate concentrations, can increase droplet viscosity because of their own slow relaxation times. In fact, when very short RNA strands of 10 nt are added in FUS condensates (*red cross symbol*), the viscosity of the phase-separated droplets is almost two times lower than that of condensates containing 250 nt at the same poly-U/FUS ratio ($\sim$0.06 mg poly-U/mg FUS). In the Supporting materials and methods, Section SVII and Tables S4 and S5, we provide the values of η for the different RBP condensates as a function of poly-U concentration and length, as well as details on the statistical analysis for estimating the uncertainty

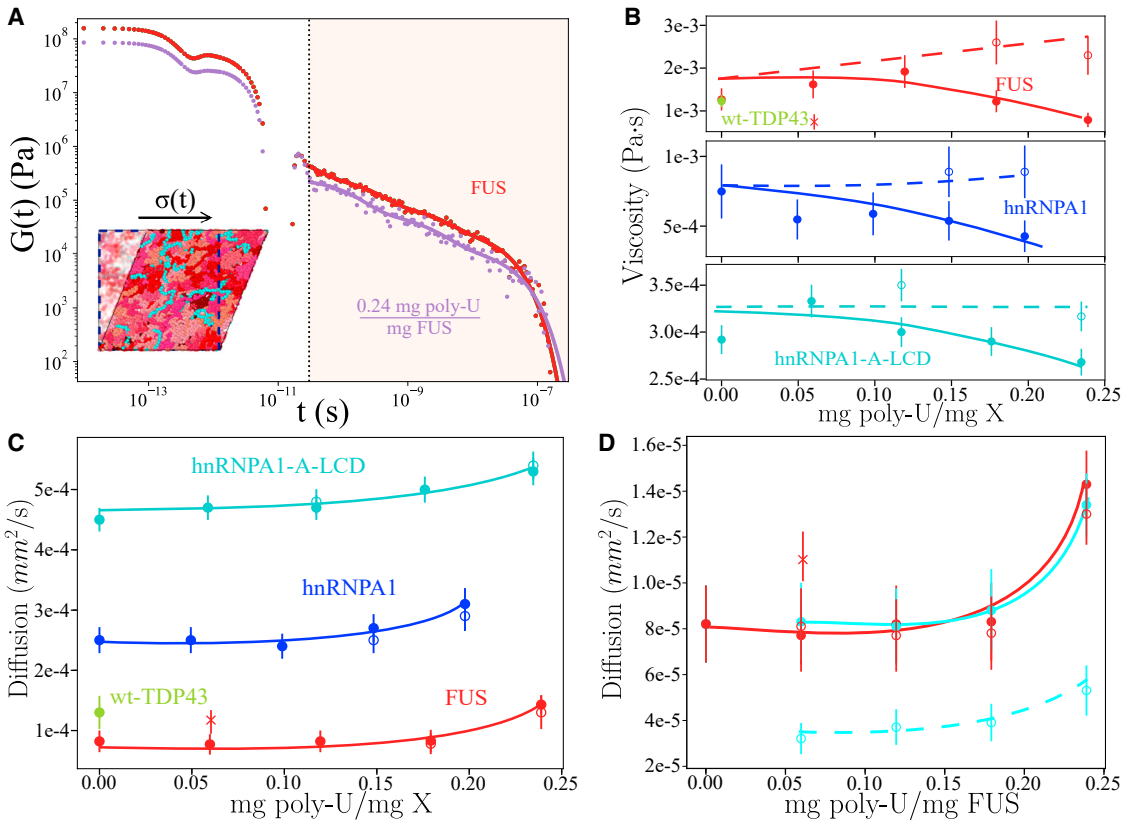


FIGURE 5 RNA critically regulates the dynamical properties of RBP condensates. (*A*) Shear stress relaxation modulus of FUS condensates in the absence (*red*) versus presence (*purple*) of poly-U strands of 125 nucleotides at 0.24 mg poly-U/mg FUS mass fraction, $T/T_c = 0.97$ (where T_c refers to the critical temperature of FUS pure condensates) and the corresponding equilibrium bulk density of each droplet at such conditions. The vertical dotted line separates the fast-decay relaxation mode regime (*white*) and the slow-decay relaxation mode one (*beige*). A snapshot illustrating a shear stress relaxation experiment over a poly-U/FUS condensate simulation box is included. (*B*) Viscosity of FUS (at $T/T_c = 0.97$), hnRNPA1 (at $T/T_c = 0.985$), and hnRNPA1-A-LCD (at $T/T_c = 0.98$) condensates as a function of the poly-U/protein mass ratio. An estimate of wt-TDP-43 viscosity in absence of poly-U at $T/T_c = 0.97$ is also included (*green circle*). Solid circles depict viscosities when poly-U strands of 125 nt were used and open circles when strands 250 nucleotides were added. The cross symbol in the FUS panel indicates the viscosity of FUS-poly-U condensates when strands of 10 nt were included. Continuous and dashed lines are plotted as a guide for the eye for strands of 125 and 250 nt, respectively. The error bars in viscosity have been estimated using the fit of G(t) to Maxwell modes (as described in Section SVII of the Supporting material). Note that T_c refers to the pure component critical temperature of each protein. (*C*) Protein diffusion coefficient (*solid circles*) as a function of poly-U (125 nt)/protein (X) mass ratio. Open circles account for the protein diffusion coefficient when poly-U strands of 250 nt were added, and solid ones for 125 nucleotide RNA strands. The same system conditions described in (*B*) are applied on these calculations. Continuous curves are included as a guide to the eye. (*D*) Diffusion coefficient of FUS (*red*) and poly-U strands (*cyan*) as a function of the poly-U/FUS mass ratio. Open circles show the diffusion coefficient of FUS (*red*) and poly-U strands (*cyan*) when RNA strands of 250 nt were added, and solid circles correspond to values with poly-U strands of 125 nt. The red cross symbols indicate the diffusion of FUS proteins in condensates with poly-U chains of 10 nucleotides. The diffusion of poly-U chains of 10 nt is $D \sim 3 \times 10^{-3}$ mm^2/s, so it has been omitted from the panel. To see this figure in color, go online.

of these calculations. We note that our $G(t)$-values for FUS do not quantitatively match with those from experiments of (74). That is somewhat expected because our coarse-grained model has been parameterized to describe the radius of gyration (103) and most frequent molecular contacts (126) between proteins rather than dynamic properties such as transport properties within the condensates. Nevertheless, the observed behavior with RNA and between the different RBP condensates is expected to qualitatively hold despite the different model approximations (i.e., implicit solvent and amino acids or nucleotides represented by spherical beads).

Finally, we measure the protein diffusion coefficient (D) within the condensates for all previous poly-U concentrations and strand lengths of 125 and 250 nt. In all cases, we find a coherent correlation between viscosity and protein mobility, the latter being considerably higher at moderate poly-U/protein ratios than for pure protein condensates (Fig. 5 *C*). Strikingly, protein diffusion hardly depends on poly-U strand length (*open symbols*) as viscosity does (Fig. 5 *B*). Only when extremely short RNA chains of 10 nt are added, as those tested in FUS condensates (Fig. 5 *D*, *cross symbol*), protein diffusion noticeably increases. Although the shear stress relaxation modulus of the condensates crucially depends on the RNA strand length (longer RNAs imply longer $G(t)$ relaxation decay), the protein diffusion coefficient does not. The latter mainly depends on droplet density (and temperature), and, as shown in Fig. 4 *A*, condensate densities

remain similar when using strands of either 125 or 250 nucleotides. However, when 10 nt strands are added at the same poly-U concentration, the droplet density critically decreases from ~0.28 g/cm^3 (for 125 and 250 nt chains) to ~0.20 g/cm^3. Therefore, our simulations suggest that the condensate dynamics dependence on RNA concentration is intimately related to the droplet density decrease as a function of poly-U concentration and length shown in Figs. 3 *B* and 4 *A*. To better comprehend the underlying mechanism behind such behavior, we also measure D for poly-U strands of different lengths (i.e., 125 and 250 nt) within the condensates. In Fig. 5 *D*, we observe the severe impact of RNA chain length on its own mobility, as expected. Whereas FUS D (and also condensate stability and density) barely depends on the poly-U length (at least between 125 and 250 nucleotides), a twofold decrease in the RNA diffusion coefficient when adding 250 nt chains instead of 125 nt is behind the augment of droplet viscosity at high RNA concentration shown in Fig. 5 *B*. Moreover, when adding 10 nt chains, FUS (and RNA) diffusion considerably increases with respect to those with 125 or 250 nt (at the same RNA concentration) because of the droplet density reduction (Fig. 5 *D*). Interestingly, we also note that FUS, despite having the lower critical temperature to phase separate and thus weaker LLPS-stabilizing interactions than wt-TDP-43 and hnRNPA1 (Fig. 1 *B*), displays the lowest protein diffusion of the set in the absence of poly-U. Such an intriguing fact, which might be related to patterning sequence effects (158) or protein length (106), highlights how beyond stability, condensate dynamics also entail intricate processes that need to be further investigated. In fact, methods promoting LLPS at lower protein concentration or enhancing protein mobility, such as by short RNA inclusion, could play therapeutic roles in preventing the emergence of pathological solid-like aggregates (by decreasing droplet density and viscosity) related to some neurodegenerative disorders such as amyotrophic lateral sclerosis or multisystem proteinopathy (15,50,62).

DISCUSSION

Here, we investigate the dual effect of RNA in controlling the stability and dynamics of RNA-binding protein condensates. By means of a high-resolution sequence-dependent CG model for proteins and RNA (103,119,126), we explore via MD simulations the underlying molecular and thermodynamic mechanisms enabling liquid-liquid phase separation of FUS, hnRNPA1, and TDP-43 along their corresponding prion-like and RRM domains in the presence versus absence of RNA poly-U strands. After validating the model by comparing the relative ability of the aforementioned proteins (without RNA) to phase separate against their experimental protein saturation concentration—finding a remarkable qualitative agreement between both simulations and experiments—we characterize the condensates by determining their surface tension, the key molecular contacts sustaining LLPS, and the protein conformational ensemble in both phases. We find that highly inhomogeneous sequence contact maps, such as those of wt-TDP-43, can lead to the emergence of largely heterogeneous droplets with low surface tension, in which the exposure of PLD regions to the droplet interface deeply contributes to lowering γ and favoring multidroplet emulsions (142,159,160). However, such condensate heterogeneities can be significantly relieved when α-α helical PLD interactions are present, as recently hypothesized by Wang et al. (134). Moreover, the analysis of the intermolecular contact maps within our droplets reveals the major importance of certain sequence domains of these RBPs in LLPS, such as the hnRNPA1 PLD-PLD interactions or the FUS PLD-RGG interactions. Additionally, amino acid contacts such as G-G, R-Y, G-S, G-Y, K-F, and K-Y have been shown (Figs. S7–S9) to play a leading role in phase separation, highlighting the relevance of cation-π and electrostatic forces besides hydrophobicity in the physiological salt regime (61). Also, the conformational protein ensemble inside the condensates has been demonstrated to be almost independent of temperature, in contrast to those measured in the diluted phase (86). However, along the protein diluted-to-condensed transition, a significant enrichment towards more extended conformational ensembles (to maximize protein molecular connectivity (116)) has been observed.

Our simulations with poly-U RNA also reveal how the formation of protein condensates is clearly enhanced at low poly-U concentration (50), whereas inhibited at high poly-U/protein ratios (43,56,57). The RNA concentration that promotes the highest increase in droplet stability is near the point at which poly-U/FUS and poly-U/hnRNPA1 mixtures are electroneutral (and also for both hnRNPA1 RRMs and A-LCD regions separately), in agreement with findings for LAF-1-PLD condensates (119). We show how such a boost in droplet stability is related to an increase of the condensate surface tension and liquid-network connectivity at low RNA ratios. In contrast, neither of the two studied TDP-43 variants, nor their RRMs together or individually, exhibited significantly LLPS enhancement through poly-U addition with this model. Besides, we demonstrate that beyond a certain strand length of ~100 nucleotides, the stability of the droplets for a given RNA concentration reaches a plateau, whereas below that minimal chain length, as for very short lengths (i.e., ~10 nt), it can even hinder phase separation (57). These results have been shown to be related to the conformational structure and radius of gyration that RNA chains can adopt, which enable intermolecular binding between distinct proteins within the condensates. Overall, our results evidence how RBP condensate stability can be critically modulated by varying RNA concentration and length.

Finally, we focus on the transport properties of the RBP condensates as a function of RNA concentration

and length. Our simulations demonstrate that although viscosity severely depends on the length of the added RNA chains—i.e., poly-U strands of 10 and 125 nt reduce droplet viscosity (43), whereas 250-nucleotide strands moderately increase viscosity at high RNA concentration (49,58) (Fig. 5)—protein diffusion hardly depends on poly-U length and mainly depends on droplet density, which in turn is mainly controlled by RNA concentration. The droplet viscosity gain with RNA length comes from the slower relaxation times and RNA diffusion in crowded environments by long RNA chains (Fig. 5 *D*). However, the addition of moderately short RNA strands (i.e., with a similar or slightly lower R_g than those of the proteins) could help in promoting condensate dynamics without significantly destabilizing phase separation (Fig. 3 *B*). Our results suggest that the enhanced droplet dynamics at high RNA concentrations is mediated by a density reduction upon poly-U addition due to electrostatic repulsion. Taken together, our observations shed light on the crucial role of RNA (concentration and length) on the formation and phase behavior of RNA-protein complexes (54,58,125). Moreover, this work provides a novel, estimation of the transport properties of protein condensates by means of computer simulations, which could pave the way for future studies characterizing protein-RNA mobility in other relevant systems. Expanding our understanding of LLPS and the role of RNA in this process may drive solutions to precisely modulate aberrant liquid-to-solid transitions in the cell.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2021.11.003.

AUTHOR CONTRIBUTIONS

J.R.E. and J.R. designed the research. A.R.T. and A.G. built the protein models, A.R.T. performed the simulations. A.R.T., J.R.E., A.G., and J.R. analyzed the data. A.R.T. and J.R.E. wrote the initial version of the manuscript. All authors contributed and edited the manuscript. J.R.E. and J.R. supervised the research.

ACKNOWLEDGMENTS

We acknowledge Dr R. Collepardo-Guevara and Dr J. A. Joseph for useful discussions. We also acknowledge the constructive comments from the reviewers for helping us to improve the manuscript.

This project has received funding from the Oppenheimer Research Fellowship of the University of Cambridge. A.R.T. is funded by Universidad Politécnica de Madrid (ESTANCIAS-PIF-20-TYOSR8-13-4N0WPQ), A.G. is funded by an Engineering and Physical Sciences Research Council (EPSRC) studentship (EP/N509620/1) and a Winton scholarship. J.R. acknowledges funding from the Spanish Ministry of Economy and Competitivity (PID2019-105898GA-C22). J.R.E. also acknowledges funding from the Roger Ekins Research Fellowship of Emmanuel College. This work has been performed using resources provided by the Cambridge Tier-2 system operated by the University of Cambridge Research Computing Service (http://www.hpc.cam.ac.uk) funded by EPSRC Tier-2 capital grant EP/P020259/1. The authors gratefully acknowledge the Universidad Politécnica de Madrid (www.upm.es) for providing computing resources on the Magerit Supercomputer.

REFERENCES

1. Brangwynne, C. P., C. R. Eckmann, …, A. A. Hyman. 2009. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science.* 324:1729–1732.
2. Li, P., S. Banjade, …, M. K. Rosen. 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature.* 483:336–340.
3. Hyman, A. A., C. A. Weber, and F. Jülicher. 2014. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* 30:39–58.
4. Banani, S. F., H. O. Lee, …, M. K. Rosen. 2017. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18:285–298.
5. Shin, Y., and C. P. Brangwynne. 2017. Liquid phase condensation in cell physiology and disease. *Science.* 357:eaaf4382.
6. Alberti, S., A. Gladfelter, and T. Mittag. 2019. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell.* 176:419–434.
7. Nott, T. J., E. Petsalaki, …, A. J. Baldwin. 2015. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell.* 57:936–947.
8. Narlikar, G. J. 2020. Phase-separation in chromatin organization. *J. Biosci.* 45:5.
9. Brangwynne, C. P., P. Tompa, and R. V. Pappu. 2015. Polymer physics of intracellular phase transitions. *Nat. Phys.* 11:899–904.
10. Leblond, C. S., H. M. Kaneb, …, G. A. Rouleau. 2014. Dissection of genetic factors associated with amyotrophic lateral sclerosis. *Exp. Neurol.* 262:91–101.
11. Protter, D. S. W., B. S. Rao, …, R. Parker. 2018. Intrinsically disordered regions can contribute promiscuous interactions to RNP granule assembly. *Cell Rep.* 22:1401–1412.
12. Boeynaems, S., S. Alberti, …, M. Fuxreiter. 2018. Protein phase separation: a new phase in cell biology. *Trends Cell Biol.* 28:420–435.
13. Zhou, H., Z. Song, …, L. Lai. 2019. Mechanism of DNA-induced phase separation for transcriptional repressor VRN1. *Angew. Chem. Int.Engl.* 58:4858–4862.
14. Saha, S., C. A. Weber, …, A. A. Hyman. 2016. Polar positioning of phase-separated liquid compartments in cells regulated by an mRNA competition mechanism. *Cell.* 166:1572–1584.e16.
15. Molliex, A., J. Temirov, …, J. P. Taylor. 2015. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell.* 163:123–133.
16. Gui, X., F. Luo, …, D. Li. 2019. Structural basis for reversible amyloids of hnRNPA1 elucidates their role in stress granule assembly. *Nat. Commun.* 10:2006.
17. Su, X., J. A. Ditlev, …, R. D. Vale. 2016. Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science.* 352:595–599.
18. Larson, A. G., and G. J. Narlikar. 2018. The role of phase separation in heterochromatin formation, function, and regulation. *Biochemistry.* 57:2540–2548.
19. Larson, A. G., D. Elnatan, …, G. J. Narlikar. 2017. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature.* 547:236–240.
20. Strom, A. R., A. V. Emelyanov, …, G. H. Karpen. 2017. Phase separation drives heterochromatin domain formation. *Nature.* 547:241–245.

21. Sabari, B. R., A. Dall'Agnese, …, R. A. Young. 2018. Coactivator condensation at super-enhancers links phase separation and gene control. *Science.* 361:eaar3958.

22. Yoo, H., C. Triandafillou, and D. A. Drummond. 2019. Cellular sensing by phase separation: using the process, not just the products. *J. Biol. Chem.* 294:7151–7159.

23. Klosin, A., F. Oltsch, …, C. Zechner. 2020. Phase separation provides a mechanism to reduce noise in cells. *Science.* 367:464–468.

24. Sheu-Gruttadauria, J., and I. J. MacRae. 2018. Phase transitions in the assembly and function of human miRISC. *Cell.* 173:946–957.e16.

25. Franzmann, T. M., and S. Alberti. 2019. Prion-like low-complexity sequences: key regulators of protein solubility and phase behavior. *J. Biol. Chem.* 294:7128–7136.

26. Kroschwald, S., M. C. Munder, …, S. Alberti. 2018. Different material states of Pub1 condensates define distinct modes of stress adaptation and recovery. *Cell Rep.* 23:3327–3339.

27. Bouchard, J. J., J. H. Otero, …, T. Mittag. 2018. Cancer mutations of the tumor suppressor SPOP disrupt the formation of active, phase-separated compartments. *Mol. Cell.* 72:19–36.e8.

28. Moser, J. J., and M. J. Fritzler. 2010. Cytoplasmic ribonucleoprotein (RNP) bodies and their relationship to GW/P bodies. *Int. J. Biochem. Cell Biol.* 42:828–843.

29. Feric, M., and C. P. Brangwynne. 2013. A nuclear F-actin scaffold stabilizes ribonucleoprotein droplets against gravity in large cells. *Nat. Cell Biol.* 15:1253–1259.

30. Gall, J. G. 2003. The centennial of the Cajal body. *Nat. Rev. Mol. Cell Biol.* 4:975–980.

31. Brangwynne, C. P., T. J. Mitchison, and A. A. Hyman. 2011. Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci. USA.* 108:4334–4339.

32. Lu, H., D. Yu, …, Q. Zhou. 2018. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature.* 558:318–323.

33. Chen, Y., and A. S. Belmont. 2019. Genome organization around nuclear speckles. *Curr. Opin. Genet. Dev.* 55:91–99.

34. Chong, P. A., and J. D. Forman-Kay. 2016. Liquid-liquid phase separation in cellular signaling systems. *Curr. Opin. Struct. Biol.* 41:180–186.

35. Welsh, T. J., Y. Shen, …, T. P. J. Knowles. 2018. Mechanobiology of protein droplets: force arises from disorder. *Cell.* 175:1457–1459.

36. Shin, Y., Y.-C. Chang, …, C. P. Brangwynne. 2018. Liquid nuclear condensates mechanically sense and restructure the genome. *Cell.* 175:1481–1491.e13.

37. Riback, J. A., and C. P. Brangwynne. 2020. Can phase separation buffer cellular noise? *Science.* 367:364–365.

38. Qamar, S., G. Wang, …, P. St George-Hyslop. 2018. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interactions. *Cell.* 173:720–734.e15.

39. Murthy, A. C., G. L. Dignon, …, N. L. Fawzi. 2019. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* 26:637–648.

40. Rhoads, S. N., Z. T. Monahan, …, F. P. Shewmaker. 2018. The role of post-translational modifications on prion-like aggregation and liquid-phase separation of FUS. *Int. J. Mol. Sci.* 19:886.

41. Li, H.-R., W.-C. Chiang, …, J. R. Huang. 2018. TAR DNA-binding protein 43 (TDP-43) liquid-liquid phase separation is mediated by just a few aromatic residues. *J. Biol. Chem.* 293:6090–6098.

42. McGurk, L., E. Gomes, …, N. M. Bonini. 2018. Poly (ADP-ribose) prevents pathological phase separation of TDP-43 by promoting liquid demixing and stress granule localization. *Mol. Cell.* 71:703–717.e9.

43. Elbaum-Garfinkle, S., Y. Kim, …, C. P. Brangwynne. 2015. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl. Acad. Sci. USA.* 112:7189–7194.

44. Yang, P., C. Mathieu, …, J. P. Taylor. 2020. G3BP1 is a tunable switch that triggers phase separation to assemble stress granules. *Cell.* 181:325–345.e28.

45. Sanders, D. W., N. Kedersha, …, C. P. Brangwynne. 2020. Competing protein-RNA interaction networks control multiphase intracellular organization. *Cell.* 181:306–324.e28.

46. Guillén-Boixet, J., A. Kopach, …, T. M. Franzmann. 2020. RNA-induced conformational switching and clustering of G3BP drive stress granule assembly by condensation. *Cell.* 181:346–361.e17.

47. Cléry, A., M. Blatter, and F. H. Allain. 2008. RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.* 18:290–298.

48. Sanulli, S., M. J. Trnka, …, G. J. Narlikar. 2019. HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature.* 575:390–394.

49. Wei, M.-T., S. Elbaum-Garfinkle, …, C. P. Brangwynne. 2017. Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat. Chem.* 9:1118–1125.

50. Burke, K. A., A. M. Janke, …, N. L. Fawzi. 2015. Residue-by-residue view of in vitro FUS granules that bind the C-terminal domain of RNA polymerase II. *Mol. Cell.* 60:231–241.

51. Guo, L., and J. Shorter. 2015. It's raining liquids: RNA tunes viscoelasticity and dynamics of membraneless organelles. *Mol. Cell.* 60:189–192.

52. Alshareedah, I., T. Kaur, …, P. R. Banerjee. 2019. Interplay between short-range attraction and long-range repulsion controls reentrant liquid condensation of ribonucleoprotein-RNA complexes. *J. Am. Chem. Soc.* 141:14593–14602.

53. Alshareedah, I., G. M. Thurston, and P. R. Banerjee. 2021. Quantifying viscosity and surface tension of multicomponent protein-nucleic acid condensates. *Biophys. J.* 120:1161–1169.

54. Yu, M., and E. A. Lemke. 2021. There is plenty of room in protein-RNA condensates. *Biophys. J.* 120:1121–1122.

55. Schwartz, J. C., X. Wang, …, T. R. Cech. 2013. RNA seeds higher-order assembly of FUS protein. *Cell Rep.* 5:918–925.

56. Banerjee, P. R., A. N. Milin, …, A. A. Deniz. 2017. Reentrant phase transition drives dynamic substructure formation in ribonucleoprotein droplets. *Angew. Chem. Int.Engl.* 56:11354–11359.

57. Maharana, S., J. Wang, …, S. Alberti. 2018. RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science.* 360:918–921.

58. Zhang, H., S. Elbaum-Garfinkle, …, A. S. Gladfelter. 2015. RNA controls PolyQ protein phase transitions. *Mol. Cell.* 60:220–230.

59. Muzzopappa, F., M. Hertzog, and F. Erdel. 2021. DNA length tunes the fluidity of DNA-based condensates. *Biophys. J.* 120:1288–1300.

60. Zacco, E., R. Graña-Montes, …, A. Pastore. 2019. RNA as a key factor in driving or preventing self-assembly of the TAR DNA-binding protein 43. *J. Mol. Biol.* 431:1671–1688.

61. Wang, J., J.-M. Choi, …, A. A. Hyman. 2018. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell.* 174:688–699.e16.

62. Patel, A., H. O. Lee, …, S. Alberti. 2015. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell.* 162:1066–1077.

63. Li, Y. R., O. D. King, …, A. D. Gitler. 2013. Stress granules as crucibles of ALS pathogenesis. *J. Cell Biol.* 201:361–372.

64. Afroz, T., E.-M. Hock, …, M. Polymenidou. 2017. Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nat. Commun.* 8:45.

65. Lu, Y., L. Lim, and J. Song. 2017. RRM domain of ALS/FTD-causing FUS characteristic of irreversible unfolding spontaneously self-assembles into amyloid fibrils. *Sci. Rep.* 7:1043.

66. Wegmann, S., B. Eftekharzadeh, …, B. T. Hyman. 2018. Tau protein liquid-liquid phase separation can initiate tau aggregation. *EMBO J.* 37:e98049.

67. Choi, J.-M., A. S. Holehouse, and R. V. Pappu. 2020. Physical principles underlying the complex biology of intracellular phase transitions. *Annu. Rev. Biophys.* 49:107–133.

68. Portz, B., B. L. Lee, and J. Shorter. 2021. FUS and TDP-43 phases in health and disease. *Trends Biochem. Sci.* 46:550–563.

69. Sakon, J. J., and K. R. Weninger. 2010. Detecting the conformation of individual proteins in live cells. *Nat. Methods.* 7:203–205.

70. König, I., A. Zarrine-Afsar, …, B. Schuler. 2015. Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat. Methods.* 12:773–779.

71. LeBlanc, S. J., P. Kulkarni, and K. R. Weninger. 2018. Single molecule FRET: a powerful tool to study intrinsically disordered proteins. *Biomolecules.* 8:140.

72. Riback, J. A., M. A. Bowman, …, T. R. Sosnick. 2017. Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. *Science.* 358:238–241.

73. Bremer, A., M. Farag, …, T. Mittag. 2021. Deciphering how naturally occurring sequence features impact the phase behaviors of disordered prion-like domains. *bioRxiv* https://doi.org/10.1101/2021.01.01.425046.

74. Jawerth, L., E. Fischer-Friedrich, …, F. Jülicher. 2020. Protein condensates as aging Maxwell fluids. *Science.* 370:1317–1323.

75. Ryan, V. H., G. L. Dignon, …, N. L. Fawzi. 2018. Mechanistic view of hnRNPA2 low-complexity domain structure, interactions, and phase separation altered by mutation and arginine methylation. *Mol. Cell.* 69:465–479.e7.

76. Monahan, Z., V. H. Ryan, …, N. L. Fawzi. 2017. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* 36:2951–2967.

77. Kato, M., T. W. Han, …, S. L. McKnight. 2012. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell.* 149:753–767.

78. Dignon, G. L., R. B. Best, and J. Mittal. 2020. Biomolecular phase separation: from molecular driving forces to macroscopic properties. *Annu. Rev. Phys. Chem.* 71:53–75.

79. Roberts, S., V. Miao, …, A. Chilkoti. 2020. Complex microparticle architectures from stimuli-responsive intrinsically disordered proteins. *Nat. Commun.* 11:1342.

80. Paloni, M., R. Bailly, …, A. Barducci. 2020. Unraveling molecular interactions in liquid-liquid phase separation of disordered proteins by atomistic simulations. *J. Phys. Chem. B.* 124:9009–9016.

81. Feig, M., and Y. Sugita. 2019. Whole-cell models and simulations in molecular detail. *Annu. Rev. Cell Dev. Biol.* 35:191–211.

82. Shaw, D. E., P. Maragakis, …, W. Wriggers. 2006. The science of crystallization: microscopic phenomena and defect generation. *Proc. Natl. Acad. Sci. USA.* 82:36.

83. Lindorff-Larsen, K., S. Piana, …, D. E. Shaw. 2011. How fast-folding proteins fold. *Science.* 334:517–520.

84. Collepardo-Guevara, R., G. Portella, …, M. Orozco. 2015. Chromatin unfolding by epigenetic modifications explained by dramatic impairment of internucleosome interactions: a multiscale computational study. *J. Am. Chem. Soc.* 137:10205–10215.

85. Krainer, G., T. J. Welsh, …, T. P. J. Knowles. 2021. Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun.* 12:1085.

86. Dignon, G. L., W. Zheng, …, J. Mittal. 2018. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA.* 115:9929–9934.

87. Schuster, B. S., G. L. Dignon, …, J. Mittal. 2020. Identifying sequence perturbations to an intrinsically disordered protein that determine its phase-separation behavior. *Proc. Natl. Acad. Sci. USA.* 117:11421–11431.

88. Joseph, A. J., A. Reinhardt, …, R. Collepardo-Guevara. 2021. Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nat. Comput. Sci,* https://doi.org/10.1038/s43588-021-00155-3.

89. Garaizar, A., and J. R. Espinosa. 2021. Salt dependent phase behavior of intrinsically disordered proteins from a coarse-grained model with explicit water and ions. *J. Chem. Phys.* 155:125103.

90. Lin, Y.-H., J. P. Brady, …, H. S. Chan. 2017. Charge pattern matching as a 'fuzzy' mode of molecular recognition for the functional phase separations of intrinsically disordered proteins. *New J. Phys.* 19:115003.

91. McCarty, J., K. T. Delaney, …, J.-E. Shea. 2019. Complete phase diagram for liquid–liquid phase separation of intrinsically disordered proteins. *J. Phys. Chem. Lett.* 10:1644–1652.

92. Choi, J.-M., A. A. Hyman, and R. V. Pappu. 2020. Generalized models for bond percolation transitions of associative polymers. *Phys. Rev. E.* 102:042403.

93. Weber, C. A., D. Zwicker, …, C. F. Lee. 2019. Physics of active emulsions. *Rep. Prog. Phys.* 82:064601.

94. Harmon, T. S., A. S. Holehouse, …, R. V. Pappu. 2017. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife.* 6:e30294.

95. Das, S., A. Eisen, …, H. S. Chan. 2018. A lattice model of charge-pattern-dependent polyampholyte phase separation. *J. Phys. Chem. B.* 122:5418–5431.

96. Harmon, T. S., A. S. Holehouse, and R. V. Pappu. 2018. Differential solvation of intrinsically disordered linkers drives the formation of spatially organized droplets in ternary systems of linear multivalent proteins. *New J. Phys.* 20:045002.

97. Choi, J.-M., F. Dar, and R. V. Pappu. 2019. LASSI: a lattice model for simulating phase transitions of multivalent proteins. *PLoS Comput. Biol.* 15:e1007028.

98. Liu, H., S. K. Kumar, and F. Sciortino. 2007. Vapor-liquid coexistence of patchy models: relevance to protein phase behavior. *J. Chem. Phys.* 127:084902.

99. Nguemaha, V., and H.-X. Zhou. 2018. Liquid-liquid phase separation of patchy particles illuminates diverse effects of regulatory components on protein droplet formation. *Sci. Rep.* 8:6728.

100. Chou, H.-Y., and A. Aksimentiev. 2020. Single-protein collapse determines phase equilibria of a biological condensate. *J. Phys. Chem. Lett.* 11:4923–4929.

101. Joseph, J. A., J. R. Espinosa, …, R. Collepardo-Guevara. 2021. Thermodynamics and kinetics of phase separation of protein-RNA mixtures by a minimal model. *Biophys. J.* 120:1219–1230.

102. Garaizar, A., J. R. Espinosa, …, R. Collepardo-Guevara. 2021. Kinetic interplay between droplet maturation and coalescence modulates shape of aged protein condensates. *bioRxiv* https://doi.org/10.1101/2021.10.07.463530.

103. Dignon, G. L., W. Zheng, …, J. Mittal. 2018. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.* 14:e1005941.

104. Perdikari, T. M., N. Jovic, …, J. Mittal. 2021. A predictive coarse-grained model for position-specific effects of post-translational modifications. *Biophys. J.* 120:1187–1197.

105. Blas, F. J., L. G. MacDowell, …, G. Jackson. 2008. Vapor-liquid interfacial properties of fully flexible Lennard-Jones chains. *J. Chem. Phys.* 129:144703.

106. Silmore, K. S., M. P. Howard, and A. Z. Panagiotopoulos. 2017. Vapour–liquid phase equilibrium and surface tension of fully flexible Lennard–Jones chains. *Mol. Phys.* 115:320–327.

107. Das, R. K., and R. V. Pappu. 2013. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. USA.* 110:13392–13397.

108. Hazra, M. K., and Y. Levy. 2020. Charge pattern affects the structure and dynamics of polyampholyte condensates. *Phys. Chem. Chem. Phys.* 22:19368–19375.

109. Bianchi, E., J. Largo, …, F. Sciortino. 2006. Phase diagram of patchy colloids: towards empty liquids. *Phys. Rev. Lett.* 97:168301.

110. Banjade, S., Q. Wu, …, M. K. Rosen. 2015. Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. *Proc. Natl. Acad. Sci. USA.* 112:E6426–E6435.

111. Martin, E. W., A. S. Holehouse, …, T. Mittag. 2020. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science.* 367:694–699.

112. Ruff, K. M., F. Dar, and R. V. Pappu. 2021. Ligand effects on phase separation of multivalent macromolecules. *Proc. Natl. Acad. Sci. U S A.* 118:e2017184118.

113. Espinosa, J. R., A. Garaizar, …, R. Collepardo-Guevara. 2019. Breakdown of the law of rectilinear diameter and related surprises in the liquid-vapor coexistence in systems of patchy particles. *J. Chem. Phys.* 150:224510.

114. Russo, J., J. M. Tavares, …, F. Sciortino. 2011. Re-entrant phase behaviour of network fluids: a patchy particle model with temperature-dependent valence. *J. Chem. Phys.* 135:034501.

115. Blas, F. J., A. Galindo, and C. Vega. 2003. Study of the solid-liquid-vapour phase equilibria of flexible chain molecules using Wertheim's thermodynamic perturbation theory. *Mol. Phys.* 101:449–458.

116. Garaizar, A., I. Sanchez-Burgos, …, J. R. Espinosa. 2020. Expansion of intrinsically disordered proteins increases the range of stability of liquid-liquid phase separation. *Molecules.* 25:4705.

117. Espinosa, J. R., J. A. Joseph, …, R. Collepardo-Guevara. 2020. Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components. *Proc. Natl. Acad. Sci. U S A.* 117:13238–13247.

118. Dar, F., and R. V. Pappu. 2020. Multidimensional phase diagrams for multicomponent systems comprising multivalent proteins. *Biophys. J.* 118:213a.

119. Regy, R. M., G. L. Dignon, …, J. Mittal. 2020. Sequence dependent phase separation of protein-polynucleotide mixtures elucidated using molecular simulations. *Nucleic Acids Res.* 48:12593–12603.

120. Sanchez-Burgos, I., J. R. Espinosa, …, R. Collepardo-Guevara. 2021. Valency and binding affinity variations can regulate the multilayered organization of protein condensates with many components. *Biomolecules.* 11:278.

121. Damman, R., S. Schütz, …, M. Baldus. 2019. Atomic-level insight into mRNA processing bodies by combining solid and solution-state NMR spectroscopy. *Nat. Commun.* 10:4536.

122. Šponer, J., M. Krepl, …, M. Otyepka. 2017. How to understand atomistic molecular dynamics simulations of RNA and protein-RNA complexes? *Wiley Interdiscip. Rev. RNA.* 8:e1405.

123. Reyes, C. M., and P. A. Kollman. 2000. Structure and thermodynamics of RNA-protein binding: using molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. *J. Mol. Biol.* 297:1145–1158.

124. Kaur, T., M. Raju, …, P. R. Banerjee. 2020. Sequence-encoded and composition-dependent protein-RNA interactions control multiphasic condensate topologies. *bioRxiv* https://doi.org/10.1101/2020.08.30.273748.

125. Alshareedah, I., M. M. Moosa, …, P. R. Banerjee. 2020. Phase transition of RNA-protein complexes into ordered hollow condensates. *Proc. Natl. Acad. Sci. USA.* 117:15650–15658.

126. Das, S., Y.-H. Lin, …, H. S. Chan. 2020. Comparative roles of charge, π, and hydrophobic interactions in sequence-dependent phase separation of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA.* 117:28795–28805.

127. Welsh, T. J., G. Krainer, …, T. P. Knowles. 2020. Single particle zeta-potential measurements reveal the role of electrostatics in protein condensate stability. *bioRxiv* https://doi.org/10.1101/2020.04.20.047910.

128. Rowlinson, J. S., and B. Widom. 2013. Molecular Theory of Capillarity. Courier Corporation, Chelmsford, MA.

129. Farahi, N., T. Lazar, …, R. Pancsa. 2021. Concentration and dosage sensitivity of proteins driving liquid-liquid phase separation. *bioRxiv* https://doi.org/10.1101/2021.02.19.430946.

130. Li, Q., X. Peng, …, Z. Zhang. 2020. LLPSDB: a database of proteins undergoing liquid-liquid phase separation in vitro. *Nucleic Acids Res.* 48:D320–D327.

131. Lin, Y., D. S. Protter, …, R. Parker. 2015. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell.* 60:208–219.

132. Babinchak, W. M., R. Haider, …, W. K. Surewicz. 2019. The role of liquid-liquid phase separation in aggregation of the TDP-43 low-complexity domain. *J. Biol. Chem.* 294:6306–6317.

133. Lim, L., Y. Wei, …, J. Song. 2016. ALS-causing mutations significantly perturb the self-assembly and interaction with nucleic acid of the intrinsically disordered prion-like domain of TDP-43. *PLoS Biol.* 14:e1002338.

134. Wang, A., A. E. Conicella, …, N. L. Fawzi. 2018. A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J.* 37:e97452.

135. Conicella, A. E., G. L. Dignon, …, N. L. Fawzi. 2020. TDP-43 α-helical structure tunes liquid-liquid phase separation and function. *Proc. Natl. Acad. Sci. U S A.* 117:5883–5894.

136. Ladd, A. J., and L. V. Woodcock. 1977. Triple-point coexistence properties of the lennard-jones system. *Chem. Phys. Lett.* 51:155–159.

137. García Fernández, R., J. L. Abascal, and C. Vega. 2006. The melting point of ice Ih for common water models calculated from direct coexistence of the solid-liquid interface. *J. Chem. Phys.* 124:144506.

138. Espinosa, J. R., E. Sanz, …, C. Vega. 2013. On fluid-solid direct coexistence simulations: the pseudo-hard sphere model. *J. Chem. Phys.* 139:144502.

139. Choi, U. B., H. Sanabria, …, K. R. Weninger. 2019. Spontaneous switching among conformational ensembles in intrinsically disordered proteins. *Biomolecules.* 9:114.

140. Choi, U. B., J. J. McCann, …, M. E. Bowen. 2011. Beyond the random coil: stochastic conformational switching in intrinsically disordered proteins. *Structure.* 19:566–576.

141. Vis, M., E. M. Blokhuis, …, H. N. W. Lekkerkerker. 2018. Interfacial tension of phase-separated polydisperse mixed polymer solutions. *J. Phys. Chem. B.* 122:3354–3362.

142. Sanchez-Burgos, I., J. A. Joseph, …, J. R. Espinosa. 2021. Size conservation emerges spontaneously in biomolecular condensates formed by scaffolds and surfactant clients. *Sci. Rep.* 11:15241.

143. Majumdar, A., P. Dogra, …, S. Mukhopadhyay. 2019. Liquid-liquid phase separation is driven by large-scale conformational unwinding and fluctuations of intrinsically disordered protein molecules. *J. Phys. Chem. Lett.* 10:3929–3936.

144. Emmanouilidis, L., L. Esteban-Hofer, …, F. H. Allain. 2021. NMR and EPR reveal a compaction of the RNA-binding protein FUS upon droplet formation. *Nat. Chem. Biol.* 17:608–614.

145. Mann, J. R., and C. J. Donnelly. 2021. RNA modulates physiological and neuropathological protein phase transitions. *Neuron.* 109:2663–2681.

146. Kuo, P.-H., C.-H. Chiang, …, H. S. Yuan. 2014. The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic Acids Res.* 42:4712–4722.

147. Lukavsky, P. J., D. Daujotyte, …, F. H. Allain. 2013. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat. Struct. Mol. Biol.* 20:1443–1449.

148. Zacco, E., S. R. Martin, …, A. Pastore. 2018. The RNA-recognition motifs of TAR DNA-binding protein 43 may play a role in the aberrant self-assembly of the protein. *Front. Mol. Neurosci.* 11:372.

149. Bhardwaj, A., M. P. Myers, …, F. E. Baralle. 2013. Characterizing TDP-43 interaction with its RNA targets. *Nucleic Acids Res.* 41:5062–5074.

150. Piana, S., P. Robustelli, …, D. E. Shaw. 2020. Development of a force field for the simulation of single-chain proteins and protein–protein complexes. *J. Chem. Theory Comput.* 16:2494–2507.

151. Krepl, M., A. Cléry, …, J. Sponer. 2016. Synergy between NMR measurements and MD simulations of protein/RNA complexes: application to the RRMs, the most common RNA recognition motifs. *Nucleic Acids Res.* 44:6452–6470.

152. Boeynaems, S., A. S. Holehouse, …, A. D. Gitler. 2019. Spontaneous driving forces give rise to protein-RNA condensates with coexisting phases and complex material properties. *Proc. Natl. Acad. Sci. USA.* 116:7889–7898.

153. Mason, T. G., and D. A. Weitz. 1995. Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids. *Phys. Rev. Lett.* 74:1250–1253.

154. Fisher, R. S., and S. Elbaum-Garfinkle. 2020. Tunable multiphase dynamics of arginine and lysine liquid condensates. *Nat. Commun.* 11:4628.

155. Lavalette, D., C. Tétreau, …, Y. Blouquit. 1999. Microscopic viscosity and rotational diffusion of proteins in a macromolecular environment. *Biophys. J.* 76:2744–2751.

156. Rubinstein, M., and R. H. Colby. 2003. Polymer Physics, volume 23. Oxford University Press, Oxford, UK.

157. Ramírez, J., S. K. Sukumaran, …, A. E. Likhtman. 2010. Efficient on the fly calculation of time correlation functions in computer simulations. *J. Chem. Phys.* 133:154103.

158. Statt, A., H. Casademunt, …, A. Z. Panagiotopoulos. 2020. Model for disordered proteins with strongly sequence-dependent liquid phase behavior. *J. Chem. Phys.* 152:075101.

159. Garcia-Jove Navarro, M., S. Kashida, …, Z. Gueroui. 2019. RNA is a critical element for the sizing and the composition of phase-separated RNA-protein condensates. *Nat. Commun.* 10:3230.

160. Lu, T., and E. Spruijt. 2020. Multiphase complex coacervate droplets. *J. Am. Chem. Soc.* 142:2905–2914.

Biophysical Journal, Volume *120*

Supplemental information

`RNA modulation of transport properties and stability in phase-separated condensates

Andrés R. Tejedor, Adiran Garaizar, Jorge Ramírez, and Jorge R. Espinosa

RNA modulation of transport properties and stability in phase separated condensates (Supporting Information)

Andrés R. Tejedor and Jorge Ramírez*

Department of Chemical Engineering,

Universidad Politécnica de Madrid,

José Gutiérrez Abascal 2, 28006, Madrid, Spain

Adiran Garaizar and Jorge R. Espinosa†

Maxwell centre, Cavendish Laboratory, Department of physics,

University of Cambridge, J J Thomson Avenue,

Cambridge CB3 0HE, United Kingdom

(Dated: November 18, 2021)

* jorge.ramirez.upm.es

† jr752@cam.ac.uk

SI. MODEL AND METHODS

A. Protein/RNA HPS model

In order to simulate the different studied proteins and RNA, we employ the LAMMPS Molecular Dynamics simulation package [1] with the recent reparameterization by Das *et al.* [2] of the chemically-accurate coarse-grained (CG) HPS protein model proposed by Dignon *et al.* [3]. For RNA, we use the new HPS-compatible CG model proposed by Regy *et al.* [4]. The coarse-grained model resolution, both for proteins and RNA, is of one bead per amino acid and nucleotide. In the model, the intrinsically disordered regions (IDRs) of the proteins are considered as fully flexible polymers, and the structured globular domains are treated as rigid bodies (where their conformations are taken from the Protein Data Bank (PDB) crystalline structure — see SII for the PBD codes) by using the rigid body integrator of LAMMPS [1]. Moreover, the interactions of the structured globular domains are scaled down by a 30% to account for the 'buried' amino acids as shown by Krainer *et al.* [5]. Also, in this model RNA strands are treated as flexible polymers.

The potential energy of the coarse-grained force field is given by:

$$E = E_{\text{Bonds}} + E_{\text{Electrostatic}} + E_{\text{Hydrophobic}} + E_{\text{Cation}-\pi}, \tag{S1}$$

where $E_{\text{Hydrophobic}}$, $E_{\text{Cation}-\pi}$ and $E_{\text{Electrostatic}}$ interactions are only applied between non-bonded beads and E_{Bonds} between subsequent beads directly bonded to each other.

Bonded interactions between subsequent amino acid protein beads or consecutive RNA nucleotides are described by the harmonic potential:

$$E_{\text{Bonds}} = \sum_{\text{Protein/RNA bonds}} k(r - r_0)^2, \tag{S2}$$

where the equilibrium bond length is $r_0 = 5.0$Å between subsequent nucleotides and $r_0 = 3.81$Å between bonded amino acid beads. The spring constant is $k = 10\,\text{kJ}/(\text{mol}\text{Å}^2)$. Please see section SI C for further details on the value of this model parameter.

The electrostatic interactions, $E_{\text{Electrostatic}}$, among charged amino acids and RNA nucleotides are described by a Yukawa/Debye-Hückel potential of the form:

$$E_{\text{Electrostatic}} = \sum_i \sum_{j<i} \frac{1}{4\pi D} \frac{q_i q_j}{r} e^{-r/\kappa}, \tag{S3}$$

where q_i and q_j represent the charges of the beads i and j (amino acids or nucleotides), $D = 80$ is the dielectric constant of water, r is the distance between the ith and jth beads, and κ= 1 nm is the Debye screening length that mimics the implicit solvent (water and ions) at physiological salt concentration (∼150mM of NaCl) [3].

The hydrophobic interactions between different amino acid types and nucleotides are built upon a scale of amino acid and RNA nucleotide hydrophobicity based on a statistical potential derivation from contacts in PDB structures, and implemented through the functional form of an Ashbaugh/Hatch potential (see further details on these References [3, 4, 6, 7]):

$$E_{\text{Hydrophobic}} = \sum_i \sum_{j<i} \begin{cases} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] + (1-\lambda_{ij})\epsilon_{ij}, & r < 2^{1/6}\sigma_{ij} \\ \lambda_{ij} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right], & \text{otherwise,} \end{cases} \tag{S4}$$

where λ_i and λ_j are parameters that account for the hydrophobicity of the ith and jth interacting particles respectively, being $\lambda_{ij} = (\lambda_i + \lambda_j)/2$. The excluded volume of the different residues/nucleotides is given by σ_i and σ_j, where $\sigma_{ij} = (\sigma_i + \sigma_j)/2$, and r is the distance between the ij particles. ϵ_{ij} (0.2 kcal/mol) is a fitting parameter to reproduce experimental single-IDR radius of gyration (R_g) [3]. When at least one of the ij amino acids is part of a structured globular domain, λ_{ij} is scaled by a factor of 0.7 to account for 'buried' amino acids in globular domains. The specific values for each amino acid and nucleotide σ, q, and λ parameters can be found in References: Dignon *et al.* for proteins [3] and Regy *et al.* for RNA [4].

Finally, we consider an extra term for describing cation-π interactions (only for the following set of pairs of amino acids (c-π:{Arg-Phe, Arg-Trp, Arg-Tyr, Lys-Phe, Lys-Trp and Lys-Tyr}):

$$E_{\text{cation-}\pi} = \sum_{i\in\text{c-}\pi} \sum_{j\in\text{c-}\pi \& j<i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right], \tag{S5}$$

where σ_{ij} is the same as in the hydrophobic interactions and ϵ_{ij} is $\epsilon_{ij} = 3.0\text{kcal}\,\text{mol}^{-1}$ for all six cation-π pairs as proposed in Ref. [2] (Approach 1). Consistently, the interaction of these amino acids is scaled down by a 30% when they are found in structured globular

domains.

B. Simulation details

All simulations were carried out using LAMMPS [1] software. Direct Coexistence simulations (described in Section III) were carried out in the NVT ensemble using a Nosé-Hover thermostat [8] for the rigid bodies (crystalline structured domains of the proteins), and a Langevin thermostat [9] for the rest of the particles, both with a relaxation time of 5 ps. The timestep for the Verlet integration of the equations of motion was chosen to be of 10 fs. NPT simulations for pure bulk protein liquids (see section V) were carried out at p=1 bar using a Nosé-Hover barostat [1] and thermostat [8] with relaxation times of 50 ps and 5 ps respectively. For computational efficiency, we use a cut-off of $3\sigma_{ij}$ for the cation-π and hydrophobic interactions and 3.5 nm for the electrostatic ones [3]. We also turn off the interactions between particles that are part of the same globular structured domain (i.e., within the same rigid body).

C. HPS model spring constant

The spring constant used in our simulations is $k = 10\,\mathrm{kJ/(mol\,\AA^2)}$, as described in the reparameterization of the HPS model with cation-π interactions proposed by Das *et al* [2]. However, the HPS model was previously formulated with spring constants of $k = 10\,\mathrm{kJ/(mol\,\AA^2)}$ [3] and $k = 10\,\mathrm{kcal/(mol\,\AA^2)}$ [10]. To quantify the precise effect of k in the phase behavior of IDRs, we perform Direct Coexistence simulations using the HPS model without cation-π interactions to directly compare our results for different spring constants (k) with those provided by Dignon *et al.* for FUS40 [3] (Fig. S1).

We find that, despite the value of the spring constant does not dramatically change the phase diagram, the value that best reproduces the results by Dignon *et al.* [3] is $k = 10\,\mathrm{kJ/(mol\AA^2)}$. For this reason, and accordingly to the work of Das *et al.* in the HPS-Cation-π model [2], we use a spring constant of $k = 10\,\mathrm{kJ/(mol\AA^2)}$ or equivalently $k = 2.4\,\mathrm{kcal/(mol\AA^2)}$.

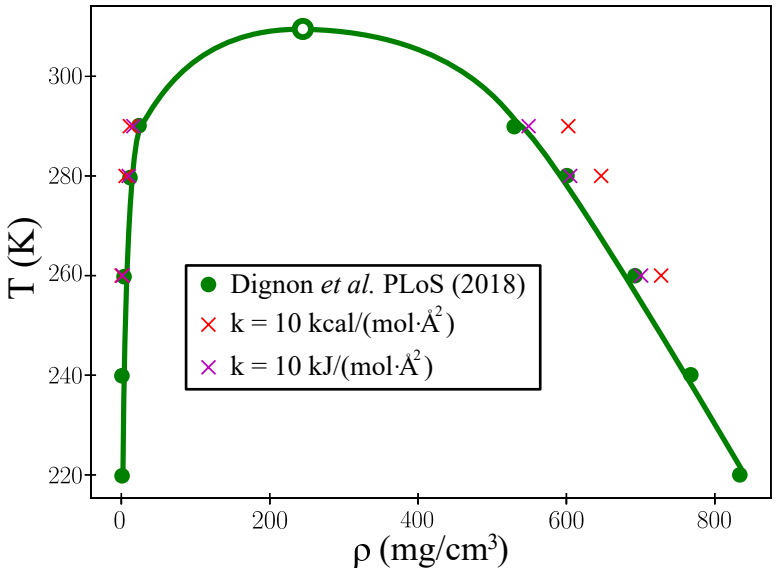


FIG. S1: Temperature-density phase diagram of FUS40 for two different spring constants (k) between bonded interactions, as indicated in the legend (see Eq. (S2)). For comparison, we also include the phase diagram obtained by Dignon *et al* [3] (green circles). The empty circle shows the critical temperature obtained by Dignon *et al* in their original work ($T_c = 309.6K$). The solid line is included to help visualizing the phase diagram.

D. Experimental validation of the HPS model without the Cation–π reparameterization

Here, we reproduce the comparison between the critical temperature measured in simulations and the experimental protein saturation concentration (as in Fig. 1 of the main text), including results from the HPS model without the additional cation–π enhanced interactions proposed by Das *et al.* [2]. We have run simulations of the HPS model for FUS, hnRNPA1, and wt-TDP-43, and we have computed their critical temperature using the law of rectilinear diameters and critical exponents [11]. The results for FUS-PLD can be directly adopted from the cation–π reparameterization since the FUS-PLD sequence does not contain pairs of amino acids involved in cation–π interactions. We show these results in Fig. S2.

Please recall that, although the absolute critical temperature of FUS-PLD is the same for both models, the critical temperature T_c' of wt-TDP-43, used to normalize the data of each set, is not. As it can be seen in Figure S2, the correlation between the critical temperature from simulations and the experimental saturation concentration of the proteins exhibited by the HPS model+cation–π (Figure 1 of the main text), is no longer present in the case of the HPS model (cross symbols). More notably, the critical temperature of

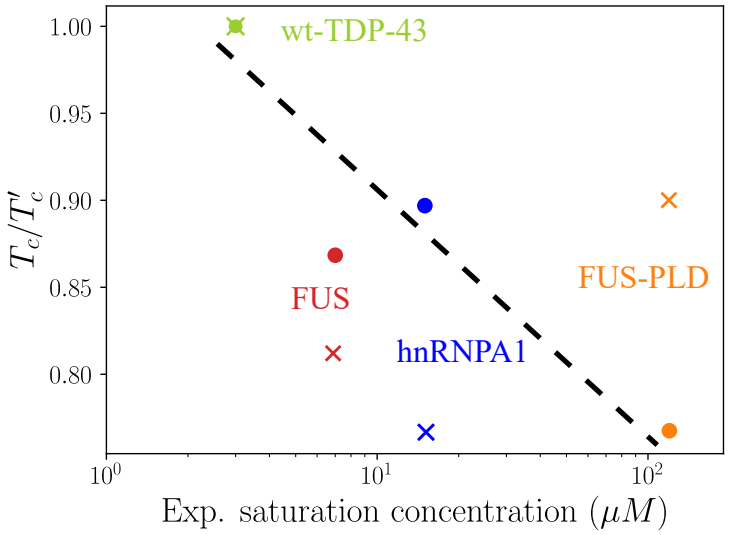


FIG. S2: Comparison of the renormalized critical temperature measured in simulations with the HPS+cation–π (filled circles) and the HPS models (crosses) against the experimental saturation concentration of the proteins. The experimental saturation concentrations at physiological salt conditions were obtained from the references provided in the main text (Fig. 1). Temperature is renormalized by the highest critical temperature T_c' of each set, corresponding to wt-TDP-43 for both models.

FUS is way below the critical point of FUS-PLD, which clearly contradicts the experimental trend. The saturation concentration of FUS-PLD is approximately one order of magnitude greater than the corresponding full-sequenced FUS, and thus, the critical point of the full protein is expected to be greater than the T_c of FUS-PLD. These results show the necessity of including further reparameterization in the HPS model (i.e., an extra term for cation–π interactions) as proposed by Das *et al.* [2].

SII. SEQUENCES AND PDBS OF THE STUDIED PROTEINS

FUS

MASNDYTQQATQSYGAYPTQPGQGYSQQSSQPYGQQSYSGYSQSTDTSGYGQSSYSSYGQSQNTG
YGTQSTPQGYGSTGGYGSSQSSQSSYGQQSSYPGYGQQPAPSSTSGSYGSSSQSSSYGQPQSGSYSQ
QPSYGGQQQSYGQQQSYNPPQGYGQQNQYNSSSGGGGGGGGGGNYGQDQSSMSSGGGSGGGYG
NQDQSGGGGSGGYGQQDRGGRGRGGSGGGGGGGGGGYNRSSGGYEPRGRGGGRGGRGGMGGS
DRGGFNKFGGPRDQGSRHDSEQDNSDNNTIFVQGLGENVTIESVADYFKQIGIIKTNKKTGQPMIN
LYTDRETGKLKGEATVSFDDPPSAKAAIDWFDGKEFSGNPIKVSFATRRADFNRGGGNGRGGRGR
GGPMGRGGYGGGGSGGGGRGGFPSGGGGGGGGQQRAGDWKCPNPTCENMNFSWRNECNQCKA
PKPDGPGGGPGGSHMGGNYGDDRRGGRGGYDRGGYRGRGGDRGGFRGGRGGGDRGGFGPGK
MDSRGEHRQDRRERPY

FUS-PLD

MASNDYTQQATQSYGAYPTQPGQGYSQQSSQPYGQQSYSGYSQSTDTSGYGQSSYSSYGQSQNTG
YGTQSTPQGYGSTGGYGSSQSSQSSYGQQSSYPGYGQQPAPSSTSGSYGSSSQSSSYGQPQSGSYSQ
QPSYGGQQQSYGQQQSYNPPQGYGQQNQYNS

hnRNPA1

MSKSESPKEPEQLRKLFIGGLSFETTDESLRSHFEQWGTLTDCVVMRDPNTKRSRGFGFVTYATVE
EVDAAMNARPHKVDGRVVEPKRAVSREDSQRPGAHLTVKKIFVGGIKEDTEEHHLRDYFEQYGKI
EVIEIMTDRGSGKKRGFAFVTFDDHDSVDKIVIQKYHTVNGHNCEVRKALSKQEMASASSSQRGRS
GSGNFGGGRGGGFGGNDNFGRGGNFSGRGGFGGSRGGGGYGGSGDGYNGFGNDGGYGGGGPG
YSGGSRGYGSGGQGYGNQGSGYGGSGSYDSYNNGGGGGFGGGSGSNFGGGGSYNDFGNYNNQSS
NFGPMKGGNFGGRSSGPYGGGGGQYFAKPRNQGGYGGSSSSSSYGSGRRF

hnRNPA1-PLD

GDGYNGFGNDGGYGGGGPGYSGGSRGYGSGGQGYGNQGSGYGGSGSYDSYNNGGGGGFGGGSG
SNFGGGGSYNDFGNYNNQSSNFGPMKGGNFGGRSSGPYGGGGGQYFAKPRNQGGYGGSSSSSSYGS
GRRF

hnRNPA1-RRMs

MSKSESPKEPEQLRKLFIGGLSFETTDESLRSHFEQWGTLTDCVVMRDPNTKRSRGFGFVTYATVE
EVDAAMNARPHKVDGRVVEPKRAVSREDSQRPGAHLTVKKIFVGGIKEDTEEHHLRDYFEQYGKI
EVIEIMTDRGSGKKRGFAFVTFDDHDSVDKIVIQKYHTVNGHNCEVRKALSKQ

hnRNPA1-A-LCD

MASASSSQRGRSGSGNFGGGRGGGFGGNDNFGRGGNFSGRGGFGGSRGGGGYGGSGDGYNGFG
NDGSNFGGGGSYNDFGNYNNQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGYGGSSSSSS
YGSGRRF

TDP-43

MSEYIRVTEDENDEPIEIPSEDDGTVLLSTVTAQFPGACGLRYRNPVSQCMRGVRLVEGILHAPDAG
WGNLVYVVNYPKDNKRKMDETDASSAVKVKRAVQKTSDLIVLGLPWKTTEQDLKEYFSTFGEVL
MVQVKKDLKTGHSKGFGFVRFTEYETQVKVMSQRHMIDGRWCDCKLPNSKQSQDEPLRSRKVFV
GRCTEDMTEDELREFFSQYGDVMDVFIPKPFRAFAFVTFADDQIAQSLCGEDLIIKGISVHISNAEPK
HNSNRQLERSGRFGGNPGGFGNQGGFGNSRGGGAGLGNNQGSNMGGGMNFGAFSINPAMMAAA
QAALQSSWGMMGMLASQQNQSGPSGNNQNQGNMQREPNQAFGSGNNSYSGSNSGAAIGWGSASN
AGSGSGFNGGFGSSMDSKSSGWGMMSEYIRVTEDENDEPIEIPSEDDGTVLLSTVTAQFPGACGLR
YRNPVSQCMRGVRLVEGILHAPDAGWGNLVYVVNYPKDNKRKMDETDASSAVKVKRAVQKTSD
LIVLGLPWKTTEQDLKEYFSTFGEVLMVQVKKDLKTGHSKGFGFVRFTEYETQVKVMSQRHMID
GRWCDCKLPNSKQSQDEPLRSRKVFVGRCTEDMTEDELREFFSQYGDVMDVFIPKPFRAFAFVTF
ADDQIAQSLCGEDLIIKGISVHISNAEPKHNSNRQLERSGRFGGNPGGFGNQGGFGNSRGGGAGLGN
NQGSNMGGGMNFGAFSINPAMMAAAQAALQSSWGMMGMLASQQNQSGPSGNNQNQGNMQREP
NQAFGSGNNSYSGSNSGAAIGWGSASNAGSGSGFNGGFGSSMDSKSSGWGM

TDP-43-PLD

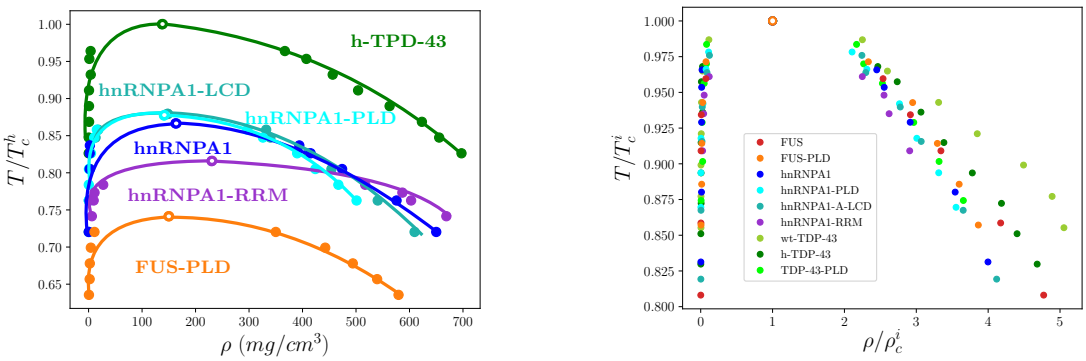
GRFGGNPGGFGNQGGFGNSRGGGAGLGNNQGSNMGGGMNFGAFSINPAMMAAAQAALQSSWG
MMGMLASQQNQSGPSGNNQNQGNMQREPNQAFGSGNNSYSGSNSGAAIGWGSASNAGSGSGFNG
GFGSSMDSKSSGWGM

The following Protein Data Bank (PDB) codes were used to build the globular structured domains of: FUS (residues from 285–371 (PDB code: 2LCW) and from 422–453 (PDB code:

6G99)), wt-TDP-43 (residues 2-38, 40-49 and 51-79 all included in the same PDB, (PDB code: 5MDI) and from residues 193-267 (PDB code: 1WF0)), h-TDP-43 (additionally to the structured domains of wt-TDP-43, this variant has an α–helical domain from residues 307-349 (PBD code: 2N2C)) and hnRNPA1 (residues from 8-91 and 103-181 in the same PDB (PDB code: 1L3K)). The intrinsically disordered regions not included in the PDBs were built using the VMD software [12].

SIII. PHASE DIAGRAM CALCULATIONS VIA DIRECT COEXISTENCE SIMULATIONS

We perform Direct Coexistence (DC) simulations [13–16] to compute the phase diagram of the different proteins and protein/RNA mixtures (See Table S1 for the employed system sizes). Within DC simulations, the two coexisting phases of the system are simulated in the same simulation box. In our case, we place a high-density protein liquid and a very low-density one. We employ a rectangular box, with an elongated side perpendicular to the interfaces (long enough to capture the bulk density of each phase), while the parallel sides are chosen such that proteins and RNA cannot interact with themselves along the periodic boundary conditions. We run NVT simulations until equilibrium is reached. Then, we measure the equilibrium coexisting densities of both phases along the long side of the box, excluding the fluctuations of the interfaces and keeping the center of mass of the slab fixed. We repeat this procedure at different temperatures until we reach the critical temperature. To avoid finite system-size effects close to the critical point, we evaluate the critical temperature (T_c) and density (ρ_c) using the law of critical exponents and rectilinear diameters [11]:



(a) **(b)**

FIG. S3: **(a)** Temperature-density phase diagram (normalised by T_c^h, where T_c^h is the critical temperature of h-TDP-43, T_c^h=472K) for FUS-PLD (orange), hnRNPA1 (dark blue), hnRNPA1-RRM (purple), hnRNPA1-A-LCD (turquoise), hnRNPA1-PLD (cyan) and h-TDP-43 (green). The critical density and temperature (depicted by empty circles) have been obtained using the law of rectilinear diameters and critical exponents, Eqs. (S6) and (S7). **(b)** Temperature-density phase diagram for all the studied proteins where both temperature and density are renormalized by their own critical temperature (T_c^i) and critical density (ρ_c^i). The shown sequences are: FUS (red), hnRNPA1 (dark blue), wt-TDP-43 (lime green), h-TDP-43 (dark green), TDP-43-PLD (light green), hnRNPA1-A-LCD (turquoise), hnRNPA1-PLD (cyan), FUS-PLD (orange) and hnRNPA1-RRM (purple). Both critical density and temperature for each protein were obtained as described in (a).

Protein	Chains (N_c)	Protein residues
FUS	24	526
FUS-PLD	80	163
hnRNPA1	40	372
hnRNPA1-PLD	100	132
hnRNPA1-RRM	80	184
hnRNPA1-A-LCD	200	135
TDP-43	32	414
TDP-43-PLD	100	141

TABLE S1: Employed systems sizes in Direct Coexistence simulations of each protein type (in pure component) including the number of protein replicas (N_c) and the number of amino acids per protein. Simulations including poly-U contain the same number of proteins, except for hnRNPA1-A-LCD, which was reduced to $N_c = 100$ to keep similar RNA/protein ratios in all systems.

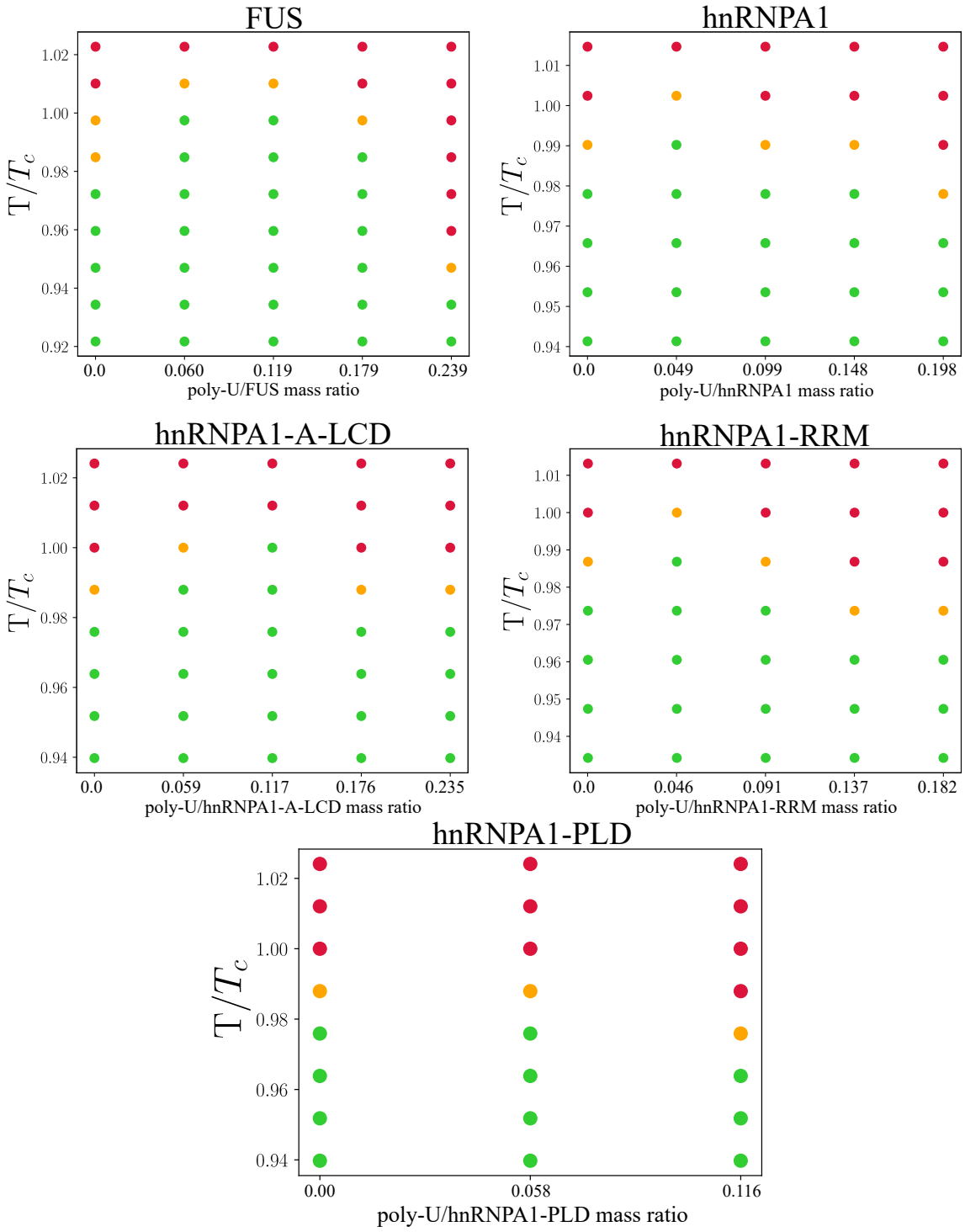


FIG. S4: Phase behaviour of FUS, hnRNPA1, hnRNPA1-A-LCD, hnRNPA1-RRM and hnRNPA1-PLD condensates at different poly-U concentrations as a function of temperature (renormalized by the critical temperature (T_c) in absence of poly-U (Table S2). Green circles indicate temperatures/concentrations where phase separation was observed in our DC simulations, and red circles where no phase separation was observed. Orange dots show the limit between both regimes. Note that LLPS cannot be directly observed by means of DC simulations just below the critical temperature due to finite size effects, and thus, the highest temperature at which we can usually observe LLPS by means of DC simulations is $T/T_c \sim 0.98$.

$$(\rho_l - \rho_v)^\alpha = s_1 \left(1 - \frac{T}{T_c}\right) \tag{S6}$$

$$\frac{\rho_l + \rho_v}{2} = \rho_c + s_2(T_c - T), \tag{S7}$$

where ρ_l and ρ_v refer to the densities of the condensed and the diluted phases respectively, s_1 and s_2 are fitting parameters, and $\alpha = 3.06$ accounts for the critical exponent of the three dimensional Ising model [11].

Protein	Method 1 (T_c/K)	Method 2 (T_c/K)
FUS	396	400
FUS-PLD	350	347
hnRNPA1	409	407
hnRNPA1-RRM	385	390
hnRNPA1-PLD	414	411
hnRNPA1-A-LCD	415	413
h-TDP-43	472	468
wt-TDP-43	456	448
TDP-43-PLD	366	367

TABLE S2: Comparison of the critical temperatures estimated by using the law of rectilinear diameters and critical exponents (Method 1), and by fitting the surface tension data as a function of temperature using the fit given in Eq. (S9) (Method 2).

poly-U / Protein	0 (T_c/K)	1 (T_c/K)	2 (T_c/K)	3 (T_c/K)	4 (T_c/K)
FUS	396	405	401	398	386
hnRNPA1	409	412	417	407	405
hnRNPA1-A-LCD	415	418	419	411	409
hnRNPA1-PLD	414	416	406	-	-
hnRNPA1-RRM	385	391	382	378*	375*

TABLE S3: Comparison of the estimated critical temperatures (using Method 1) for the different poly-U/protein mixtures as a function of poly-U concentration, given in number of added poly-U strands of 250 nucleotides. The number of proteins in each system is that given in Table S1, and is constant for all concentrations. The equivalence in poly-U/protein mass ratio with the number of added of 250 nt poly-U strands can be extracted by comparison with Fig. S4. Marked temperatures with asterisk have been estimated according to the general tendency since there are not enough data to use Eq. (S6).

SIV. INTERFACIAL FREE ENERGY CALCULATIONS

As previously explained, in DC simulations two phases coexist in the same simulation box. Since one of the box sides (*e.g.* z-axis) is longer than the other two, the bulk density of the two phases can be conveniently measured along the corresponding axis. Moreover, when the system is equilibrated, the surface tension can be also evaluated from the computed pressure tensor. By means of the following expression, the interfacial free energy (γ) can be evaluated [17]:

$$\gamma = \frac{L_N}{2}(p_N - p_T), \tag{S8}$$

where p_N denotes the normal component of the pressure tensor perpendicular to the interface, p_T represents the average of the tangential components of the pressure tensor, L_N denotes the length of the long side of the simulation box and the 2 factor accounts for the presence of two interfaces in the simulation box. The surface tensions shown in Fig. S5 (and Figs. 2(A) and 3(F) of the main text) have been computed using this expression. Furthermore, γ values can be also used to alternatively estimate the critical temperature by assuming the following scaling [11]:

$$\gamma = A(T_c - T)^{1.26}, \tag{S9}$$

where T_c and A are fitting parameters. The critical temperature can be estimated as the temperature at which γ becomes zero [18].

SV. CONFORMATIONAL PROTEIN ENSEMBLE

We analyse the protein conformational ensemble of the different sequences along the condensed and diluted liquid phases. For this purpose, we compute the histogram of the radius of gyration (R_g) distribution function of the proteins in both phases. To measure R_g in the diluted phase, we run NVT simulations with a single protein at different temperatures and at the coexisting density of each temperature according to the phase diagram. Upon reaching equilibrium, we compute the R_g histograms along the simulation. To measure R_g inside the condensates, we first equilibrate the liquid of the given protein at the desired

temperature and density, and then, we run NVT simulations in which we compute the R_g of all the protein replicas along time. The averaged R_g distributions over time and protein replicas are represented in Fig. S6 for the different studied sequences within the condensates (continuous lines) and in the diluted phase (dashed ones) for the lowest and highest temperature at which phase separation was observed.

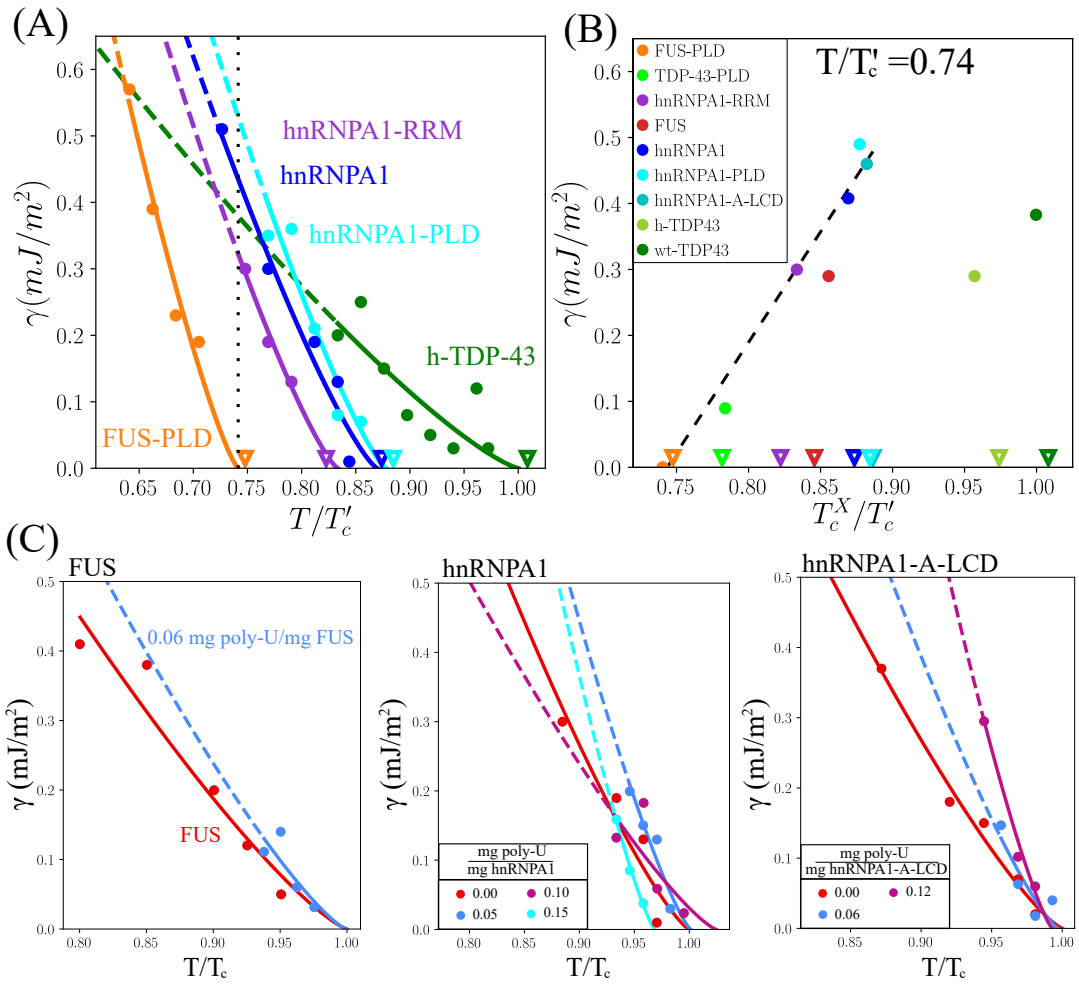


FIG. S5: (A): Droplet surface tension (γ) of hnRNPA1-RRM, hnRNPA1, hnRNPA1-PLD, h-TDP-43 and FUS-PLD condensates as a function of T/T_c' where T_c' is the critical temperature of h-TDP-43 as obtained by the extrapolation method explained above (Eq. S9). Filled circles indicate the value of γ as obtained from DC simulations and solid curves are the representation of the curves recovered by fitting our data to Eq. (S9). Dashed curves depict the predicted surface tension at low T as extrapolated from the fit. Empty triangles represent the predicted (renormalized by T_c') critical point of each system (T_c^X) using the laws of rectilinear diameters and critical exponents. See Table S2 for further details on these estimates. Black dotted line indicates the temperature selected for panel (B). (B) Interfacial tension measured at $T/T_c' = 0.74$ (see panel (A)) for all the systems as a function of the critical temperature obtained from DC simulations using the fit of data to Eq. (S9). The critical temperature of each system (T_c^X) is normalized by T_c' which corresponds to the critical temperature of h-TDP-43. Empty triangles show the critical temperatures (normalized by T_c') using the law of rectilinear diameters and critical exponents. The black dashed line shows the trend in γ *vs.* T_c^X/T_c which works for all proteins except for the two variants of TDP-43 as expected. (C): Surface tension of the poly-U-protein condensates for FUS, hnRNPA1 and hnRNPA1-A-LCD. Filled circles show the data obtained from simulations and solid lines denote the scaling fit given in Eq. (S9). The results for the surface tension of the RNA-protein mixtures are noisier than those of the pure condensates due to the proximity to the critical point, and thus, the estimation of their T_c is slightly less accurate. Dashed curves represent the extrapolation of the interfacial tension at low temperatures for each system.

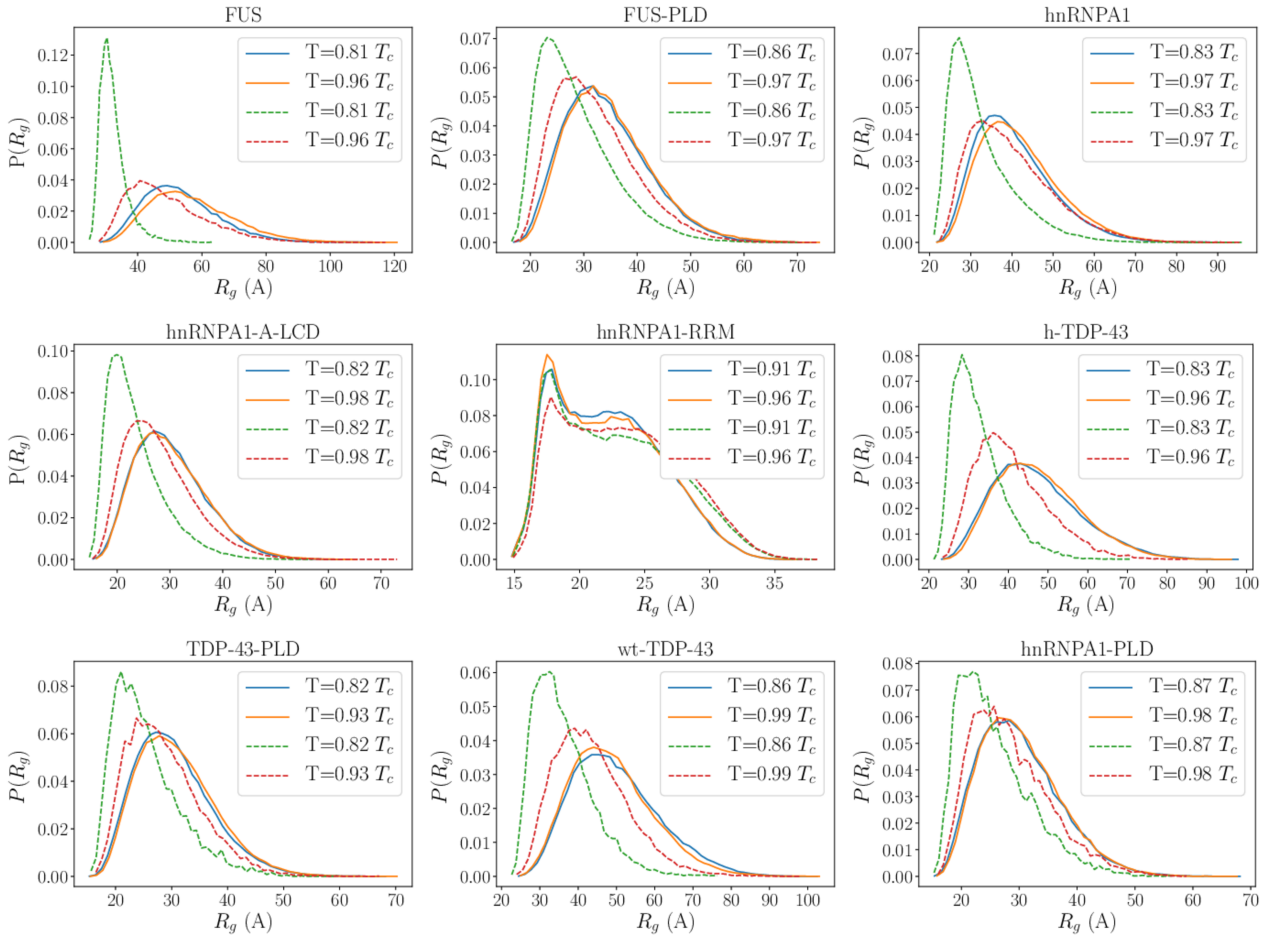


FIG. S6: Normalised radius of gyration histograms of the different studied proteins within the condensate (solid curves) and within the diluted phase (dashed lines) at the temperatures indicated in the legend. Note that T_c refers to the critical temperature of each system in pure component.

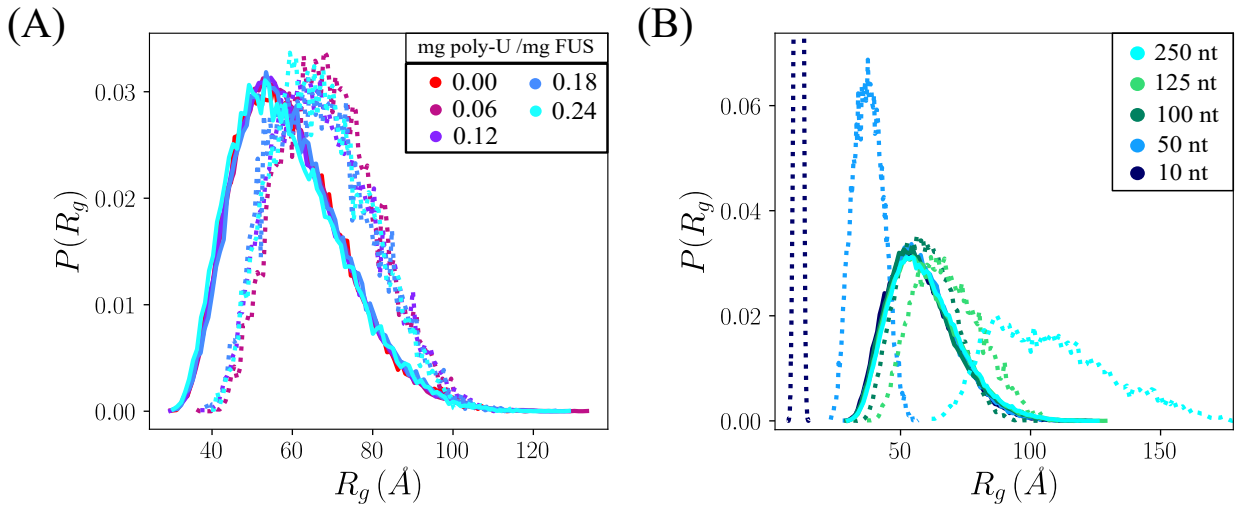


FIG. S7: (A) Normalised radius of gyration histograms of FUS (solid lines) and poly-U strands (dashed lines) within the condensate for different RNA-FUS mass ratios (as indicated in the legend) at $T \approx 0.97T_c$, where T_c refers to the critical temperature of FUS in absence of RNA. (B) Normalised radius of gyration histograms of FUS (solid lines) and poly-U strands (dashed lines) within the condensate for different poly-U strand lengths and for a constant poly-U/FUS mass ratio of 0.06. The temperature of the studied systems is $T \approx 0.97T_c$, where T_c is the critical temperature in absence of RNA.

SVI. PROTEIN/RNA CONTACT MAPS

The intermolecular contact maps between proteins inside the condensates are computed from DC trajectories. For all systems, the contacts were computed at $T/T_c \approx 0.9$ in absence of poly-U and at $T/T_c \approx 0.95$ in presence of poly-U, being T_c the critical temperature of each corresponding system (when poly-U is present, T_c refers to the critical temperature of the mixture at the considered poly-U concentration). Typically, molecular contacts are tracked through a distance criterion, and it is normally assumed that the contact map relative frequency (not absolute frequency) is in general independent of the chosen cut-off distance (for reasonable cut-off values) used in the calculations. However, in order to accurately compute the most relevant and frequent residue-residue contact pairs enabling LLPS, it is essential to specifically consider the actual parametrization of each amino acid in terms of excluded volume and minimum potential energy interacting distance. For that reason, we used a 'smart' sequence dependent cut-off distance equal to $1.2\sigma_{ij}$, where σ_{ij} accounts for the mean excluded volume of the specific ith and jth amino acids. Since the minimum of the employed potential is located at $2^{1/6}\sigma_{ij} \approx 1.122\sigma_{ij}$, we set our cut-off distance slightly beyond

that point, at $1.2\sigma_{ij}$, to ensure significant binding. Using this novel sequence-dependent cut-off scheme for each amino acid pair interaction, we can better exclude neighboring contacts that are coincidentally close to real interacting amino acids along the sequence that indeed are positively contributing to sustain LLPS.

In Figs. S8 and S12 we show the protein contact maps (averaged over all the equilibrium configurations and protein replicas) in absence (S8) *versus* presence of poly-U (at the poly-U/protein mass ratio that maximises droplet stability, see caption of Fig. S12 for the specific poly-U/protein mass fraction of each system). Given that the poly-U sequence is only composed by uridines, we also include side bars in each corresponding map of Fig. S12, which show all the contacts between uridines and the different protein amino acids. In Fig. S12 (B), we also plot the total number of intermolecular contacts between FUS (top) and hnRNPA1 (middle) (per domain) with poly-U averaged over all the equilibrium configurations and protein regions as labelled in the maps.

Moreover, the ten most frequent intermolecular protein contacts within the condensates (according to the force field [2, 3]) are provided in Figures S9-S11 for pure component droplets, and in Fig. S13 for droplets with poly-U. These figures include the top ten most repeated contacts evaluated with the sequence-dependent cut-off described above (panel (A)), and the ten most frequent contacts after renormalization by the relative abundance of each amino acid along the protein sequence (panel (B)). Furthermore, in panel (C), we provide the natural abundance of the different amino acids along each protein sequence. The top ten contacts provided in panel (A) indicates the number of intermolecular contacts between pairs of amino acids per protein and configuration (averaged over all of them). That magnitude over the number of amino acids in the protein sequence that are involved in every pairwise contact is shown in B. The same information but for the poly-U/protein mixtures described in Fig. S12 is provided in Fig. S13.

Finally, in Figure S14, we provide information of the number of contacts between FUS-FUS, FUS–poly-U and the total number of contacts of different FUS-RNA mixtures. In panel (A), the number of contacts are calculated for different poly-U concentrations and strand lengths of 125 nt. On the other hand, panel (B) shows the same analysis as a function of the strand length for 3 different cases (10 nt, 125 nt, and 250 nt) and at a constant concentration of 0.06 mg poly-U/mg FUS. All these results have been obtained from the simulations employed to compute viscosity and droplet diffusion.

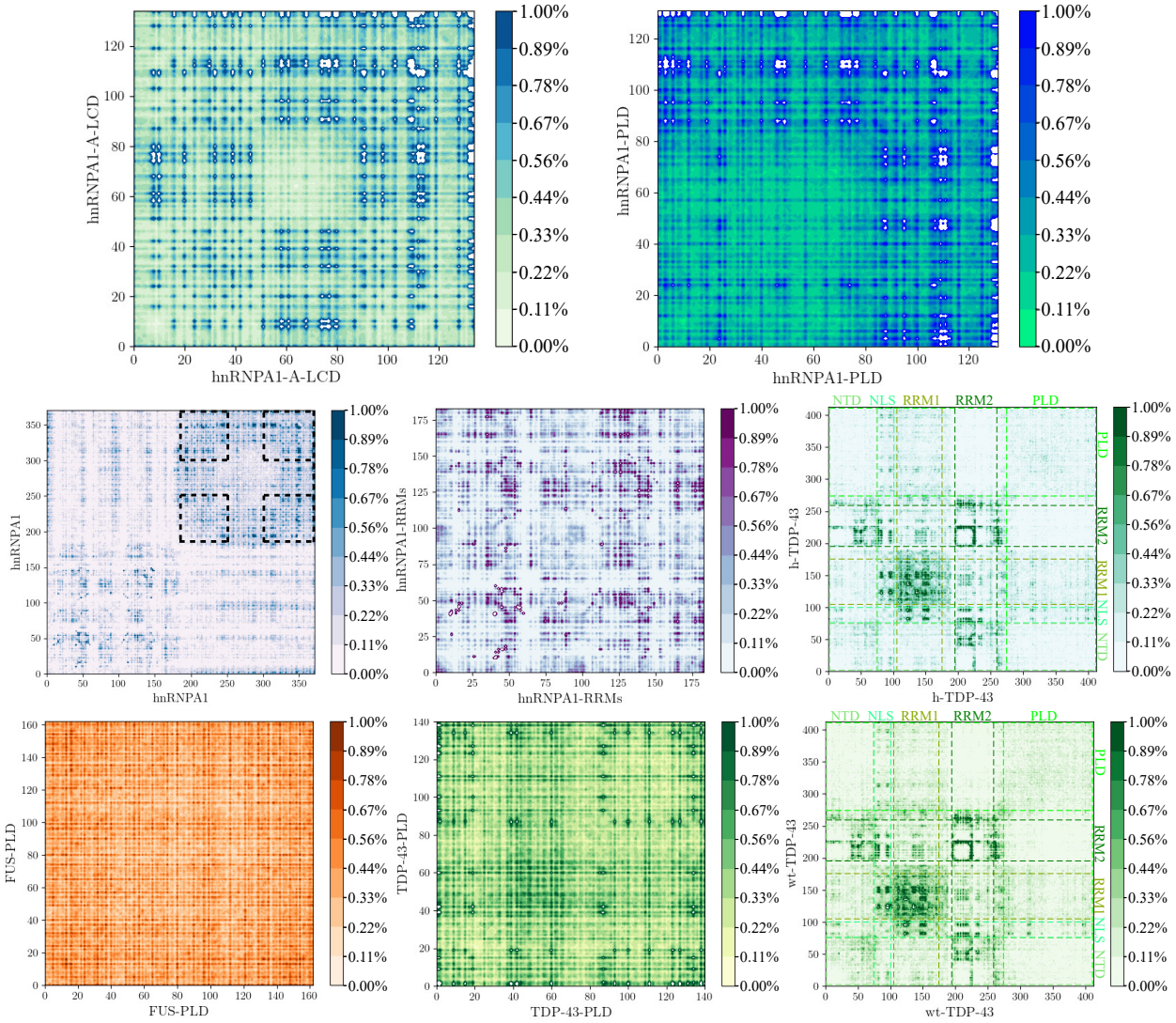


FIG. S8: Average number of intermolecular protein contacts per protein replica in percentage (where 100% would mean that all protein replicas in the condensate have a given contact at all times) for hnRNPA1-A-LCD, hnRNPA1-PLD, hnRNPA1, hnRNPA1-RRMs, FUS-PLD, TDP-43-PLD, wt-TDP-43 and h-TDP-43 condensates in pure component measured at $T/T_c = 0.95$, being T_c the corresponding critical temperature of each protein type (see Table S2). Dashed squares in the contact map of hnRNPA1 show the two corresponding sequences of hnRNPA1-A-LCD contained in hnRNPA1. The different protein domains in wt-TDP-43 and h-TDP-43 are indicated by dashed lines.

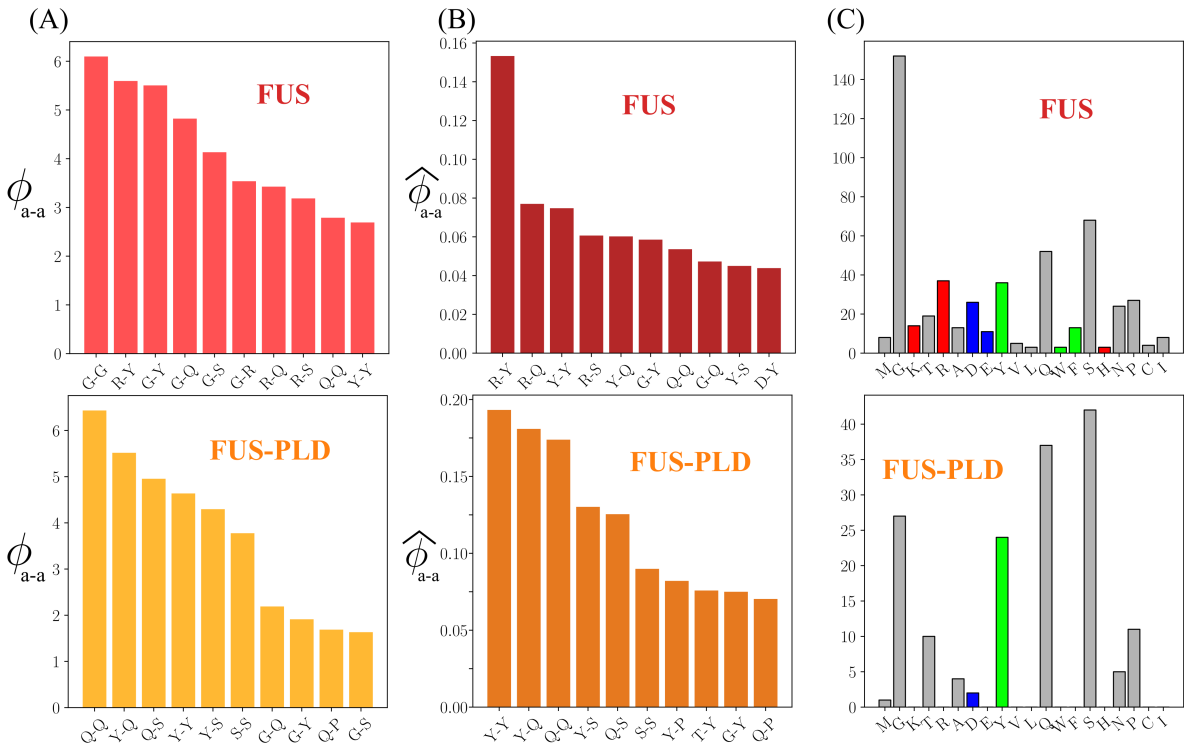


FIG. S9: (A) Average number of the ten most repeated intermolecular amino acid contacts per protein (ϕ_{a-a}) within FUS and FUS-PLD droplets (in absence of poly-U) at $T/T_c = 0.9$ (see Table S2 for the critical temperature T_c of each protein). (B) The same as in (A) but renormalized by the protein amino acid abundance ($\hat{\phi}_{a-a}$). Note that for every given pairwise contact interaction, we normalize by the sum of the amount of the two amino acids involved in the contact interaction divided by two. If the amino acid pair is between same type of amino acids, we then recover the natural abundance of that amino acid type in the sequence. (C) Abundance of each amino acid type in FUS and FUS-PLD sequences. The color code indicates positively charged (red), negatively charged (blue) and aromatic residues (green), while the rest of amino acids are labelled in grey.

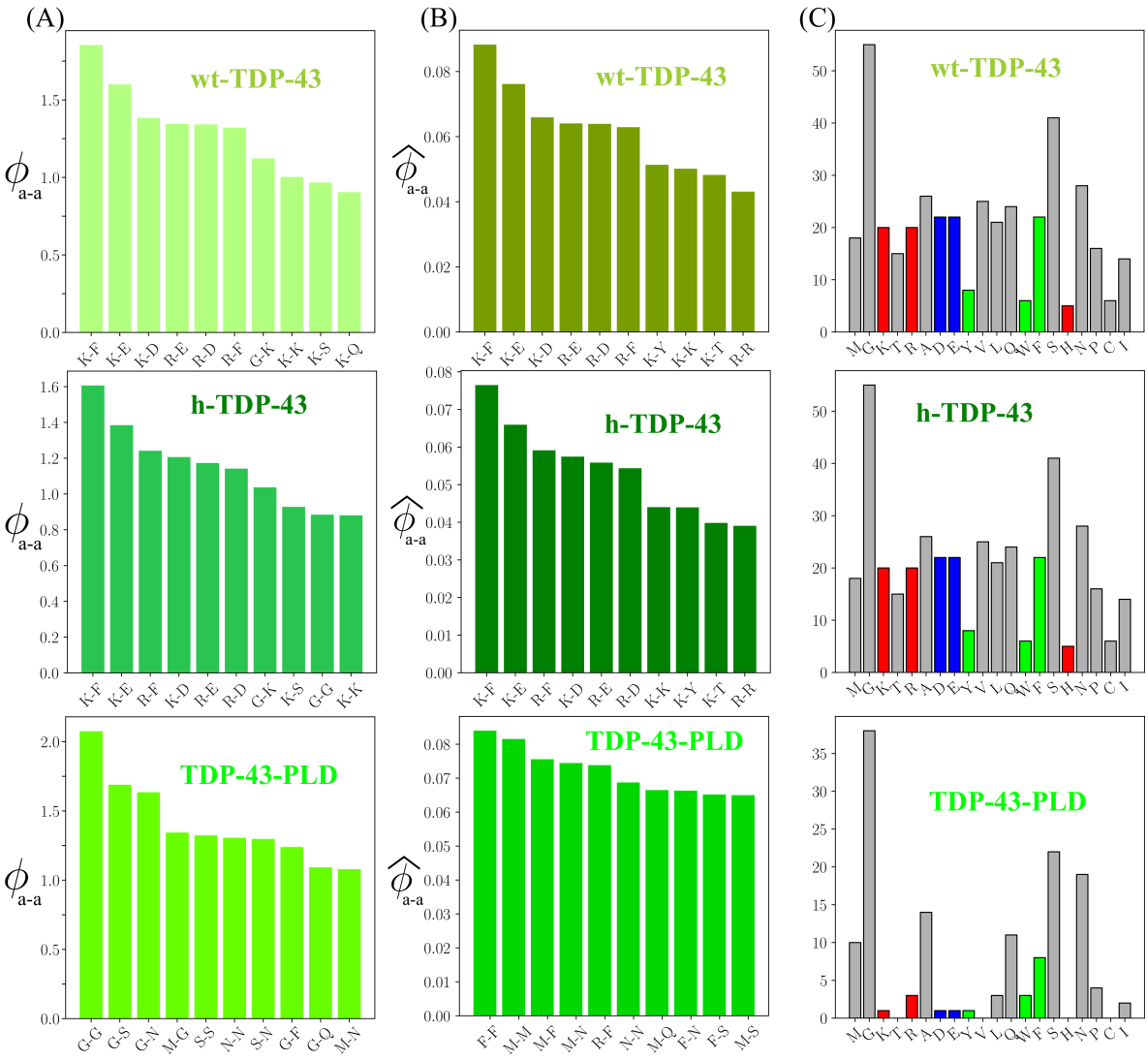


FIG. S10: (A) Average number of the ten most repeated intermolecular amino acid contacts per protein (ϕ_{a-a}) within wt-TDP-43, h-TDP-43 and TDP-43-PLD droplets (in absence of poly-U) at $T/T_c = 0.9$ (see Table S2 for the critical temperature T_c of each protein). (B) The same as in (A) but renormalized by the protein amino acid abundance ($\hat{\phi}_{a-a}$). Note that for every given pairwise contact interaction, we divide by the sum of the amount of the two amino acids involved in the contact interaction divided by two. If the amino acid pair is between the same type of amino acids, we then recover the natural abundance of that amino acid type in the sequence. (C) Abundance of each amino acid type in the sequences of wt-TDP-43, h-TDP-43 and TDP-43-PLD. The color code indicates positively charged (red), negatively charged (blue) and aromatic residues (green), while the rest are labelled in grey.

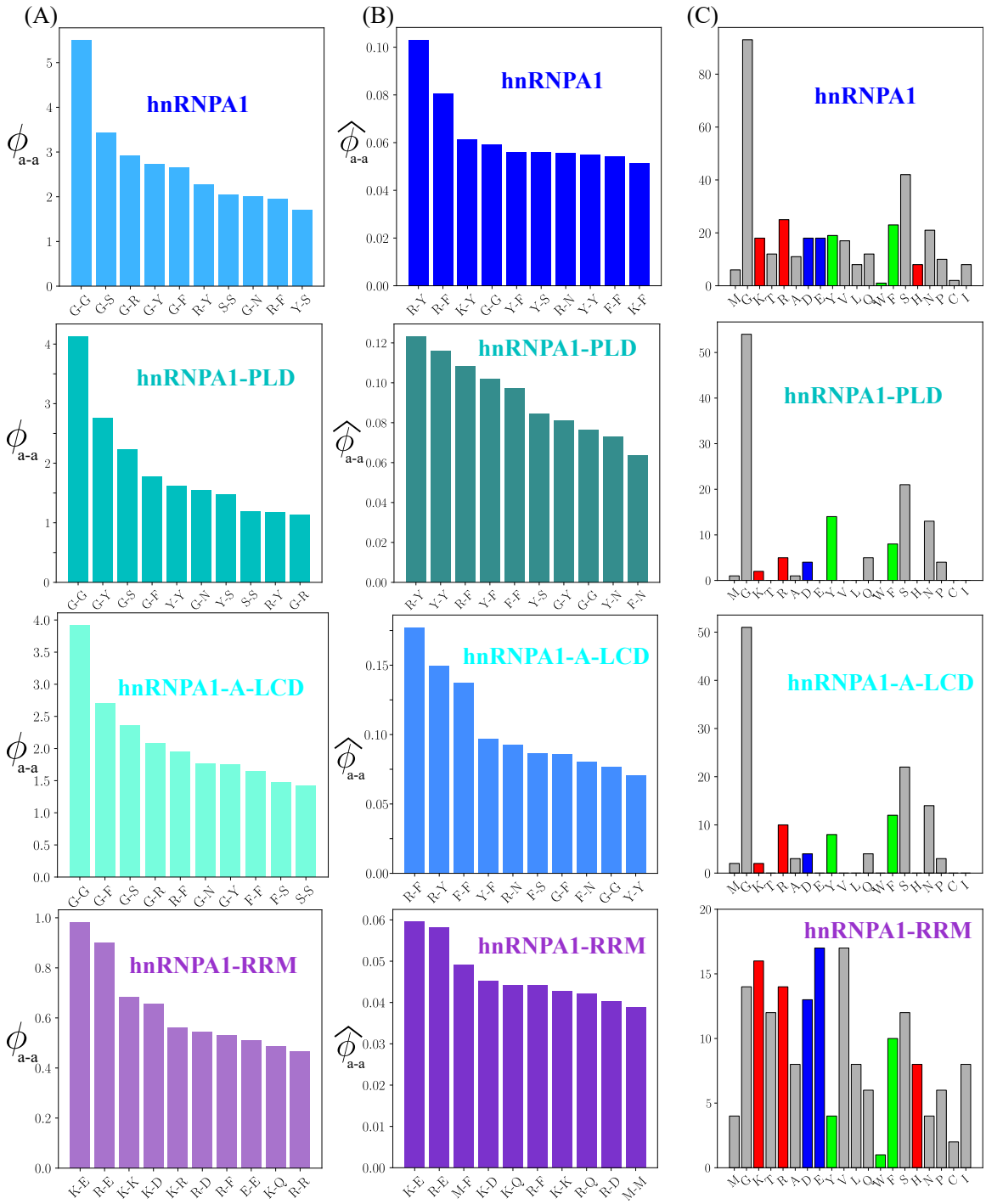


FIG. S11: (A) Average number of the ten most repeated intermolecular amino acid contacts per protein (ϕ_{a-a}) within hnRNPA1, hnRNPA1-PLD, hnRNPA1-A-LCD and hnRNPA1-RRM droplets (in absence of poly-U) at $T/T_c = 0.9$ (see Table S2 for the critical temperature of each protein). (B) The same as in (A) but renormalized by the protein amino acid abundance ($\hat{\phi}_{a-a}$). Note that for every given pairwise contact interaction, we divide by the sum of the amount of the two amino acids involved in the contact pair divided by two. If the amino acid pair is between the same type of amino acids, we then recover the natural abundance of that amino acid type in the sequence. (C) Abundance of each amino acid type in the sequences of hnRNPA1, hnRNPA1-PLD, hnRNPA1-RRM and hnRNPA1-A-LCD proteins. The color code indicates positively charged (red), negatively charged (blue) and aromatic residues (green), while the rest are labelled in grey.

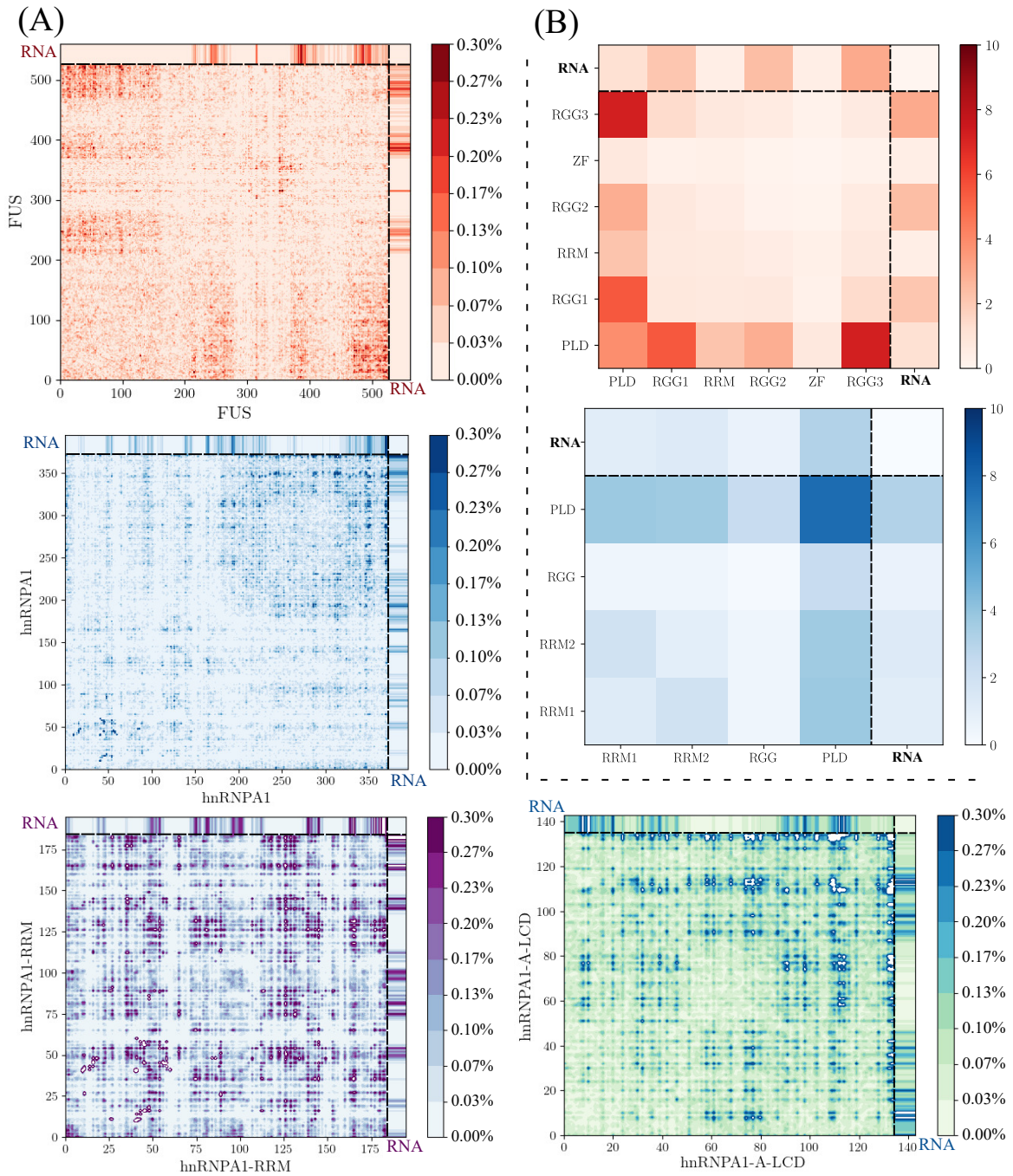


FIG. S12: (A) Average number of intermolecular contacts per protein in percentage (where 100% would mean that all protein replicas have a given contact at all times) for FUS, hnRNPA1, hnRNPA1-A-LCD and hnRNPA1-RRM condensates in presence of poly-U(250) at $T/T_c = 0.95$ and at the coexisting droplet equilibrium density at such T (see Table S3 for the corresponding critical temperatures of each system). The number of poly-U(250nt) chains are 2 in all systems except for hnRNPA1-RRM where is 1 poly-U(250nt) strand. The corresponding poly-U/protein mass fractions of each system are the following: FUS (0.119), hnRNPA1 (0.099), hnRNPA1-A-LCD (0.117) and hnRNPA1-RRM (0.046). The contacts with poly-U are also included in the upper and right side edges of the maps. (B) Number of contacts per protein domain replica in FUS and hnRNPA1 condensates in presence of poly-U at the same conditions described in (A). The bar indicates the average number of contacts that a protein domain has at all times for each replica. The protein contact domains with poly-U are also given in the edges of the maps.

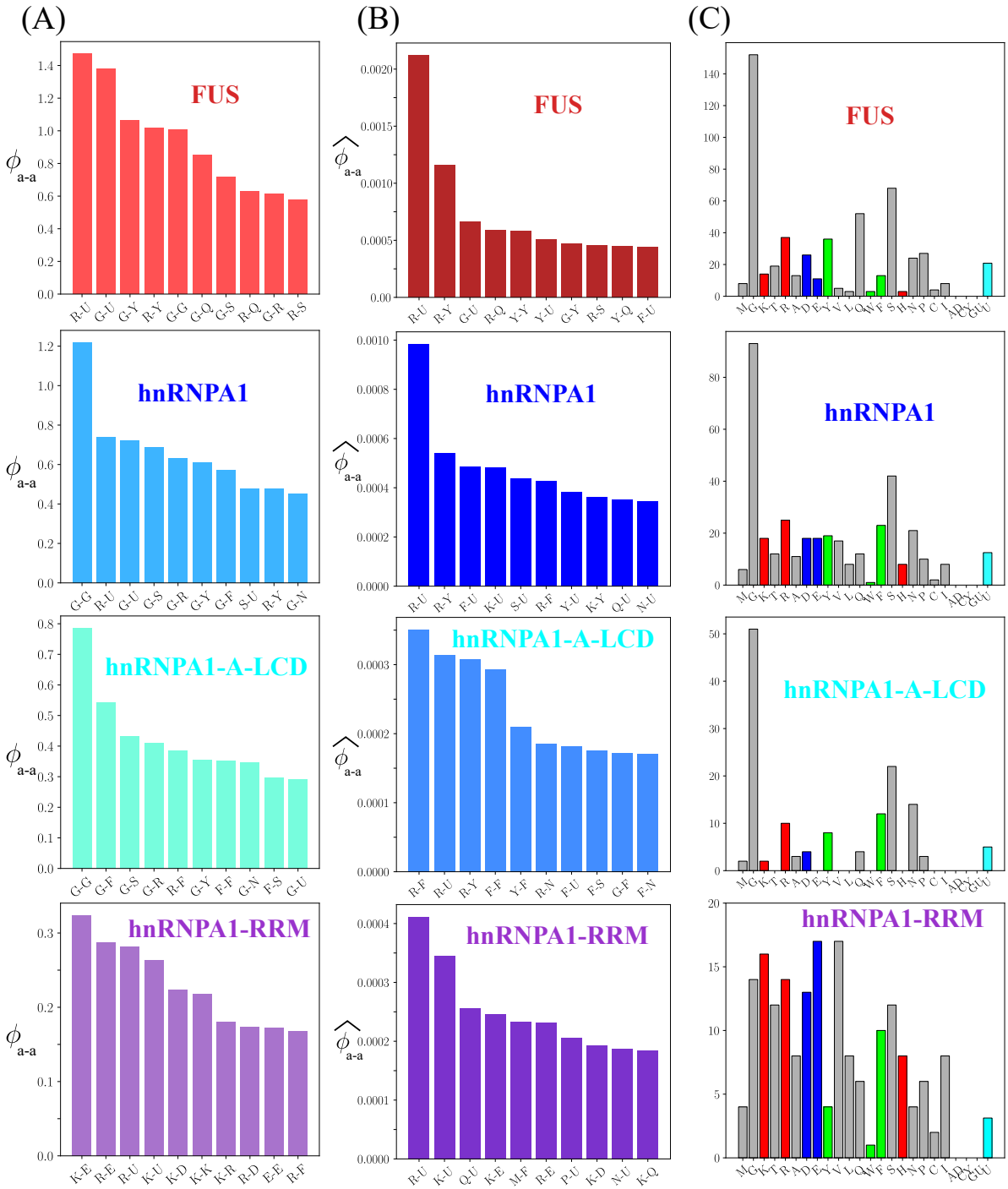


FIG. S13: (A) Average number of the ten most repeated intermolecular contacts per protein (ϕ_{a-a}) in FUS, hnRNPA1, hnRNPA1-A-LCD and hnRNPA1-RRM poly-U condensates at the same concentration and conditions described in Fig. S12. (B) The same as in (A) but renormalized by the protein amino acid and poly-U abundance ($\hat{\phi}_{a-a}$). Note that for every given pairwise contact interaction, we divide by the sum of the amount of the two residues/nucleotides involved in the contact interaction divided by two. If the amino acid pair is between the same type of amino acids, we then recover the natural abundance of that amino acid type in the sequence. (C) Abundance of each amino acid (and nucleotide) per protein replica in the system. Red bars indicate the positively charged amino acids, blue color depict negatively charged amino acids, green accounts for aromatic residues and grey for the rest of amino acids. The number of uridines per protein replica is indicated in cyan.

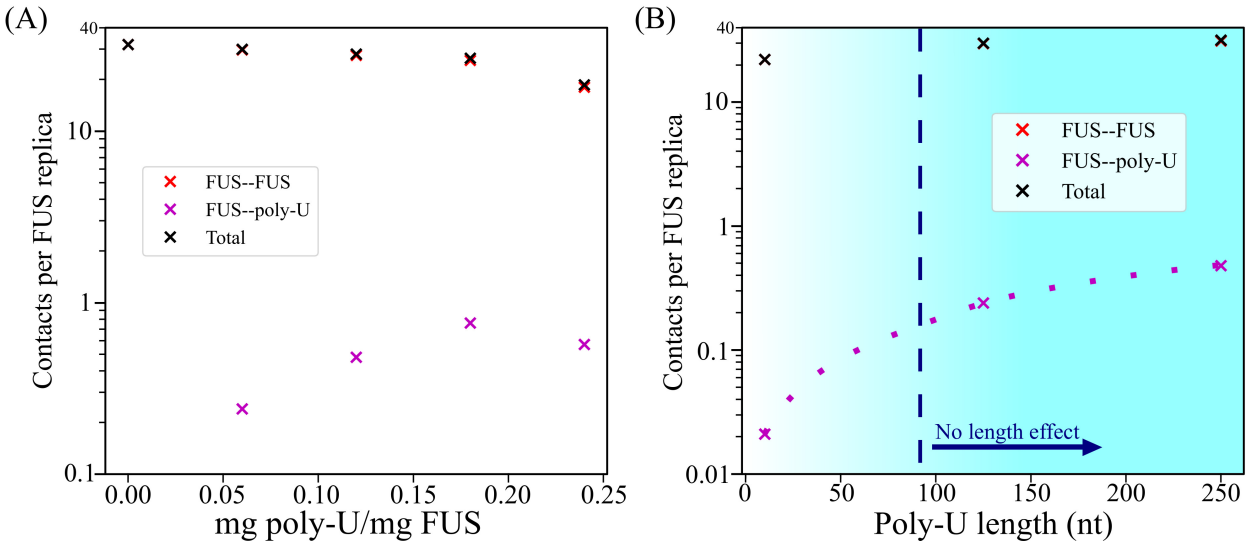


FIG. S14: (A) Number of contacts per FUS protein measured in the condensate at $T/T_c = 0.97$ for different concentrations of poly-U with 125 nt chain length. We include FUS–FUS, FUS–poly-U and the total number of contacts. Poly-U–poly-U contacts have been omitted since they are close to zero due to repulsive electrostatic interactions. (B) Number of contacts measured in the condensate at $T/T_c = 0.97$ for different lengths of poly-U strands and for a constant poly-U/FUS mass ratio of ~ 0.06. We include FUS–FUS, FUS–poly-U and the total number of contacts. The limit of the RNA critical length adopted from the obtained critical temperatures in Fig. 4 of the main manuscript is depicted by a dashed vertical line. Please note that red symbols representing FUS-FUS contacts overlap for almost every state with the total number of contacts (black symbols).

SVII. CALCULATION OF TRANSPORT PROPERTIES WITHIN THE CONDENSATES

We evaluate droplet viscosity and protein diffusion inside the condensates just below the pure component critical temperature from absence to moderately high poly-U concentration for FUS, hnRNPA1, and hnRNPA1-A-LCD proteins. Furthermore, we characterize the viscoelastic properties of wt-TDP-43 condensates without poly-U. We perform NVT simulations in a cubic box at the equilibrium bulk droplet density corresponding to the temperature and poly-U concentration of each system, taken from the phase diagram. We prepare the initial configuration of these systems by compressing the long side of the DC simulation box until obtaining a cubic box, and then the system is relaxed in the NpT ensemble until the bulk equilibrium droplet density is reached. Then, systems are further equilibrated for ∼100 nanoseconds in the NVT ensemble and, finally, production runs entail from 3 to 6 microseconds depending on the system (around 6 microseconds were simulated for each poly-U/FUS

mixture, and almost 5 for each poly-U/hnRNPA1 and poly-U/hnRNPA1-A-LCD condensates).

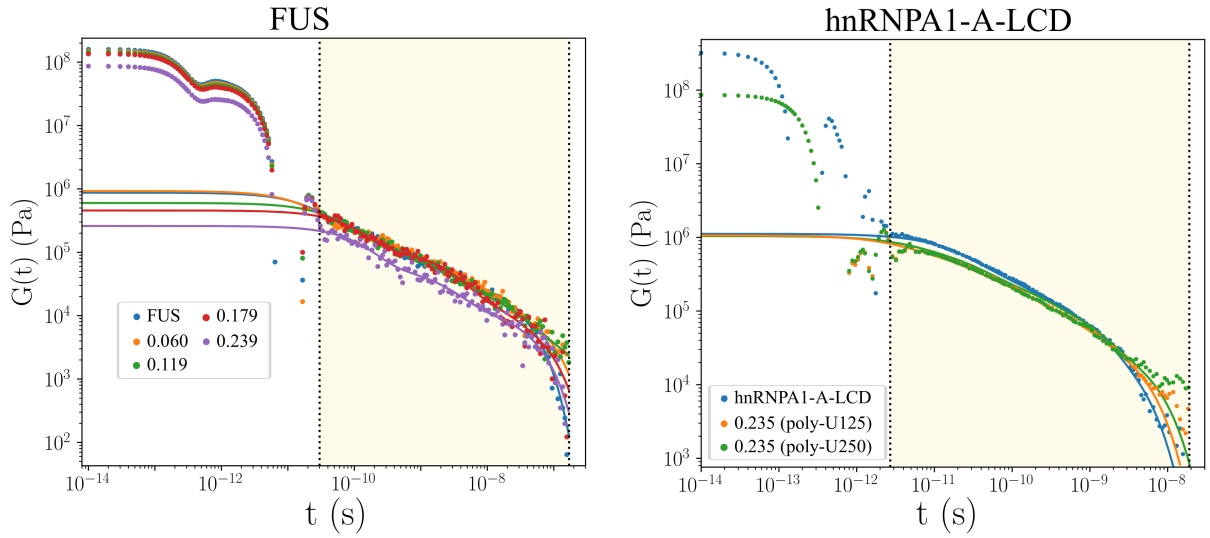


FIG. S15: Shear stress relaxation modulus of FUS and hnRNPA1-A-LCD droplets at different poly-U/protein mass fractions as indicated in the legend. Circles show the computed data from NVT simulations at $T/T_c = 0.97$ for FUS mixtures and $T/T_c = 0.98$ for hnRNPA1-A-LCD mixtures (where T_c refers to the critical temperature of the pure condensates of each protein type) and at the corresponding equilibrium density of each droplet at such conditions. The shadowed regime (light yellow) indicates the region in which the Maxwell modes fit is applied (solid lines). Note that these fits are only applied to compute the contribution to the total viscosity along the shadowed regime. While for hnRNPA1-A-LCD, we plot $G(t)$ for a given RNA concentration using strands of 125 and 250 nt, as well as for the pure protein system, for FUS, all the different concentrations were achieved by adding 125 nt poly-U strands.

From NVT simulations, we can compute both viscosity and protein diffusion in the condensate in separate ways. The shear viscosity can be straightforwardly calculated by integrating the relaxation modulus in time (see Chapter 7 of the book [19]):

$$\eta = \int_0^\infty dt\, G(t) \tag{S10}$$

In an isotropic system, we can compute the shear relaxation modulus $G(t)$ more accurately by using all the components of the pressure tensor ($\sigma_{\alpha\beta}$) as shown in Ref. [20]:

$$\begin{aligned} G(t) = &\frac{V}{5k_BT}\left[\langle\sigma_{xy}(0)\sigma_{xy}(t)\rangle + \langle\sigma_{xz}(0)\sigma_{xz}(t)\rangle + \langle\sigma_{yz}(0)\sigma_{yz}(t)\rangle\right] \\ &+ \frac{V}{30k_BT}\left[\langle N_{xy}(0)N_{xy}(t)\rangle + \langle N_{xz}(0)N_{xz}(t)\rangle + \langle N_{yz}(0)N_{yz}(t)\rangle\right], \end{aligned} \tag{S11}$$

where $N_{\alpha\beta} = \sigma_{\alpha\alpha} - \sigma_{\beta\beta}$ is the first normal stress difference. This correlation can be easily computed by using the compute ave/correlate/long in the USER-MISC package of LAMMPS [1]. In all cases, the relaxation modulus presents an initial regime that mainly accounts for the protein/poly-U intramolecular interactions, followed by a terminal region which corresponds to much slower relaxation modes, as those coming from intermolecular interactions and the relaxation of the protein and RNA conformations. Due to the very wide range time-scale involved in the calculation and the noisy nature of the relaxation modulus in the terminal region obtained in the simulations, we follow a particular strategy to calculate our estimate of viscosity. At short times, $G(t)$ is smooth and the integral can be computed using numerical integration (trapezoidal rule). However, at longer times $G(t)$ presents more noise, and hence, we calculate the integral in that regime by first fitting $G(t)$ to a series of Maxwell modes ($G_i \exp(-t/\tau)$) equidistant in logarithmic time [21] and then calculating the integral analytically. The fit to Maxwell modes is carried out with the help of the open-source RepTate software [22]. Finally, viscosity is obtained by adding the two terms:

$$\eta = \eta(t_0) + \int_{t_0}^{\infty} dt\, G_M(t), \tag{S12}$$

where $\eta(t_0)$ corresponds to the computed term for short time-scales, $G_M(t)$ is the part evaluated via the Maxwell mode fit at long time-scales, and t_0 is the time that separates both regimes. In Fig. S15 we plot the stress relaxation function $G(t)$ for FUS and hnRNPA1-A-LCD for different poly-U concentrations and lengths. We only fit $G(t)$ in the shadowed regime, and then we integrate it as explained in Eq. (S12). The time t_0 is defined by the left dotted vertical line. Note that, due to the finite size of the simulation box and the finite length of the run, the relaxation modulus $G(t)$ shown in Fig. S15 is noisy in the region closer to the terminal time (the time where the modulus decays exponentially to zero). The reported value of the viscosity is the result of the Maxwell mode fit to the noisy $G(t)$ domain, and hence, has some level of uncertainty. The error bars shown in Fig. 5 of the main paper have been estimated from the error of the Maxwell mode fits to the value of $G(t)$ obtained in our simulations.

In Tables S4 and S5, we provide the values of the viscosity measured as a result of the contribution at short times and the integral of the Maxwell mode fit. This results are plotted in Fig. 5 of the main text, but here we also provide the values for the pure systems and

poly-U/protein mixtures with RNA strands of 10, 125 and 250 nt.

Protein \ poly-U	No RNA	1x250 (nt)	2x250 (nt)	3x250 (nt)	4x250 (nt)
FUS	1.27×10^{-3}	1.80×10^{-3}	2.30×10^{-3}	2.60×10^{-3}	2.30×10^{-3}
hnRNPA1	7.5×10^{-4}	7.0×10^{-4}	6.8×10^{-4}	8.9×10^{-4}	8.9×10^{-4}
hnRNPA1-A-LCD	2.92×10^{-4}	3.12×10^{-4}	3.50×10^{-4}	2.90×10^{-4}	3.16×10^{-4}

TABLE S4: Viscosity ($Pa \cdot s$) of FUS (at $T/T_c = 0.97$), hnRNPA1 (at $T/T_c = 0.985$) and hnRNPA1-A-LCD (at $T/T_c = 0.98$) condensates as a function of poly-U concentration, given in number of added poly-U strands of 250 nucleotides. The number of proteins in each system is that given in Table S1, and is constant for all concentrations. The equivalence in poly-U/protein mass ratio can be extracted from Fig. S4

Protein \ poly-U	2x125 (nt)	4x125 (nt)	6x125 (nt)	8x125 (nt)	26x10 (nt)
FUS	1.62×10^{-3}	1.92×10^{-3}	1.22×10^{-3}	0.97×10^{-3}	0.07×10^{-3}
hnRNPA1	5.5×10^{-4}	5.9×10^{-4}	5.4×10^{-4}	4.3×10^{-4}	-
hnRNPA1-A-LCD	3.33×10^{-4}	3.00×10^{-3}	2.90×10^{-3}	2.68×10^{-3}	-

TABLE S5: Viscosity ($Pa \cdot s$) of FUS (at $T/T_c = 0.97$), hnRNPA1 (at $T/T_c = 0.985$) and hnRNPA1-A-LCD (at $T/T_c = 0.98$) condensates as a function of poly-U concentration, given in number of poly-U strands of 125 nucleotides and 10 nucleotides for the last column. The number of proteins in each system is that given in Table S1, and is constant for all concentrations. The equivalence in poly-U/protein mass ratio is provided in Fig. S4

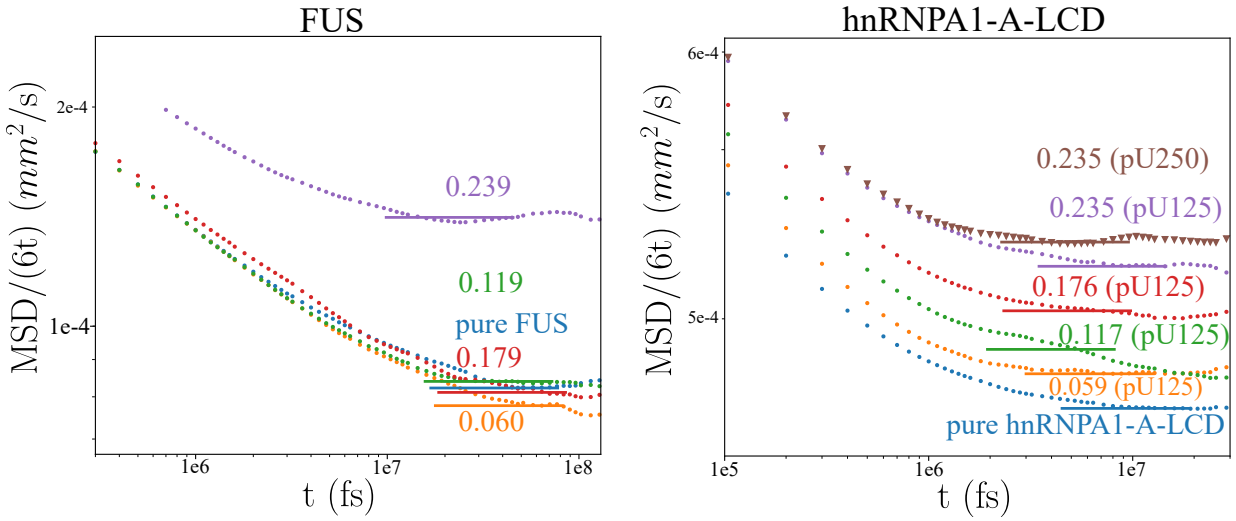


FIG. S16: Diffusion coefficients (MSD/6t) of FUS and hnRNPA1-A-LCD proteins inside the condensates at different poly-U/protein mass fractions as indicated in the different curves. The data shown for FUS correspond to poly-U strands of 125 nucleotides length, while for hnRNPA1-A-LCD condensates, strands of 125 and 250 nt were introduced as specified in each curve. Simulations were performed in the NVT ensemble at $T/T_c = 0.97$ for FUS, and at $T/T_c = 0.98$ for hnRNPA1-A-LCD proteins (where T_c refers the critical temperature of each pure condensate) and at the corresponding equilibrium density of each droplet at such conditions. The horizontal solid lines depict where the diffusive regime starts, and thus, where diffusion coefficients can be measured.

The protein diffusion coefficient inside the condensates is obtained through the mean squared displacement (MSD) of the proteins center of mass. After a subdiffusive regime (i.e., ∼1 molecular diameter), proteins exhibit a diffusive behavior and then the MSD of the center of mass can be measured via:

$$\left\langle (\mathbf{R}_{\mathrm{CM}}(t) - \mathbf{R}_{\mathrm{CM}}(0))^2 \right\rangle = 6 D_c t, \tag{S13}$$

where R_{CM} indicates the center of mass of a given protein at different times, and D_c accounts for the diffusion coefficient. In order to get an accurate estimate of the MSD, the same correlator technique employed in the calculation of the relaxation modulus has been used [20]. By plotting the MSD divided by $6t$ (as shown in Fig. S16), the function shows a plateau at long times that provides the value of the diffusion coefficient of the proteins (D_c). In contrast to viscosity, protein diffusion coefficients barely depend on the length of the added poly-U strands, but depend on the induced droplet density by each poly-U

concentration and length.

[1] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, Journal of computational physics **117**, 1 (1995).

[2] S. Das, Y.-H. Lin, R. M. Vernon, J. D. Forman-Kay, and H. S. Chan, Comparative roles of charge, π, and hydrophobic interactions in sequence-dependent phase separation of intrinsically disordered proteins, Proceedings of the National Academy of Sciences **117**, 28795 (2020).

[3] G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, Sequence determinants of protein phase behavior from a coarse-grained model, PLoS computational biology **14**, e1005941 (2018).

[4] R. M. Regy, G. L. Dignon, W. Zheng, Y. C. Kim, and J. Mittal, Sequence dependent phase separation of protein-polynucleotide mixtures elucidated using molecular simulations, Nucleic Acids Research **48**, 12593 (2020).

[5] G. Krainer, T. J. Welsh, J. A. Joseph, J. R. Espinosa, S. Wittmann, E. de Csilléry, A. Sridhar, Z. Toprakcioglu, G. Gudiškytė, M. A. Czekalska, *et al.*, Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions, Nature Communications **12**, 1 (2021).

[6] H. S. Ashbaugh and H. W. Hatch, Natively unfolded protein stability as a coil-to-globule transition in charge/hydropathy space, *Journal of the American Chemical Society*, Journal of the American Chemical Society **130**, 9536 (2008).

[7] L. H. Kapcha and P. J. Rossky, A simple atomic-level hydrophobicity scale reveals protein interfacial structure, Journal of molecular biology **426**, 484 (2014).

[8] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, The Journal of chemical physics **81**, 511 (1984).

[9] T. Schneider and E. Stoll, Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions, Physical Review B **17**, 1302 (1978).

[10] G. L. Dignon, W. Zheng, Y. C. Kim, and J. Mittal, Temperature-controlled liquid–liquid phase separation of disordered proteins, ACS central science **5**, 821 (2019).

[11] J. S. Rowlinson and B. Widom, *Molecular theory of capillarity* (Courier Corporation, 2013).

[12] W. Humphrey, A. Dalke, and K. Schulten, Vmd: visual molecular dynamics, Journal of molec-

ular graphics **14**, 33 (1996).

[13] A. Ladd and L. Woodcock, Triple-point coexistence properties of the lennard-jones system, Chemical Physics Letters **51**, 155 (1977).

[14] R. García Fernández, J. L. Abascal, and C. Vega, The melting point of ice i h for common water models calculated from direct coexistence of the solid-liquid interface, The Journal of chemical physics **124**, 144506 (2006).

[15] F. J. Blas, L. G. MacDowell, E. de Miguel, and G. Jackson, Vapor-liquid interfacial properties of fully flexible lennard-jones chains, The Journal of chemical physics **129**, 144703 (2008).

[16] J. R. Espinosa, E. Sanz, C. Valeriani, and C. Vega, On fluid-solid direct coexistence simulations: The pseudo-hard sphere model, The Journal of chemical physics **139**, 144502 (2013).

[17] J. Walton, D. Tildesley, J. Rowlinson, and J. Henderson, The pressure tensor at the planar surface of a liquid, Molecular physics **48**, 1357 (1983).

[18] J. M. Prausnitz, R. N. Lichtenthaler, and E. G. De Azevedo, *Molecular thermodynamics of fluid-phase equilibria* (Pearson Education, 1998).

[19] M. Rubinstein, R. H. Colby, *et al.*, *Polymer physics*, Vol. 23 (Oxford university press New York, 2003).

[20] J. Ramírez, S. K. Sukumaran, B. Vorselaars, and A. E. Likhtman, Efficient on the fly calculation of time correlation functions in computer simulations, The Journal of chemical physics **133**, 154103 (2010).

[21] A. E. Likhtman, Single-Chain Slip-Link Model of Entangled Polymers:˜Simultaneous Description of Neutron Spin-Echo, Rheology, and Diffusion, Macromolecules **38**, 6128 (2005).

[22] V. A. Boudara, D. J. Read, and J. Ramírez, Reptate rheology software: Toolkit for the analysis of theories and experiments, Journal of Rheology **64**, 709 (2020).