Emergence of order in condensates composed of multi-valent, multi-domain proteins

Srivastav Ranganathan and Eugene Shakhnovich
Department of Chemistry and Chemical Biology,
Harvard University, Cambridge, MA 02138
Email: shakhnovich@chemistry.harvard.edu

Abstract.
Many RNA-binding proteins (RBPs) that assemble into membraneless organelles, have a common architecture including disordered prion-like domain (PLD) and folded RNA-binding domain (RBD). An enrichment of PLD within the condensed phase gives rise to formation, on longer time scales, amyloid-like fibrils (aging). In this study, we employ coarse-grained Langevin dynamics simulations to explore the physical basis for the structural diversity in condensed phases of multi-domain RBPs. We discovered a highly cooperative first order transition between disordered structures and an ordered phase whereby chains of PLD organize in fibrils with high nematic orientational order. An interplay between homo-domain (PLD-PLD) and hetero-domain (PLD-RBD) interactions results in variety of structures with distinct spatial architectures. Interestingly, the different structural phases also exhibit vastly different intra-cluster dynamics of proteins, with diffusion coefficients 5 (disordered structures) to 50 times (ordered structures) lower than that of the dilute phase. Cooperativity of this liquid-solid transition makes fibril formation highly malleable to mutations or post-translational modifications. Our results provide a mechanistic understanding of how multi-domain RBPs could form assemblies with distinct structural and material properties.

Significance Statement.
Assembly of proteins and nucleic acids into dense, liquid-like pockets is associated with several key functions including stress response, gene-regulation, DNA-repair and RNA processing. Several RNA binding proteins such as FUS are known to form liquid-like condensates that progressively harden into more dynamically, solid-like structures, a phenomenon that gets accelerated by disease mutations. In this study, we discover the mechanistic origins of this transition and show that small mutational or posttranslational modifications could result in sharp disorder-order transitions that could characterize accelerated liquid-solid transition in disease mutants.

1 Introduction

Spatial seggregation of biomolecules is a vital mechanism by which cellular machinery and biochemical reactions get spatio-temporally localized [1,2]. While classical examples of such spatiotemporal seggregation involves physical compartments bound by a membrane, more recent evidence points towards the presence of biomolecule-rich compartments that form and disassemble in response to external cues[3–5]. These structures, known as membraneless-organelles (MLOs) have been attributed to performing diverse cellular functions, ranging from cell division,

ribosome synthesis to stress response[3,6–9]. Despite the diversity in function, the physical mechanisms resulting in their formation shows remarkable similarity. The hallmark of these structures is the ability to assemble into droplet-like assemblies via a mechanism called as liquid-liquid phase separation (LLPS)[9–11].

A key class of membraneless assemblies are ribonucleoprotein granules (RNPs) which are a heterogenous mix of RNA-binding proteins (RBPs) and RNA [12–14]. Proteins that get localized within RNP granules are often RBPs rich in low-complexity and intrinsically disordered domains[3,14–16]. Further, RBPs typically display a multi-domain structure, often with a low complexity region (also known as the prion-like domain, PLD) as well as an RNA-binding domain (RBD) [14]. Such a multi-domain architechture opens up the possibility of diverse interaction networks stabilizing the condensed phase. For instance, in their seminal work involving the FUS family of proteins, Wang et al, have demonstrated how droplets formation could be driven by more hydrophobic contacts between the prion-like domain or through pi-pi and pi-cation interactions between the PLD and the RBD[14]. The multi-domain architechture and diversity in interactions not only results in altered threshold concentrations for LLPS but also structures with different material or dynamic properties[14]. The PLD of FUS is known to form ordered amyloid-fibrillar structures[17]. On the other hand, while full length FUS proteins phase separate into spherical droplets with liquid-like properties initially[18], experimental evidence suggests that these droplets could progressively reorganize and assume more filamentous geometries which are dynamically solid-like[19]. This liquid-solid transition gains further importance in disease biology where mutations to the FUS LC domain are known to accelerate the process[20].

The scaffolding role of FUS proteins and its ubiquitous nature in several RNP condensates makes it an interesting model RBP for study. Several experimental studies have studied the ability of FUS, FUS-PLD and FUS variants to spontaneously self-assemble into liquid-like droplets in vitro[14,17,21–23]. Structural elucidation of FUS assemblies using solution and solid-state NMR studies have revealed that the FUS PLD can assume a disordered configuration (in liquid-like FUS droplets)[18] or an ordered solid-like amyloid cross-β fold in FUS hydrogels[17]. The remarkable malleability of FUS self-assembly suggests that these proteins can potentially get involved in different inter-molecular interaction networks, each resulting in distinct material properties.

Biomolecular simulations of different resolutions have previously been employed to study the thermodynamics[21,24] and dynamic properties (droplet surface tension, shear viscosities)[21] of phase separation by FUS. Simulations have also been employed to identify the key inter-molecular interactions stabilizing the condensed phase in FUS droplets[25]. A recent computational study by Garaizar et al. demonstrates how an increase in FUS PLD interaction strength could result in spatially heterogenous condensates[26].

While the thermodynamics of FUS phase separation has been extensively characterized using computational models, the factors influencing the morphology of the assemblies have not been explored in previous studies. Further, while most of the earlier computational models can

capture the spherical droplet phase in FUS, the 'fibril-like' ordered state hasn't been captured in simulations. In this study, we employ a coarse-grained (CG) representation of the full length FUS protein to understand the physical mechanisms underlying emergence of order in multi-domain protein assemblies such as those constituted by RBPs. Using this CG representation, we explore how the variation of homo-domain (PLD-PLD) and hetero-domain (PLD-RBD) interaction strengths within a narrow, biologically relevant regime (< 1 kT) could give rise to self-assembled structures with diverse morphologies. We scan a broad range of conditions to establish a relationship between interaction networks, phosphorylation, intra-cluster density and structural order within the condensed phase. Our simulations show that biomolecular assemblies could transition from a disordered to ordered phase by switching between the predominant interactions that stabilize the condensed phase. More generally we observe the emergence of an ordered phase in a first-order like phase transition where loss of entropy due to local ordering of individual polymer chains in fibrils is compensated by favorable interaction energy in a more dense ordered phase. Our simulations also show that the disorder-order structural transitions are also accompanied by dramatically altered intra-cluster dynamics of polymers making them biologically relevant.

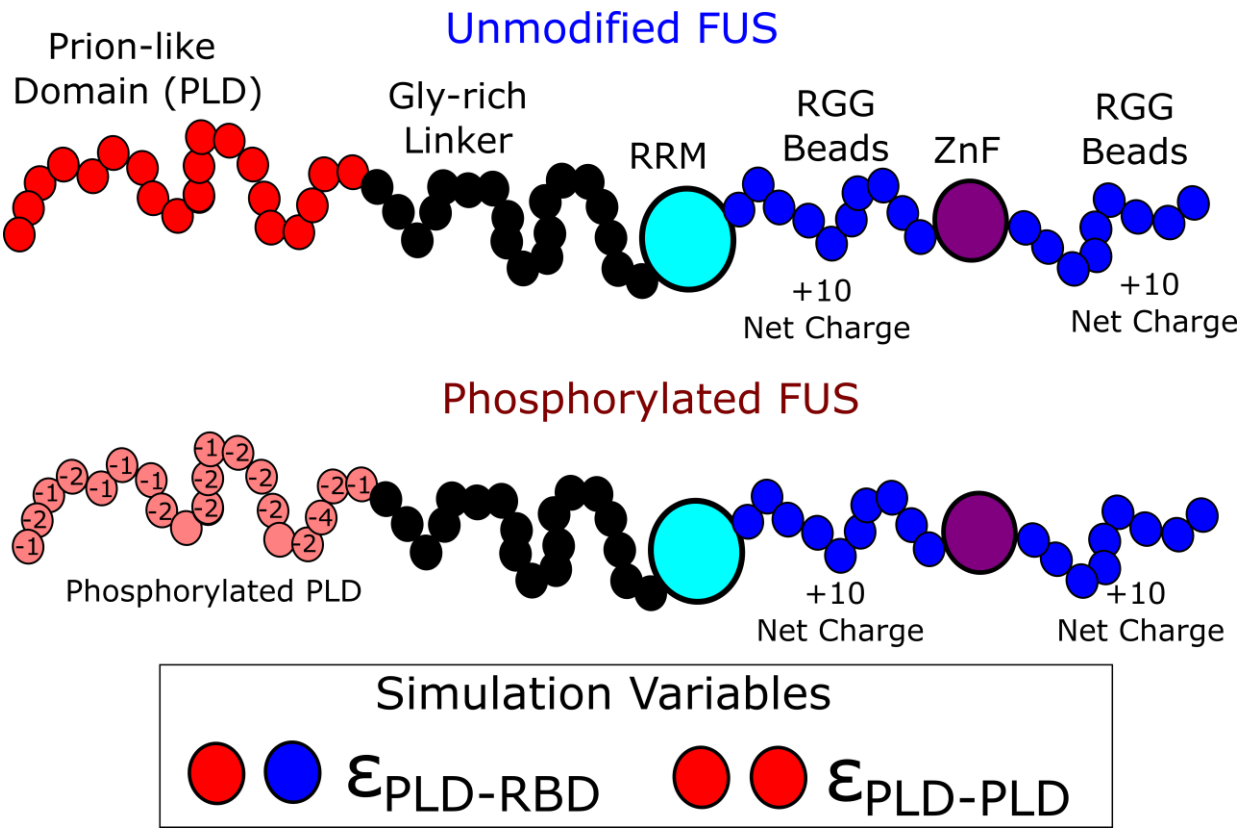

Figure 1: The Model. Coarse-grained representation of unmodified and fully phosphorylated variants of FUS. Red beads represent 8-amino acid long coarse-grained PLD stretches. Blue beads represent 5 amino-acid long stretches from the RGG region of FUS. The glycine rich linkers are shown in black. Two folded domains, the RNA-recognition motif (RRM) and the Zinc Finger Domain are modeled as large idealised beads. The two key simulation variables in this study, $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$ are shown in the box.

Model

In this work, we employ a coarse-grained model for a model scaffold protein, FUS where we coarse-grain stretches of 8 amino acids on the PLD into an effective interaction site (Fig.1). In effect, the 160-residue long PLD of FUS is modeled through 20 effective interaction beads. The RGG (arginine-glycine rich) region is represented at a resolution of 5 amino-acids per bead, with 20 effective RGG interaction beads making up the coarse-grained FUS chain. The net charge on the RBD is positive due to a high density of arginine residues in the RGG. The folded stretches of the protein, RRM and the ZnF motif, are modeled as large effective interaction domains. The size of the RGG and PLD beads was set to 6 Angstroms while the folded RRM and ZnF motifs were of 8 Angstroms. The RBD and the PLD are connected by a flexible linker which does not participate in interactions, mimicking the glycine-rich region in FUS. The glycine-rich linker (~100 residues long) is modeled with 20 beads in our study. Each effective interaction site harbors a charge based on the corresponding region of the full-length protein sequence, and based on the phosphorylation state of the protein. The persistence length (L_p) of the PLD and RGG regions was set to 4 residues while that of the glycine linker was set to 1 residue. We employ metadynamics simulations to enable the self-assembly of polymer chains into a single large cluster.

Methods

Coarse-grained models for proteins/peptides, at varying resolutions, have been used to study relationship between biomolecular structure and self-assembly[24,27,28]. Due to the fewer degrees of freedom and higher computational tractability, these coarse-grained models are efficient tools to probe structural transitions in self-assembling polymers[29]. In order to study the diversity in self-assembled states in response to varying PLD-PLD and PLD-RBD interaction strengths, we performed Langevin dynamics simulations. To simulate self-assembly, we perform simulations with 100 coarse-grained FUS chains (61-beads each) in a cubic box, with periodic boundaries enabled.

The coarse-grained FUS polymer beads interact via the following potential functions. The neighboring beads in a polymer chain are bonded via the harmonic potential with energy

$$E_{stretch} = k_s \sum_{i=1}^{M-1} (|\vec{r}_i - \vec{r}_{i+1}| - r_0)^2, \quad (1)$$

where $\vec{r}_i$ and $\vec{r}_{i+1}$ correspond to the positions of the ith and $(i+1)^{th}$ beads, respectively; r_0 is the resting bond length. k_s denotes the spring constant and has units of kT/Å^2. This interaction ensures the connectivity between the adjacent beads of the same polymer chain.

To model semi-flexible polymers, neighboring bonds in a polymer chain interact via a bending potential

$$E_{bend} = \kappa \sum_{i=1}^{M-2} (1 - cos\, \theta_i), \quad (2)$$

where θ_i refers to the angle between i^{th} and $(i+1)^{th}$ bond. Here, κ the bending stiffness, has units energy, kT. The value of κ was chosen such that the persistence length (Lp = κ/kT) of the polymer chain was ~4 FUS PLD/RGG/Linker beads. It must therefore be noted that the contour

length of the PLD (L_{PLD}) and the RBD (L_{RBD}) regions in the coarse grained polymer are 5 times that of the persistence length of the polymer. The contour length of the full length coarse-grained FUS is 15 times that of the persistence length, indicating the semiflexible nature of the modeled polymer chains.

Interactions between any pair of non-bonded beads were modeled using the Lennard-Jones potential.

$$E_{nb} = 4\epsilon \sum_{i<j} \left[\left(\frac{\sigma}{|\vec{r}_i - \vec{r}_j|} \right)^{12} - \left(\frac{\sigma}{|\vec{r}_i - \vec{r}_j|} \right)^{6} \right], \quad (3)$$

for all $|\vec{r}_i - \vec{r}_j| < r_c$. Here r_c refers to the interaction cutoff beyond which LJ interactions are neglected. The cutoff for LJ calculations were set to 2.5 times of σ.

To model electrostatic interactions, we use the standard Debye-Huckel electrostatic potential with a screening length (κ) of 1 nm and a dielectric (D) of 80. The Debye-huckel potential has the form,

$$E_{DH} = \frac{q_i q_j}{4\pi . Dr} . exp \frac{-r}{\kappa} \quad (4)$$

q_i and q_j are the charges of the two interacting beads. Each RNA bead carries a single negative charge in our simulations.

The attractive part of the LJ potential, ϵ, is tuned to vary the strength of the interaction. In this study, the strength of interaction between PLD beads is referred to as $\epsilon_{PLD-PLD}$ while PLD and RBD beads interact with a strength $\epsilon_{PLD-RBD}$. We used the LAMMPS molecular dynamics package[30] to perform the dynamic simulations, where the simulator solves Newton's equations with viscous force, and a Langevin thermostat ensuring an NVT ensemble with temperature $T = 310$ K. An integration timestep (dt) of 15 fs was used for the simulations.

Metadynamics

To enable efficient sampling of the self-assembly landscape and to ensure the formation of a single large cluster, we employed metadynamics[31]. In this method, the evolution of the self-assembling system is influenced by a history-dependent biasing potential. Energy gaussians are deposited during the course of a trajectory along the free energy space for a chosen reaction coordinate that effectively describes the process of self-assembly. The deposited potentials are then used to reconstruct the free energy profiles along this chosen reaction coordinate[31]. In order to drive self-assembly and observe coalescence of protein-rich clusters into a single large cluster, we used the radius of gyration of the peptide system, R_g^{system} as the reaction coordinate. R_g^{system} is the radius of gyration of an imaginary polymer that links together the center of masses of the individual polymer chains in the simulation. Large values of R_g^{system} correspond to a fully mixed system whereas smaller values indicate a self-assembled state. The systems were simulated till the free energy profiles showed convergence. The structures that populated the free energy minimum (single largest cluster) were used for structural analysis. The coalesced large clusters were then simulated for a further 15 microseconds using conventional Langevin dynamics simulations to compute the statistical quantities described in this paper.

Orientational Order Parameter

To characterize structural order in the FUS assemblies, we employ the nematic order parameter which is commonly used in liquid-crystal literature to quantify the extent of alignment of molecules along a reference axis[32–34].

We first define the long axis of each of the 'N' FUS chains within the self-assembly by a unit vector $\hat{z}_i$ which connects the two terminal beads of the PLD domain. In order to quantify local order, we define a second order, nematic order tensor, $Q^i_{\alpha\beta}$, for each of the 'N' FUS chains within a self-assembled cluster where the index i refers to the identity of each chain within the self-assembly.

$$Q^i_{\alpha\beta} = \frac{1}{N_{neigh}} \sum_{j=1}^{N_{neigh}} \left(\frac{3}{2} z^j_\alpha z^j_\beta - \frac{1}{2}\delta_{\alpha\beta}\right) \quad (5)$$

where z^j_α,z^j_β are component vectors (α, β =x,y or z) corresponding to each of the N_{neigh} chains that the ith chain is physically in contact with. In other words, we loop over each FUS chain within the self assembly and identify other chains that lie within an interaction radius of 2.5σ. If any part of two chains falls within this interaction cutoff, the chains are considered to be in physical contact. Therefore, we have N such order local order tensors ($Q^i_{\alpha\beta}$), one corresponding to the region defined by each FUS chain and its corresponding neighbors. As we loop over each chain, we scan through different regions of the self-assembly and compute the corresponding local order.

The N tensor order parameters, $Q^i_{\alpha\beta}$, hold information related to the extent of local order (a scalar order parameter S_i), and the direction of local alignment, a unit vector $\hat{d}_i$ known as the director. From the above nematic tensor $Q^i_{\alpha\beta}$, we can obtain S_i and $\hat{d}_i$ by computing the largest eigenvalue and its corresponding eigenvector. We employ this scalar order parameter, S_i to quantify the magnitude of local orientational order in the self-assembled clusters. This order parameter S_i ranges from 0 to 1. $S_i \to 0$ indicates weak local order, while $S_i \to 1$ suggests high local orientational order.

Equivalently, the scalar order parameter S_i is the 2nd order Legendre polynomial which can be expressed as,

$$S_i = \frac{1}{N_{neigh}} \sum_{j=1}^{N_{neigh}} \left(\frac{3}{2}(\hat{z}^j . \hat{d}_i)^2 - \frac{1}{2}\right), \text{ where i ranges from 1 to N} \quad (6)$$

In our study, we use $< S >= \frac{1}{N}\sum_{i=1}^{N} S_i$, to define the extent of local ordering within a self-assembled structure.

Results

Scaffold proteins of the FUS family (Fused in sarcoma) are particularly well studied for their phase-separation behavior[14]. These proteins typically harbor two domains – a low complexity, prion-like domain (PLD), and an RNA-binding domain. The prion-like domain and the RGG region (arginine, glycine rich region) of the RBD exhibit a high degree of disorder[35].

Earlier studies have established the ability of the full length FUS proteins[14] as well as the prion-like domains to form liquid-like droplets[14], at different saturation concentrations. Interestingly, these liquid-like droplets are often observed to progressively harden into more ordered assemblies that exhibit solid-like behavior. In this work, we vary the strengths of the PLD-PLD and PLD-RBD interactions to probe whether a switch in interaction networks could result in an emergence of ordered structures.

Interplay between homo- and hetero-domain interactions results in structures with distinct interaction networks

FUS is known to spontaneously phase separate in vitro and also act as a scaffold protein in RNA-protein condensates in the cell[15,18]. While the prion-like domain (PLD) of FUS is known to phase separate in isolation, these droplets form at much higher threshold concentrations than that of the full-length protein[14]. These results suggest that FUS droplets are stabilized by a hierarchy of interactions. Further, the PLD and the RBD of FUS also harbor several post-translational modification sites, further tuning the interaction networks. Here, we systematically vary two key interaction parameters, $\epsilon_{PLD-PLD}$ & $\epsilon_{PLD-RBD}$ to probe the interactions that stabilize the condensed phase.

In order to drive the self-assembly of the proteins into a single, large, condensed phase, we employ metadynamics simulations. We use the radius of gyration of the system (R_g^{system} of N protein chains as a collective variable to define the self-assembled state of the system. The free energy profiles were computed along this reaction coordinate, with minima around small values of R_g^{system} indicating a fully assembled equilibrium state. In Fig.2A, we demonstrate the free energy profiles for four such scenarios, with varying strengths of hetero-domain ($\epsilon_{PLD-RBD}$) and homodomain ($\epsilon_{PLD-PLD}$) interactions. As evident from Fig.2A, for all the four scenarios, we observe a distinct minima for $R_g^{system} \approx 50$ A (shaded green region), a value that corresponds to a state where there is a single large condensed phase. The yellow region in the free energy profiles corresponds to the multi-cluster state, with local minima in this region potentially indicative of kinetic barriers to cluster growth. The free energy profiles suggest that the protein concentrations and the range of interaction strengths chosen for study favor self-assembly of FUS into a large condensate (≈ 100 FUS chains). The metadynamics simulations also reveal that FUS clusters could form in the limit of weak (top panels in Fig.2A) or strong hetero-domain interactions (bottom panels in Fig.2A). Crucially, even when both homotypic interactions ($\epsilon_{PLD-PLD}$) and the heteropypic ones ($\epsilon_{PLD-RBD}$) are relatively weak (0.3 and 0.15 kT respectively), the system tends to assemble into protein-rich clusters.

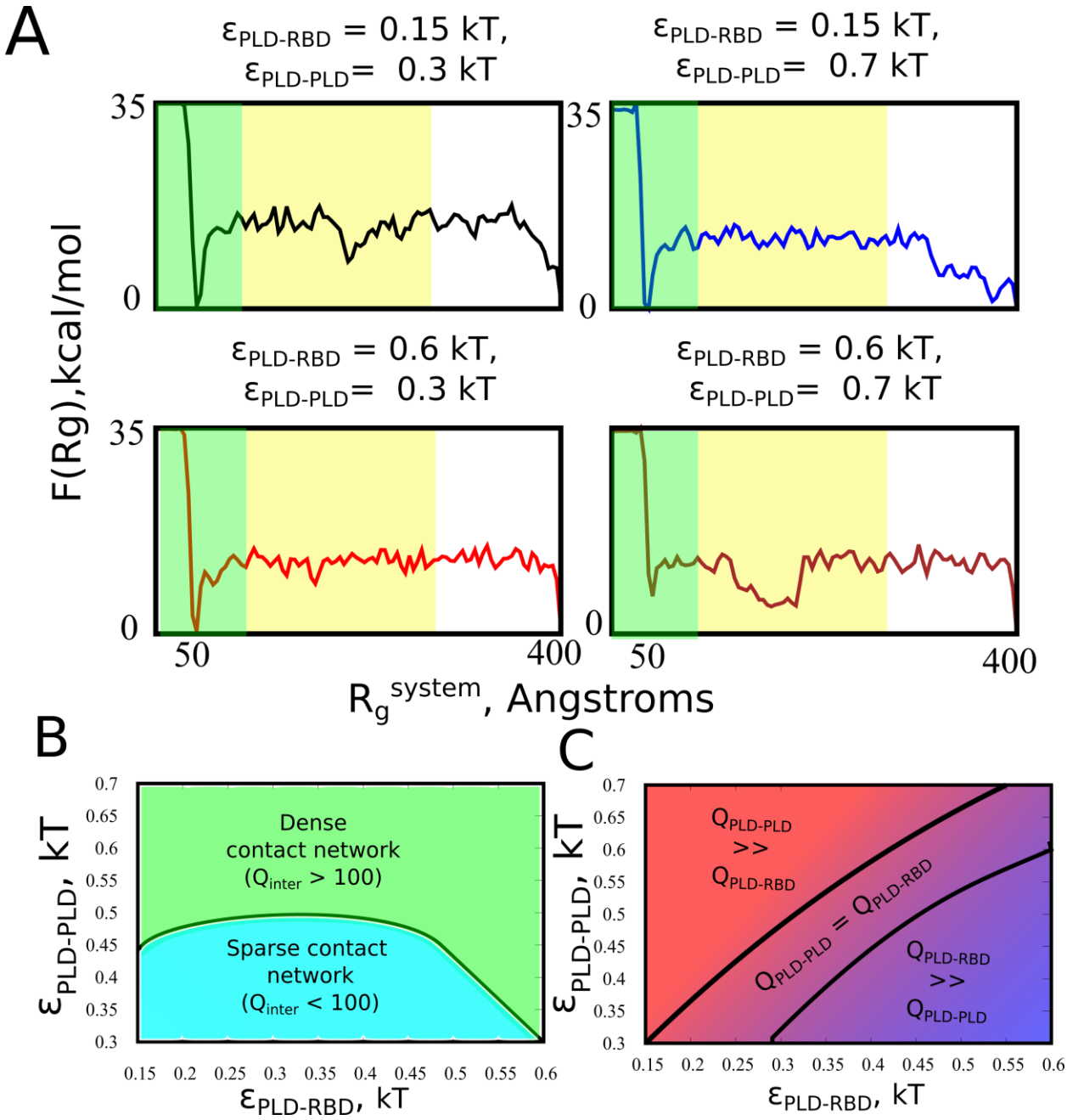

Figure 2: **Intermolecular Contacts.** A) Free energy profiles for different combinations of PLD-PLD and PLD-RBD interaction strengths. A sharp minima at ≈ 50 Å (shaded green) is indicative of an equilibrium state corresponding to a single large self-assembled cluster. As $R_g^{system} \rightarrow 400$, the system approaches the fully mixed state. B) Phase diagram describing the degree of inter-molecular connectivity within the self-assembled state quantified by the mean inter-molecular contacts per FUS chain (Q_{inter}) within the cluster. C) Switching of predominant interaction network stabilizing the condensate, as a function of varying $\epsilon_{PLD-PLD}$ & $\epsilon_{PLD-RBD}$. $Q_{PLD-PLD}$ and $Q_{PLD-RBD}$ refer to the inter-molecular contacts per chain between PLD-PLD and PLD-RBD domains, respectively.

Upon establishing that the condensed phase indeed corresponds to an equlibrium state, we proceed to characterize the interaction network of the condensed phase in a broad range of

interaction parameters. In Fig.2B, we plot the extent of inter-molecular contacts chain as a function of strengths of homotypic interactions between amino acids in the PLD domain and heterotypic interactions between PLD-RBD domains, $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$, respectively. For weak interaction strengths, the interaction network within the FUS cluster is sparse ($<= 100$ inter-molecular contacts per chain). For strong PLD-PLD interactions, we observe a dense inter-molecular contact network (> 100 contacts per chain) suggesting a different intra-cluster interaction network in this regime. Upon establishing a variation in the extent of intermolecular connectivity within the droplet upon tuning the interaction parameters, we probe the nature of interactions that make up these intermolecular contacts. In Fig.2C, we plot the ratio of PLD-PLD to PLD-RBD intermolecular contacts to elucidate the dominant intermolecular interactions that stabilize the condensed phase. For weak $\epsilon_{PLD-RBD}$ and strong $\epsilon_{PLD-PLD}$, the FUS clusters are stabilized by homotypic interactions between PLD domains from different chains. On the other hand, in the limit of weak $\epsilon_{PLD-PLD}$ and strong $\epsilon_{PLD-RBD}$, hetero-domain interactions dominate. It must, however, be noted that a variation in the interaction strengths not only results in a change in contact ratios, but also affects the density of intermolecular contacts within the cluster (Fig.2B and C). In a narrow range of interaction strengths, the contact ratios $\rightarrow 1$ ($Q_{PLD-RBD} = Q_{PLD-PLD}$). In this limit, the intra-cluster environment exhibits both homo- and hetero-domain contacts with equal likelihood. Overall, tuning the two interaction strengths within a narrow, biologically relevant regime (0.3 to 0.7 kT) results in self-assembled structures stabilized by diverse intra-cluster interaction networks.

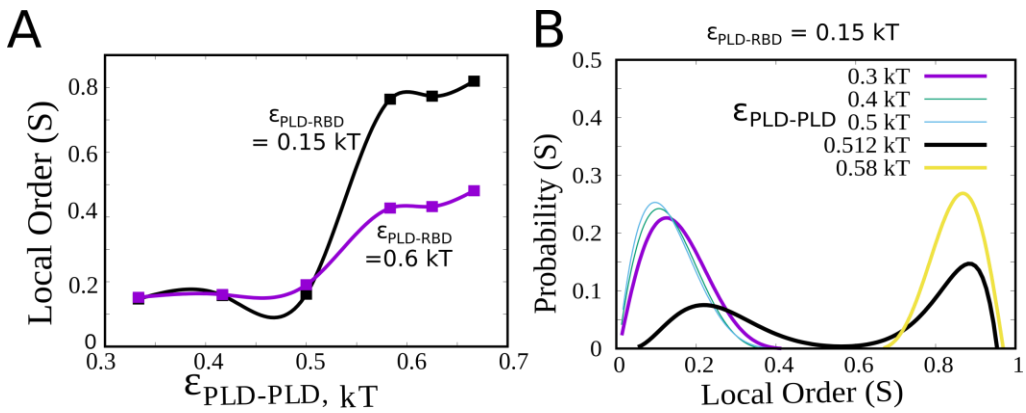

Figure 3: **Disorder-Order Transition.** A) Local orientational order as a function of $\epsilon_{PLD-PLD}$, for different strengths of $\epsilon_{PLD-RBD}$. The solid curves are provided as a guide to the eye. B) Distribution of the local orientational order parameter (S), for different values of $\epsilon_{PLD-PLD}$. The distributions were computed across 100 different local order parameter values and 50 different structures at equilibrium.

Disorder-order transition is an outcome of hetero to homo-domain interaction network switching

Tuning the strength of homo- and hetero-domain interactions in the thermally relevant regime results in structures that are stabilized by dramatically different interaction networks. In their seminal experimental study, Wang et al have previously established that a hierarchy of interactions – homo- and hetero-domain – can drive phase-separation of FUS into protein-rich

droplets[14]. Also, FUS droplets have also been observed to progressively transition into more ordered, solid-like filamentous structures upon aging[36,37]. Here, we probe whether a variation in homodomain ($\epsilon_{PLD-PLD}$) and heterodomain ($\epsilon_{PLD-RBD}$) interactions could result in protein assemblies with altered morphologies. To characterize the morphology of the assemblies, we employ the nematic orientational order parameter S which defines the extent of local orientational order for the prion like domains within a cluster. This order parameter, S can take any value from 0 to 1. $S \rightarrow 1$ indicates a high degree of orientational order. On the other hand, $S \rightarrow 0$ suggests an entirely disordered system.

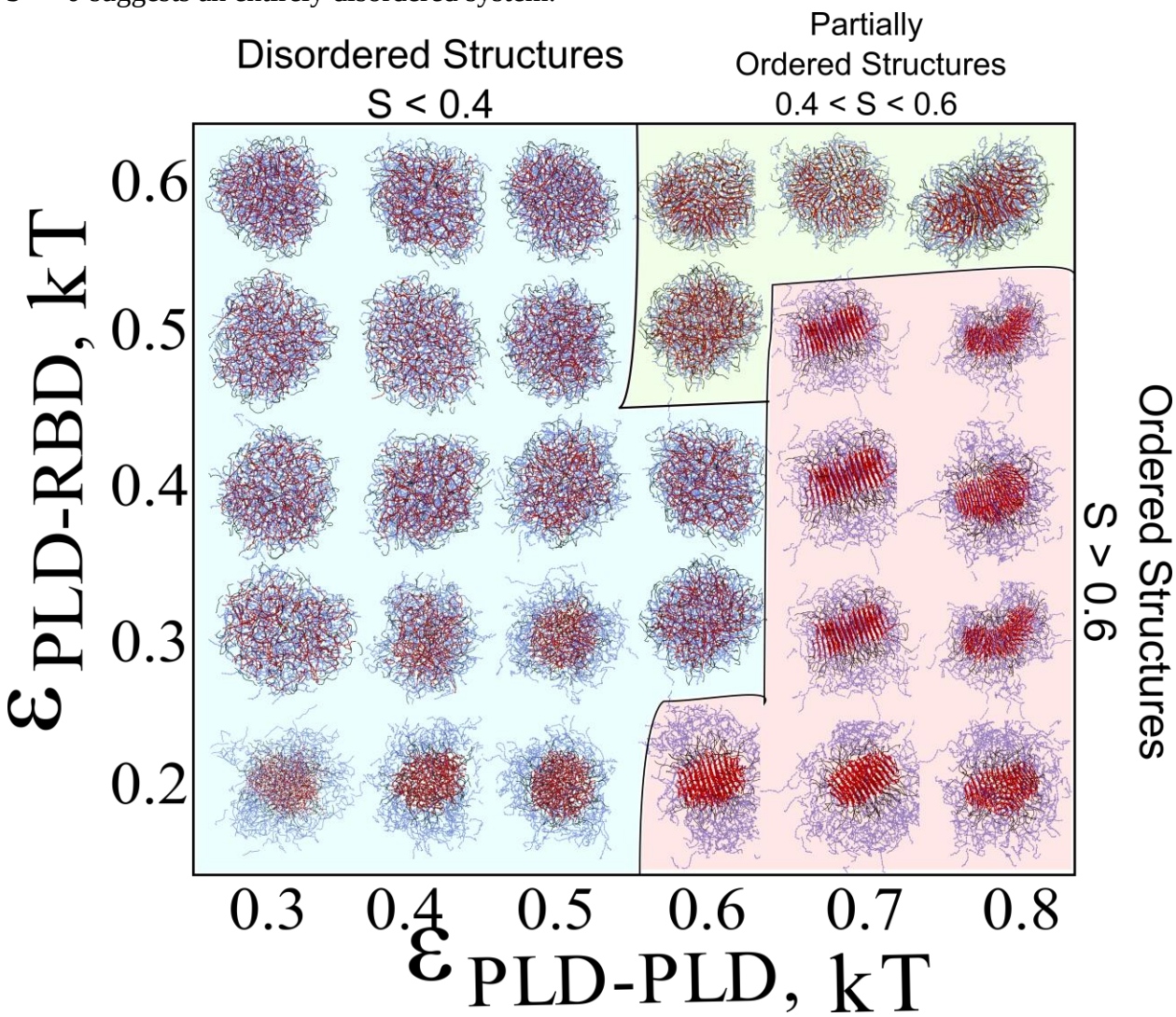

Figure 4. . **Structural Phases.** Phase diagram desribing the various structural phases in response to varying strengths of homo-domain($\epsilon_{PLD-PLD}$) and heterodomain ($\epsilon_{PLD-RBD}$) interactions. Structures are described as ordered, disordered and partially ordered based on the mean value of the nematic orientational order parameter S. Assemblies with S > 0.6 are labeled as ordered while those with S < 0.4 are labled as disordered structures. PLD residues are shown in red, while RBD regions are shown in blue.

Using this order parameter, we capture the extent of orientational order as a function of varying interaction parameters $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$. In Fig.3A, we plot S as a function of

increasing PLD-PLD interactions for two different strengths of PLD-RBD interactions. Interestingly,an increase in $\epsilon_{PLD-PLD}$ results in a first order-like, disorder → order transition for weak $\epsilon_{PLD-RBD}$. A further affirmation of the discontinuous nature of this transition could be ascertained from the evolution of the distribution of S for increasing values of (Fig.3B). For weak homodomain interactions ($\epsilon_{PLD-PLD}$<= 0.5 kT in Fig.3B), we see the distributions peaking at low values of S (~0.2), suggesting that the assemblies never access the ordered states in this interaction regime. Interestingly, an increase in $\epsilon_{PLD-PLD}$ from 0.2 to 0.5 kT results in a negligible shift in distribution. However, a further increase in homodomain interaction strength from 0.5 to 0.6 kT results in a shift from a unimodal distribution (fully disordered) to a bimodal distribution with peaks corresponding to both the ordered and disordered states. Such a bimodality is a typical signature of discontinuity. A further increase in interaction strength results in a unimodal distribution with the assemblies populating the ordered state alone. For stronger $\epsilon_{PLD-RBD}$ interaction, the <S> vs $\epsilon_{PLD-PLD}$ curve saturates at lower values of the order parameter, owing to a competition between homo-domain and heterodomain interactions.

In Fig.4, we show how a variation in these two parameters gives rise to different structural phases with varying degrees of nematic ordering. In the $\epsilon_{PLD-PLD}$-$\epsilon_{PLD-RBD}$ phase diagram in Fig.4, we define all structures with $< S > < 0.4$ as disordered assemblies. $< S > > 0.6$ were defined as ordered assemblies while those with intermediate values of $< S >$ were labeled as partially ordered structures. The S values were averaged across 100 different equilibrium configurations, and across 100 polymer chains. When both homo- and heterodomain interactions are weak ($\epsilon_{PLD-PLD}$ & $\epsilon_{PLD-RBD} < 0.5$ kT), we encounter structures which are disordered, and exhibit spherical geometry. As we increase the strength of homodomain interactions, in the limit of weak heterodomain interactions ($\epsilon_{PLD-RBD} < 0.5$ kT), we encounter ordered structures with the PLD domains displaying a high degree of orientational order with 'fibril-like' geometry. When both interactions are competitively strong ($\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$ > 0.5 kT), we observed structures with partial order. Therefore, in the regime where homotypic interactions between PLD domains dominate ($Q_{PLD-PLD} > Q_{PLD-RBD}$ in Fig.2C) the self-assembled state has a high propensity to organize into *ordered* structures stabilized by PLD-PLD contacts. An interplay between these homodomain and heterodomain interactions can therefore result in structures with varying morphologies.

Phosphorylation of the PLD is a barrier to the formation of ordered structures

Our results so far show that even within a narrow range of interaction strengths, in the thermally-relevant limit, the interplay between homo- and hetero-domain interactions can result in ordered structures. Phase-separated structures in the cell could localize proteins at concentrations as high as 50-100 times that of their average bulk concentrations[38,39]. Such a large local density of prion-like domains in these condensates makes these structures prone to forming filamentous structures. Previous experimental studies suggest that post-translational modifications could potentially act as a barrier to liquid-solid transitions in cellular condensates[37,40].

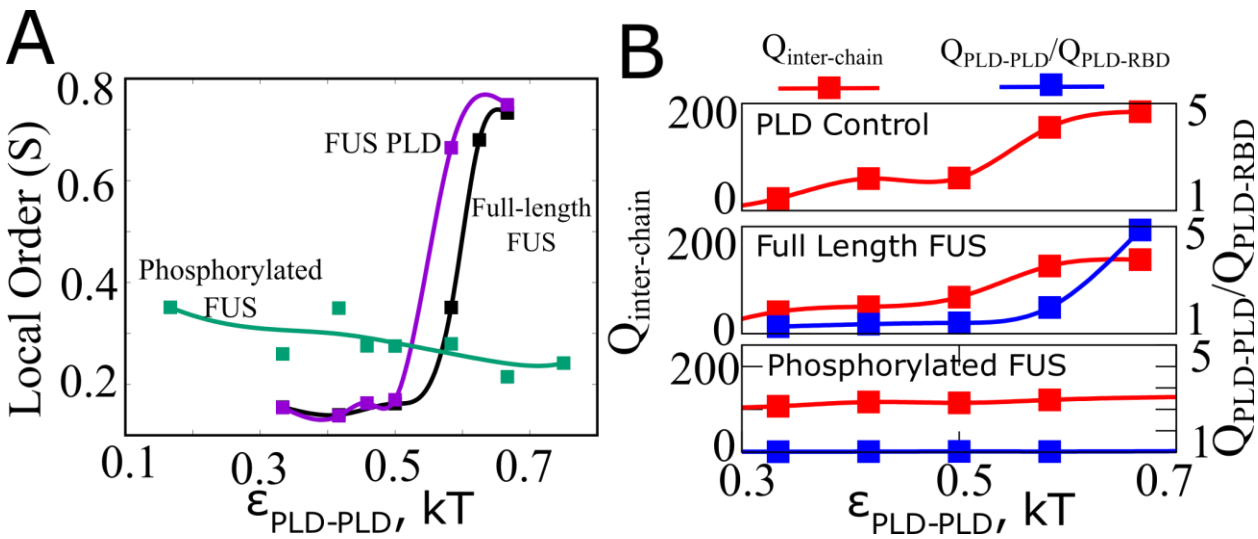

Figure 5: **Phosphorylation abrogates order.** A) Orientational order as a function of $\epsilon_{PLD-PLD}$ for the PLD alone, full-length FUS and phosphorylated FUS. Phosphorylation results in disordered structures even at strong $\epsilon_{PLD-PLD}$. B) Mean inter-chain contacts per FUS chain ($Q_{inter-chain}$) and the contact ratio ($Q_{PLD-PLD}/Q_{PLD-RBD}$) for PLD alone (top), full-length FUS (middle) and phosphorylated FUS (bottom) as a function of $\epsilon_{PLD-PLD}$. Phosphorylation results in lower values of contact ratios. The solid lines are guide to the eye.

In this context, we modeled a key post-translational modification that is typically associated with the FUS PLD, i.e phosphorylation. The FUS PLD in primates harbors 31 potential phoshposites[41]. Here, we model phosphorylation by introducing a single negative charge per phosphosite in our coarse-grained PLD beads. The net negative charge on each phosphorylation PLD bead corresponds to the number of phosphosites residing in a particular 8-residue sequence that the coarse-grained bead maps onto. The net charge on the phosphorylated variant of the PLD is -31. In Fig.5A, we compare the propensity of three variants of FUS –the PLD domain only, Full Length FUS, and phosphorylated FUS – to form ordered structures. The PLD control simulations reveal that the PLD domain can undergo a disorder-order transition on its own in the limit of $\epsilon_{PLD-PLD} > 0.5$ kT. Similarly, the Full Length FUS, for $\epsilon_{PLD-RBD} = 0.5$ kT, also shows a $\epsilon_{PLD-PLD}$-dependent disorder-order transition, with lower $< S >$ values than the PLD-only control simulations. The phosphorylated FUS, on the other hand, shows no such propensity to form ordered structures even at $\epsilon_{PLD-PLD}$ approaching 1 kT. This result suggests that the high local density of negative charges in the PLD of phosphorylated FUS results in an abrogation of order formation in these structures.

We further probe whether this loss of ordering in phosphorylated version of FUS is a result of overall loss of inter-molecular contacts within the FUS-cluster. As evident from the $Q_{inter-chain}$ vs $\epsilon_{PLD-PLD}$ curve for the phosphorylated FUS (Fig.5B, bottom panel), the modification to the PLD results in number of intermolecular contacts ($Q_{inter-chain}$) comparable to the unmodified variant. However, unlike the other two scenarios, the phosphorylated FUS shows no increase in contact ratio ($Q_{PLD-PLD}/Q_{PLD-RBD}$) even at strong $\epsilon_{PLD-PLD}$ interactions. In the corresponding limit of strong $\epsilon_{PLD-PLD}$, the Full length FUS displays contact ratio values approaching 5. This shows that phosphorylation results in an upregulation of the heterodomain

interactions over homodomain interactions, and thereby results in a more disordered condensed phase.

Disorder-order transitions are characterized by sharp density transitions

The transition from droplet-like morphology to a more fibrillar morphology in FUS protein assembly is also accompanied by a loss of dynamicity of the condensate[36]. We therefore probe how tuning $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$ alters local densities of PLD and RBDs within the clusters. In Fig.6A, we plot the pair-correlation function, g(r) for different domains in the condensed phase. The g(r) plotted in Fig.6A is the mean number of atom pairs of any type at a given radial distance, normalized by a pair of ideal gas particles for the same bulk density. The g(r) values in Fig.6 refer to the average number of atom pairs of any type (PLD-PLD, PLD-RBD, RBD-RBD) found at any radial distance (between r and r+dr), normalized by the equivalent values for a system of ideal gas particles of the same bulk density. As the peak value of g(r) → 1, the system exhibits gas-like organization. The higher this peak in g(r), ($g(r)^{peak}$ in Fig.6A), the system exhibits more solid-like organization. Also, the higher the g(r) peak, the denser the packing, and potentially less dynamic the condensed phase.

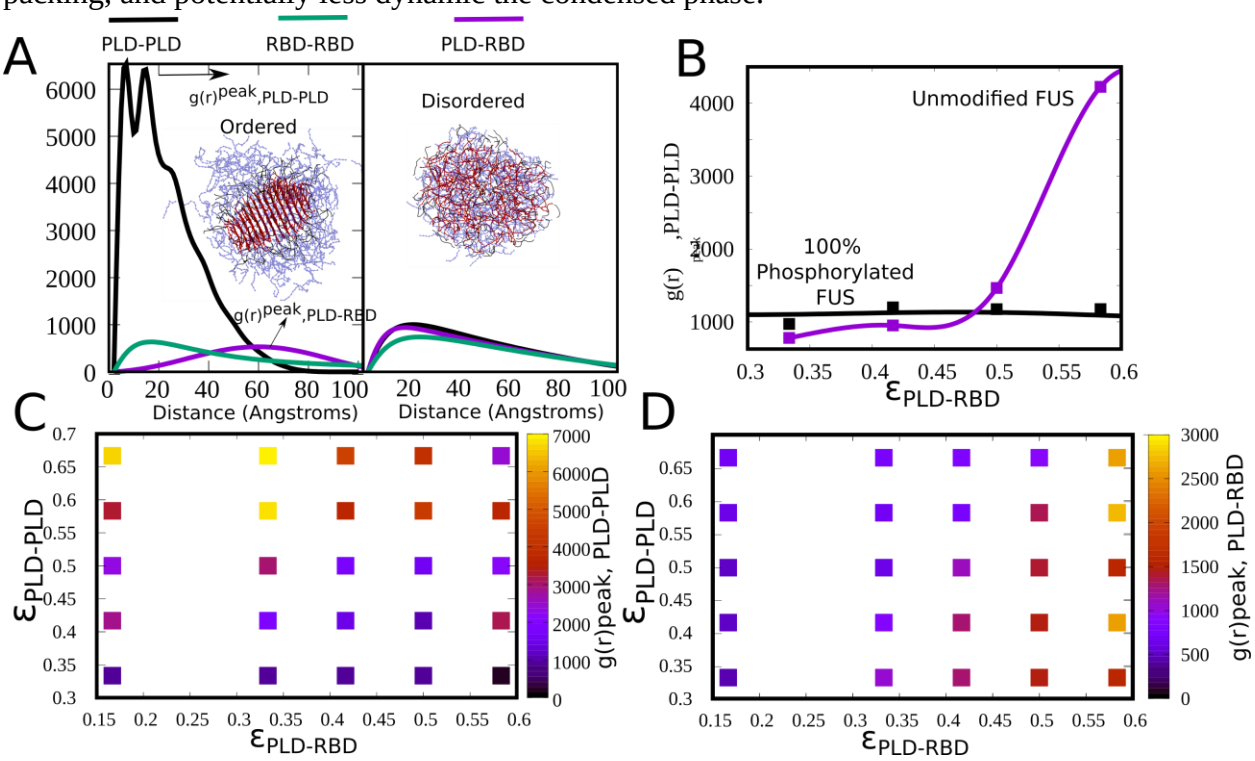

Figure 6: Local Density of Domains. A) Pair correlation function, g(r), for PLD-PLD (black), RBD-RBD (green) and PLD-RBD (purple) contacts within the largest cluster, for ordered and disordered structures. The g(r) here is essentially the probability of finding any two particle types (PLD-PLD, PLD-RBD, RBD-RBD) within a certain distance of each other normalized by corresponding value for a pair of ideal gas particles at the same bulk density. For ordered structures, the g(r) for PLD-PLD interactions peaks several-fold higher than the other two curves suggesting a high local density of PLD domains within the cluster. For disordered structures, the three curves peak at similar values indicating a more homogenous density throughout the cluster. B) The peak values of g(r) plotted as a function of $\epsilon_{PLD-PLD}$. For full-length FUS, the g(r) peak for PLD-PLD contacts shows a sharp sigmoidal transition

with increasing $\epsilon_{PLD-PLD}$, suggesting a sharp transition from a low to high density phase. For phosphorylated variants of FUS, such a transition was not observed. C) and D) Phase diagrams showing the density of PLD-PLD and PLD-RBD packing, respectively. An increase in PLD-PLD interactions results in a high packing density for PLD domains within the cluster, as evident from the g(r) peak values for PLD-PLD interactions. In contrast, little variation in PLD-RBD g(r) peaks was obseved throughout the parameter space explored.

In Fig.6A, we show representative g(r) curves for ordered and disordered structures. For ordered structures, the g(r) for PLD-PLD interactions shows peak values several-fold higher than the corresponding curves for PLD-RBD and RBD-RBD interactions. This suggests that the ordered structures have a dense PLD-core and a less dense RBD shell, as also evident from the structures shown in Fig.4. For disordered structures formed in the limit of weak interactions ($\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD} < 0.5$ kT), the g(r) curves for all three pair-wise interactions are comparable, with peak values several-fold lower than the peak g(r) values for PLD-PLD packing in the ordered structures. This result suggests that intra-cluster densities of polymer domains in the disordered structure is more uniform compared to the ordered structures which exhibit a dense core and a less dense shell.

We further probe how phosphorylation influences the peak densities of PLD domains within the FUS clusters by plotting the $g(r)^{peak}$ for PLD-PLD contacts as a function of $\epsilon_{PLD-PLD}$. Strikingly, even at higher $\epsilon_{PLD-PLD}$, the phosphorylated variant shows no change in peak densities within the condensed phase. The full length FUS, on the other hand, shows a discontinuous transition from uniform, less dense structures to the more dense core-shell structures that typify order. We further plot the peak densities for the PLD-PLD and PLD-RBD pairwise contacts as a function of varying $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$. Interestingly, the regime of interaction strengths that corresponds to a high contact ratio (Fig.2B and C) and high orientational order (Fig.4) also corresponds to large peaks in g(r) for PLD-PLD (Fig.6C). However, the peak values of g(r) for PLD-RBD interactions (Fig.6D) are significantly lower even for strong PLD-RBD interactions. This could potentially be an outcome of the cluster of positive charges that are localized within the RGG regions of the RBD, preventing a dense packing of RBD domains together. This suggests that the density transitions are driven by a change in PLD ordering and not through PLD-RBD interactions, in this narrow range of interaction strengths.

Assemblies can switch from fully-mixed to multi-phase architecture on switching interaction networks

Multi-component mixtures (composed of several protein types) can phase separate into multi-phase assemblies with sub-phases which localize different components based on interaction preferences. Diversity in interactions and interfacial tensions has been known to be a key driver in segregation of components between different sub-phases. Due to the multi-domain architecture of RBPs and the ability to form clusters stabilized by different interaction networks (Fig. 2C), we probed whether multiphase assemblies could result from single component mixtures of FUS alone. In Fig.6 we show how the interplay between the homo-domain and

hetero-domain interactions can result in hugely altered local densities of the PLD and RBD domains within the condensates. Given how density differences are a vital signature of multi-phase condensates, we looked at how different domains get packed within the single-component FUS assemblies stabilized by different interaction networks. In Fig.7, we show how FUS condensates could exhibit different sub-cluster organization for different combinations of interaction strengths. For weak interactions ($\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$ < 0.5 kT), when PLD-PLD interactions dominate ($\epsilon_{PLD-PLD}$ = 0.4 kT and $\epsilon_{PLD-RBD}$ = 0.2 kT), we observe spherical clusters with a multi-phase structure. Here, PLD segments form the core of the condensates while the RBD domains organize themselves in the outer shell. This is evident from the distributions of the radius of gyration of PLD domains which peak at a much lower value (~35 A), compared to that of RBD domains (~70 A) showing that the PLD domains form a condensed core. However, these core-shell structures formed by PLD, in the limit of weak interactions, do not exhibit nematic order. As the strengths of the two interactions become comparable, we observe single-phase clusters where the two domains are fully mixed (Fig. 6B), with the radius of gyration distributions for PLD and RBD segments showing significant overlap. In the strong interaction regime where PLD-PLD interactions dominate, we again observe core-shell structures. However, unlike the spherical disordered, core-shell structures observed when the two interactions are less than 0.5 kT, the PLD core (for $\epsilon_{PLD-PLD} = 0.6\,kT$) here shows fibril-like nematic ordering. Differential interaction strengths can thereby result in core-shell structures, even for assemblies that are made up of only one type of protein.

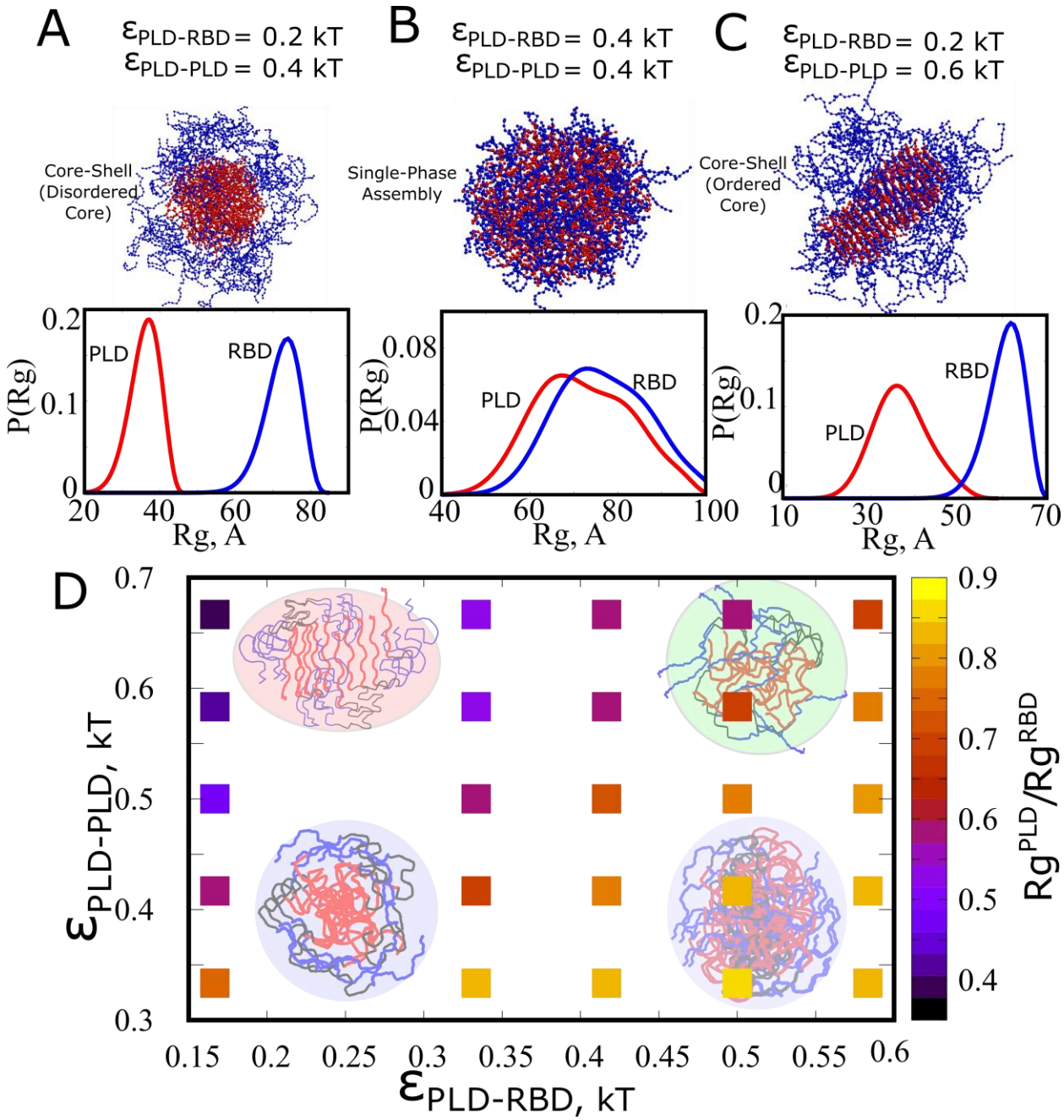

Figure 7: Different architectures of the FUS assemblies. A) For the limit of weak interactions when $\epsilon_{PLD-PLD} > \epsilon_{PLD-RBD}$, we observe disordered structures with the PLDs (red beads) phase separating from the RBDs (blue bead). B) As $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$ become comparable, we observe a disordered assembly where the PLD and RBD domains are fully mixed. C) Above a critical interaction strength, when $\epsilon_{PLD-PLD} > \epsilon_{PLD-RBD}$, we observe core shell structures with distinct PLD and RBD phases within the assembly. Unlike the disordered core-shell structures for weak interactions, the PLD core here shows a high degree of orientational order. The black and purple curves in (A),(B) and (C) show the distribution of radius of gyration of PLD and RBD domains, respectively. Small values of R_g indicate a collapsed core while large values indicate that the domains are spread out. When the two distributions do not overlap, it indicates a core-shell structure while overlapping distributions suggest a fully mixed cluster.

Dramatic slowdown of intra-droplet dynamics upon disorder to order structural transition

Our simulations reveal that the interplay between homo-domain and hetero-domain interactions can result in self-assembled structures with diverse architectures (single vs multiphase, Fig.7) and varying degrees of orientational order. However, aging of droplets and condensates in vivo and in vitro is not only characterized by structural transitions, but also a slow-down in intra-droplet dynamics. In this context, it becomes critical to ascertain whether the sharp disorder-order transitions in structural order and local densities that we observe above a critical $\varepsilon_{PLD\text{-}PLD}$ are also accompanied by altered intra-cluster dynamics. In order to probe the micro-environment within the FUS clusters, we run long 5-μs long simulations with the self-assembled clusters (localizing ~100 FUS chains) shown in Fig.4. Using these trajectories, we first plotted the mean square displacements (MS D) of the center of masses of PLD and RBD domains within the cluster (Fig.S2). Here,

$$MSD = \langle [r(t0 + t) - r(t0)]^2 \rangle \qquad (7)$$

The angular brackets in the equation for MSD indicates an ensemble average across all molecules within a cluster as well as diferent time origins along the trajectory.

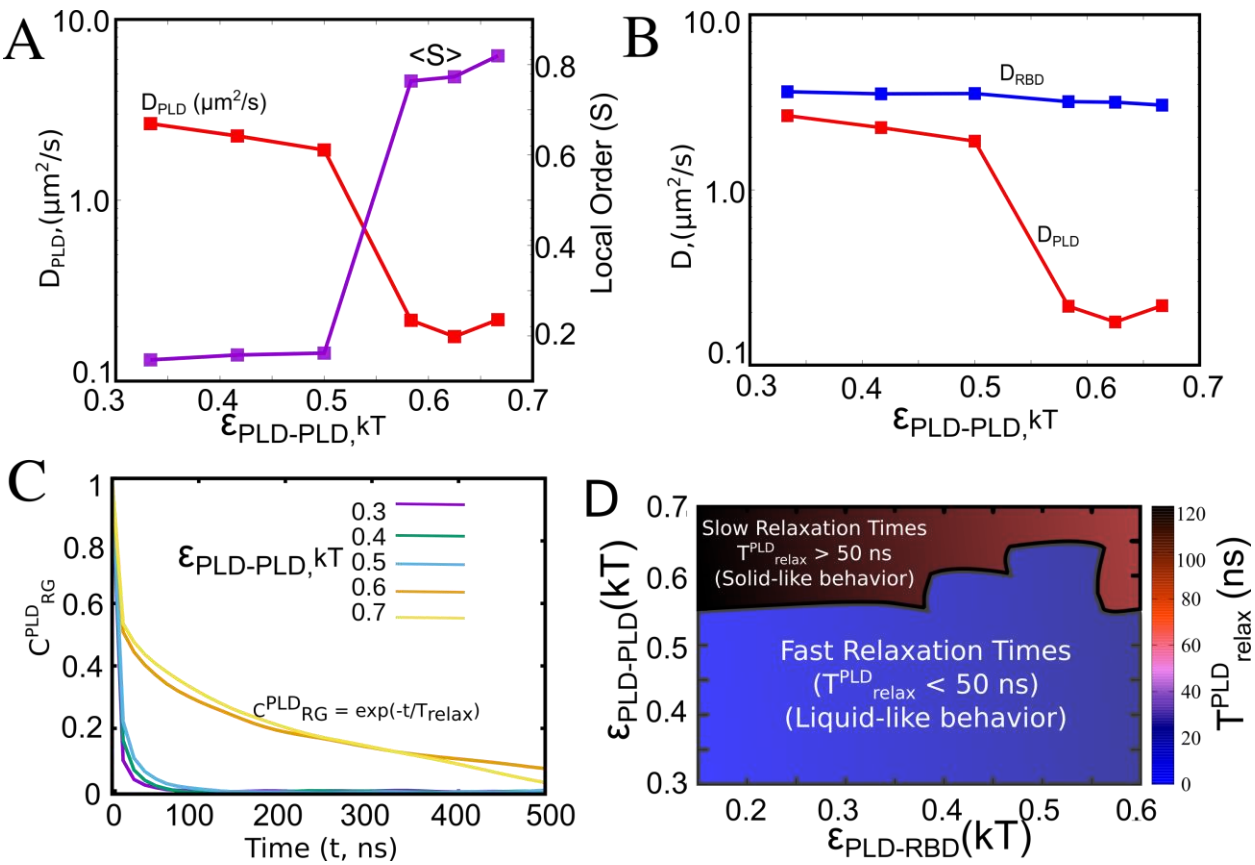

Figure 8: Intra-cluster Dynamics. A) Diffusion coefficients of the PLD domain of polymer chains within the self-assembled cluster, as a function of $\varepsilon_{PLD\text{-}PLD}$ (black curve, left-hand side Y-axis). As we increase the interaction strength for homodomain interactions above 0.5 kT, we observe a sharp slow-down in intra-droplet dynamics, with an order of magnitude decrease in D_{PLD}. This drop in D_{PLD} coincides with a sharp increase in local ordering of PLD

domains within the cluster (Purple Curve, right-hand side Y-axis). (B) Diffusion coefficients for PLD (red) and RBD (blue curve) domains within the cluster as a function of $\varepsilon_{PLD-PLD}$. At higher values of homodomain interactions, the diffusion coefficient of PLD domains shows an order of magnitude decrease. There is negligible reduction in D_{RBD}, in comparison. (C) Normalized Autocorrelation function (C^{PLD}_{RG}) for the radius of gyration of PLD domains within FUS chains as a function of time. For less viscous intra-cluster environment, the ACF curves decay rapidly to zero, suggesting that the system loses its memory at short timescales. On the other hand, the ACF profiles decay much slower at stronger values of $\varepsilon_{PLD-PLD}$ where we encounter ordered structures. $C^{PLD}_{RG} = \exp(-t/T_{relax})$ D) Measuring polymer dynamics as a function of $\varepsilon_{PLD-PLD}$ and $\varepsilon_{PLD-RBD}$. T_{relax} is the timescale that defines the rate at which autocorrelation function decays ($C^{PLD}_{RG} = \exp(-t/T_{relax})$). The higher this T_{relax}, the slower the dynamics of polymers within the protein rich cluster, and more viscous the intra-cluster environment. Low values of T_{relax}, on the other hand, indicate that the system loses its memory rapidly and signifies a relatively less viscous intradroplet environment.

We compute the diffusion coefficients (D_{PLD} and D_{RBD}) from these MSD curves (by fitting the curves to MSD = 6Dt). In Supplementary Fig. S2, we show the MSD profiles for PLD domains in various scenarios. As evident from Fig. S2, the diffusion coefficient for monomeric FUS in solution (~11 $\mu m^2/s$) shows a 5-fold reduction upon self-assembly into disordered FUS-rich clusters (~2.5 $\mu m^2/s$). However, while the disordered structures still exhibit diffusive behavior, structures with orientational order (Fig. S2, red curve) exhibit flat MSD profiles with two orders of magnitude decrease in D_{PLD}. In Fig.8A, we plot D_{PLD} (intracluster diffusion coefficient of PLD domains) as a function of homodomain interaction strengths for a fixed $\varepsilon_{PLD-RBD}$ of 0.15 kT. We observed two distinct dynamics regime and a sharp transition between them at $\varepsilon_{PLD-PLD} > 0.5$ kT. As we increase the homo-domain interaction strength from ~0.5 to ~0.6 kT, we observe a dramatic – two orders of magnitude - decrease in diffusion coefficients (black curve, Fig.8A, Supplementary Fig.S2), suggesting that the PLD domains exhibit solid-like dynamics in this regime (also evident from the flat MSD profiles in Supplementary Fig.S2 and S3). Interestingly, this sharp transition in diffusion coefficients is concomitant with a sharp increase in structural order (purple curve, Fig.8A) suggesting that solid-like intra-droplet dynamics emerges because of structural ordering. While the disorder-order structural transitions result in an order of magnitude reduction in D_{PLD}, the RBD regions show comparatively negligible drop in diffusion coefficients even in highly ordered structures, suggesting that these domains continue to exhibit liquid like dynamics within ordered structures. In Supplementary Fig.S4, we plot detailed phase diagrams that show how the diffusion coefficients of the two domains vary as a function of the homo- and hetero-domain interaction strengths.

We further computed the autocorrelation function, ACF, (computed from time traces of radius of gyration of FUS chains within the clusters) to study the characteristic timescales at which the system loses its memory[42,43]. The normalized autocorrelation function of the radius of gyration of FUS chains, C_{RG}, can be used to describe the dynamics of polymers within the dense polymer-rich clusters.

$$C_{RG} = \langle R_g(t).R_g(0)\rangle / \langle R_g(0)^2\rangle \qquad (8)$$

Here, the <> denotes an ensemble average over all the polymer chains that are part of the FUS

asssembly as well as using different timesteps as starting point along the trajectory.

As shown in Fig.8C, the ACF profiles decay with time, with faster rate of decay suggesting more liquid-like dynamics while a slower rate of decay is symbolic of more viscous intra-cluster environments. In order to quantify relaxation dynamics, we define an order parameter T_{relax} – the time at which the autocorrelation values drop to 0.01 when the system loses 99% of its memory. While, an arbitrary timescale, T_{relax} allows us to compare the relaxation dynamics of polymer chains within various self-assembled clusters and thereby offers us window into intradroplet dynamics. In Figure.8D, we show a phase diagram of polymer relaxation times as a function of the two interaction strengths. Consistent with our MSD profiles, we observed a sharp slowdown in polymer relaxation dynamics at strong $\varepsilon_{PLD\text{-}PLD}$ suggesting that the regime at which we observe ordered structures is inherently less dynamic. Overall, the different structural phases that we encounter upon tuning the two interaction parameters correspond to distinct dynamic regimes, suggesting that subtle changes in interaction strengths and alterations in interaction networks can result in arrested droplet dynamics.

Ordered structures can be seeded by strong non-specific interactions with other scaffold molecules

The contact ratios (Fig.2) , orientational order (Fig.3, 4) and peak intra-cluster densities (Fig.6) suggest that disordered-ordered transition in FUS clusters is mediated primarily by PLD-driven interactions. Further, the PLD switches from collapsed configuration to an extended ordered configuration above a critical PLD-PLD interaction strength. Therefore, could interactions with domains or regions from other biopolymers also influence ordering within FUS proteins? To address this question we performed simulations with a 200-bead long, homopolymer mimicking RNA molecules that localize within RNP condensates. We first fix the interaction strengths for RNA and RBD beads ($\epsilon_{RBD-RNA}$) to 0.7 kT and then systematically vary the strength of promiscuous RNA binding to the PLD ($\epsilon_{PLD-RNA}$). The strength of $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$ was set to 0.5 kT, a scenario that results in partially ordered structures in the absence of RNA. The ratio of protein to RNA molecules was initially set at 3:1. In Fig.9, we show how RNA-PLD interactions can influence the extent of order in the RNA-protein condensates. In the limit of weak, RNA-PLD interactions, the RNA interactions are extremely specific, binding to the RBD alone, and the system displays partial ordering in PLD domains, akin to the scenario with no RNA molecules. As we increase the strength of non-specific RNA binding to the PLD from 0.1 to 0.3 kT, we see an initial decrease in order because RNA chains compete with PLD domains from other chains for PLD binding (reduction in contact ratio, Fig.9 green curve). However, a further increase in $\epsilon_{PLD-RNA}$ results in an increase in orientational order within these structures, with the PLD and the positively charged RBD showing alignment along the negatively charged RNA backbone.

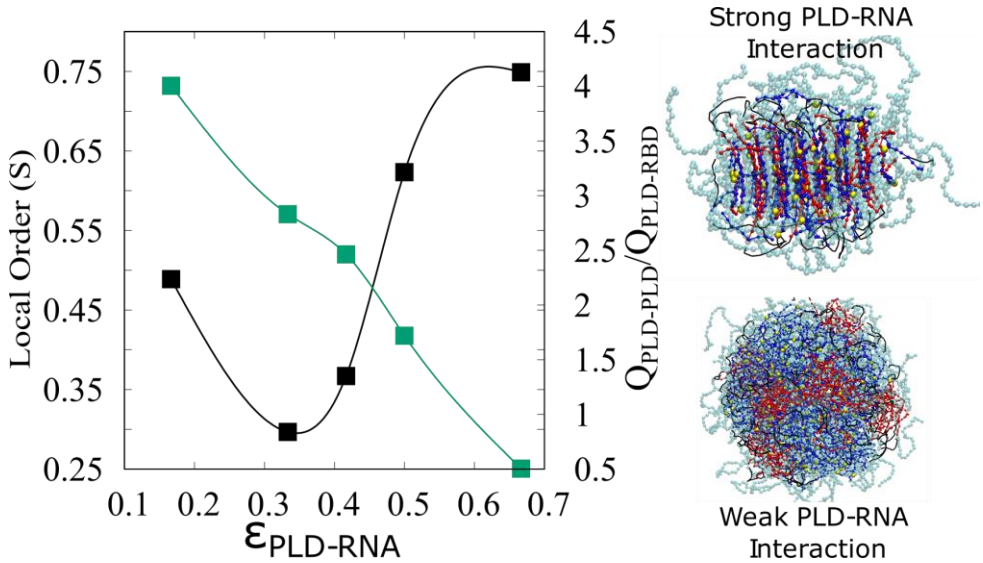

Figure 9: Ordering of FUS-RNA condensates in response to varying strengths of RNA-PLD binding. An initial increase in $\epsilon_{PLD-RNA}$ results in a decrease in orientational order and therefore disordered FUR-RNA assemblies (black curve). A futher increase in $\epsilon_{PLD-RNA}$, however, results in an increase in orientational order, with the PLD domains aligning themselves along the RNA spline. The intermolecular contact ratios (green curve) confirm that the order within the clusters is a result of RNA-PLD contacts and not PLD-PLD contacts.

We further probed the effect of RNA on on structural order in a broad range of RNA concentrations, for different regimes of protein-protein ($\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$) and protein-RNA interactions ($\epsilon_{PLD-RNA}$). When protein protein interactions are weak (Fig. 10A, $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$ = 0.4 kT), the structures that form in the absence of RNA (RNA/protein → 0 in Fig.10A and B) are disordered. In this regime, when RNA-protein interactions are weak ($\epsilon_{PLD-RNA} = 0.2\ kT$, Fig.10A) the RNA-protein assemblies remain disordered for all RNA concentrations under study. However, as RNA binding becomes stronger ($\epsilon_{PLD-RNA} = 0.4\ kT$, Fig.10B), we observe ordered structures for large RNA concentrations (Fig.10B) in the regime of weak protein-protein interactions. Therefore, for weak protein-protein and strong RNA-protein interactions, RNA molecules can scaffold ordered structures.

When PLD-PLD interactions are strong (Fig.10C), the structures observed in the absence of RNA display PLD-driven nematic order (Fig.10C, RNA/protein → 0 & Supplementary Fig.S1A). For an initial increase in RNA/protein ratio, we see a decrease in nematic order, due to a competition between RNA and PLD domains for PLD-PLD binding. A further increase in RNA/protein ratio results in an increase in nematic order (Fig.10C). However, the ordering in this regime is a result of RNA-protein interactions, with the RNA forming a backbone around which the ordered structures get stabilized (Fig.10C & Fig.S1B).

Therefore, the RNP-clusters exhibit liquid-like, disordered structures only when the strengthj of RNA-protein interactions is within a narrow range. On the other hand, interactions between the PLD domains of different, non-FUS proteins (stronger than a critical $\epsilon_{PLD-PLD}$) could also result in ordered structure formation making this an accessible self-assembled geometry for RBPs with intrinsically disordered regions.

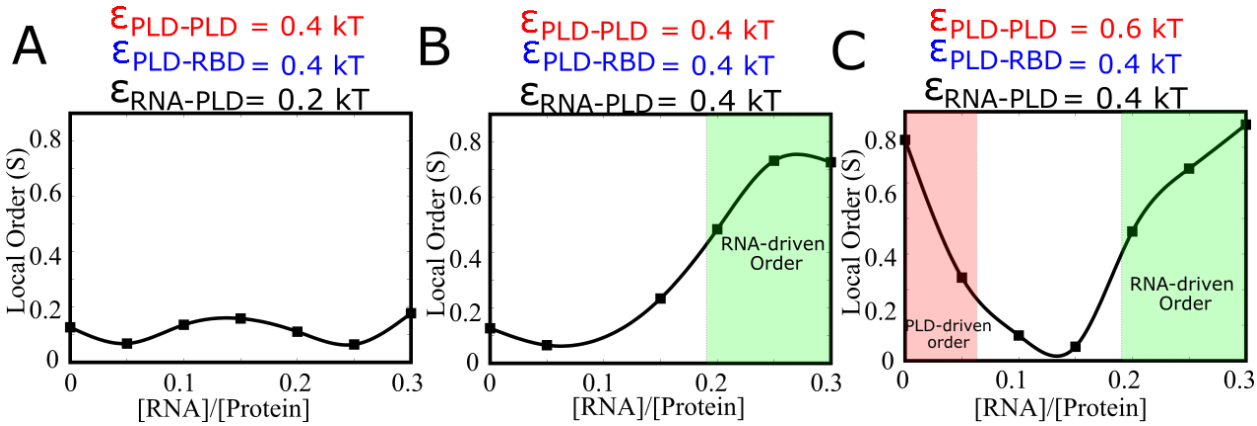

Figure 10: Ordering in FUS-RNA condensates, at varying RNA concentrations. A) When both protein-protein and protein-RNA interactions are weak, $\epsilon_{PLD-PLD}$ and $\epsilon_{PLD-RBD}$ < 0.5 kT and $\epsilon_{RNA-PLD} = 0.2\ kT$, the ensuing structures remain disordered throught the concentration range. B) As we increase RNA-protein interaction, $\epsilon_{RNA-PLD} = 0.4\ kT$, we see an RNA-driven ordering of clusters at high [RNA/Protein] ratios. C) When PLD-PLD interactions are strong, the structures that form in the absence of RNA are ordered. As we introduce RNA to the mixture, in this regime of interactions, there is an initial decrease in order. A further increase in RNA concentrations results in ordered RNA-protein clusters, with nematic order as a result of strong RNA-protein interactions. . The solid lines are guide to the eye.

Discussion

FUS and other RNA-binding proteins participating in LLPS exhibit a multi-valent, multi-domain architechture with a low complexity domain and an RNA-binding domain[44]. While the low-complexity, prion-like domain within FUS can phase-separate on its own, the threshold concentrations for LLPS of PLD alone are higher than that of the full-length protein, suggesting that a hierarchy of interactions could drive self-assembly[44]. The critical role played by the PLD can also be ascertained by a strong positive correlation between the number of Tyrosine residues in the PLD of FUS family proteins and their measured saturation concentrations for LLPS[16]. PLD domains, on their own can either exist in a disordered droplet like state or an ordered, cross-β rich amyloid-fibrillar state. This gives rise to an interesting paradox – while the low complexity domains are crucial drivers of phase separation, a high density of PLDs within the condensed phase makes these structures prone to fibril formation. Indeed, liquid-solid transitions have been widely reported in vitro and upon prolonged stress in vivo[36,37,45]. While the phenomenon of droplet aging and liquid-solid transition is ubiquitous in biological LLPS, its molecular underpinnings are largely unclear.

In this study, we aim to capture the structural diversity in FUS assemblies by exploring it within a narrow, biologically relevant range of interaction strengths and address a few key questions. Are ordered and disordered structural phases characterized by distinct interaction networks? Are morphological transitions accompanied by large changes in intra-droplet densities? Can post-translational modifications act as a barrier to ordered structure formation? To address these questions, we employ a coarse-grained, minimal representation of the multi-domain FUS protein with few interaction parameters in order to understand how the interplay between homo- and hetero-domain interactions could shape the structure of the FUS-rich

clusters. Interestingly, the minimal model can effectively capture ordered structures in the narrow range of interaction strengths ($\epsilon_{PLD-PLD}$ & $\epsilon_{PLD-RBD}$ in the range of 0.1 to 0.7 kT). Our simulations reveal that the disorder-order transition is accompanied by a switch in the predominant interaction network that stabilizes the self-assembled structures. While homo- and hetero-domain interactions are both equally likely in single-phase, disordered assemblies (Fig.2C, 4 & Fig.7D), the ordered structures are significantly enriched in PLD-PLD interactions (Fig.2C, 4). Crucially, we observe that this transition from the disordered to ordered self-assembled state has hallmarks of highly cooperative first order phase transition (Fig.3). Therefore, a small change in homodomain interaction strengths (≈ 0.1 kT per bead, Fig.3A,B and Fig.4) could result in dramatically altered structures. Mutations of single or small numbers of PLD residues, resulting in altered strength of interactions and/or interaction network could therefore cause disorder-order transitions in FUS protein droplets.

The isotropic-nematic transition to the self-assembled structures in our model occurs in the semi-flexible limit (L_{PLD} and $L_{RBD} >> L_p$) as opposed to the rod-like limit, indicating that the ordered phases are formed to maximaze favorable PLD-PLD interactions at the expense of competing less favorable PLD-RBD and RBD-RBD interaction. Interaction-driven ordering of polymer self-assemblies has been previously reported for homopolymer chains[29]. This mechanism is observed in polymer crystallization where higher density lamellae consisting of ordered chains intertwine with disordered regions[46,47]. Like in polymer crystallization, the loss of entropy in solid-like aged FUS assemblies due to local nematic ordering of the chains is compensated by stronger PLD-PLD attraction due to increase of local density of PLD monomers (Fig. 6A) (or PLD and RNA in case of RNA driven order, Fig.8 and Supplementary Fig. S1). A high local density of PLDs driven by stronger $\epsilon_{PLD-PLD}$ could result in the nucleation of nematic-like order in the condensed phase. Given that subtle changes to interaction strengths result in dramatically altered order in our simulations, these assemblies are sensitive to local density fluctuations that are most responsive to even slight variation of interaction parameters.

Further evidence that ordered structures are PLD-PLD interaction dependent comes from our results with phosphorylated PLDs that localize a high density of negative charges. Consistent with experimental findings[37], phosphorylation resulted in down-regulation of PLD-PLD contacts resulting in structures with low nematic order. Phosphorylation not only resulted in disordered structures (Fig.5) but also lower local densities (Fig.6B) of FUS PLD compared to assemblies composed of unmodified FUS. Also, ordered structures exhibited several-fold higher peak local densities compared to disordered phases, suggesting that these structures could also have vastly different material properties. Strikingly, the sharp density (and structural) transitions upon tuning the interaction parameters are also accompanied by dramatically altered intra-droplet dynamics, as shown by the intra droplet diffusion coefficients as well as polymer relaxation times (Fig.8 and Supplementary Fig.S2-S4). These results reveal that the different structural phases could also exhibit distinct material properties ranging from liquid-like (fast relaxation dynamics) to irreversible, trapped, solid-like (slow relaxation dynamics). While disordered structures exhibit liquid-like behavior (with diffusion coefficient 5-times lower than the dilute phase), the ordered

structures show a 50-fold reduction in diffusion coefficients compared to the dilute phase (Supplementary Fig.S2). This vastly altered dynamics is highly relevant to biological condensates, with irreversible phase transitions to the solid-like state being a signature of abberant behavior[48]. Indeed, FUS condensates have been identified to span several different material states, with ALS mutations accelerating their transition to the irreversible, solid-like state that results in impaired RNP-granule functions[23,48]. Our results therefore provide a minimal, mechanistic framework that link structural properties of the protein assembly to its material state.

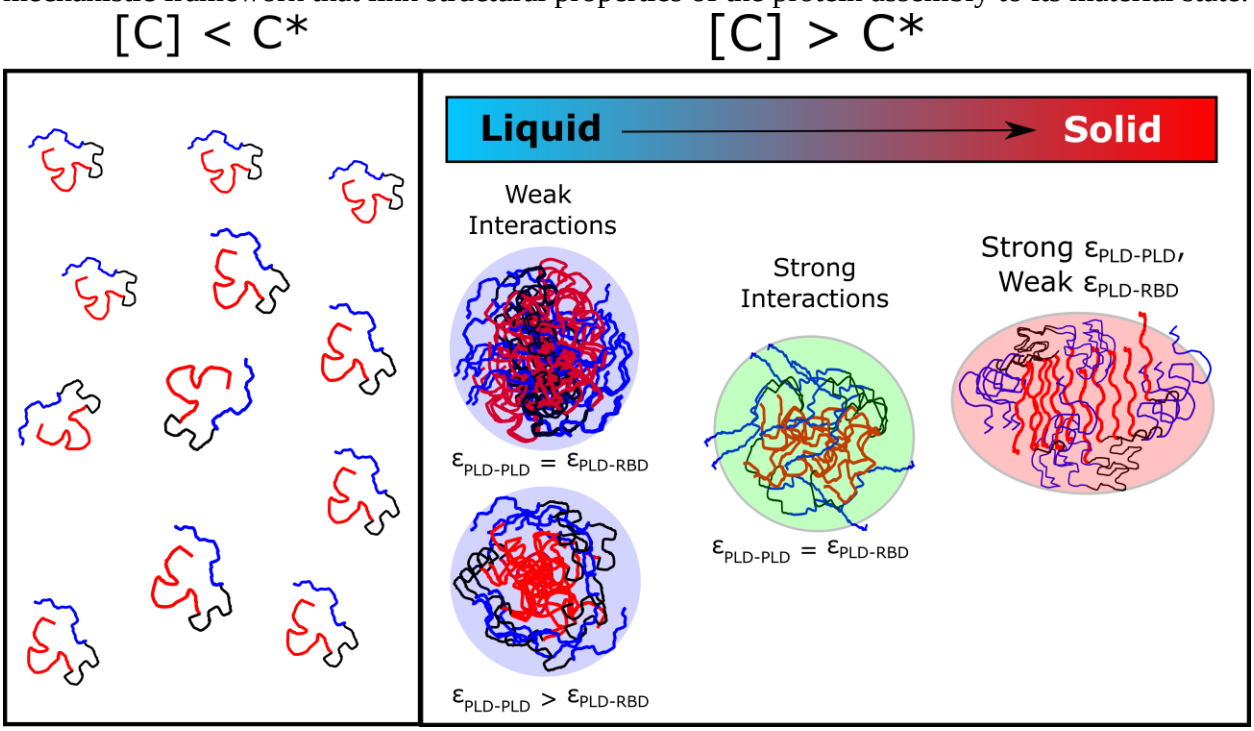

Figure 11: Graphical Summary of the findings. Below a critical concentration C*, the multi-domain RBPs remain in their soluble, fully-mixed state. However, above the critical concentration, these structures could exist in different structural, and material phases that are an outcome of the interplay between homodomain (PLD-PLD) and heterodomain (PLD-RBD) interactions.

The fact that a minimal model lacking residue-level resolution can capture diverse structural phases associated with protein condensates -- single-phase, core-shell disordered and nematically ordered structures -- suggests that the assembly of these structures could follow some generic physical mechanisms. While the current study employed a coarse-grained version of the FUS protein, the general principles for forming ordered, nematic-like structural phases could be applicable to other, related multi-domain RBPs that harbor intrinsically disordered domains (Figure. 11). The observation that a rich diversity of structural phases could exist for a narrow range of interaction strengths suggests that these transitions are amenable to tight regulatory control. Any deviation from this tight regulation could result in aberrant phase transitions that are an outcome of subtle changes to the interaction networks.

References

1. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285 (2017).
2. Choi, J. M., Holehouse, A. S. & Pappu, R. V. Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annual Review of Biophysics* (2020) doi:10.1146/annurev-biophys-121219-081629.
3. Standart, N. & Weil, D. P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. *Trends in Genetics* (2018) doi:10.1016/j.tig.2018.05.005.
4. Gonçalves, A. C. *et al.* Drug-induced Stress Granule Formation Protects Sensory Hair Cells in Mouse Cochlear Explants During Ototoxicity. *Sci. Rep.* (2019) doi:10.1038/s41598-019-48393-w.
5. Buchan, J. R., Yoon, J.-H. & Parker, R. Stress-specific composition, assembly and kinetics of stress granules in Saccharomyces cerevisiae. *J. Cell Sci.* **124**, 228–239 (2010).
6. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285 (2017).
7. Protter, D. S. W. & Parker, R. Principles and Properties of Stress Granules. *Trends Cell Biol.* **26**, 668–679 (2017).
8. Bond, U. Stressed out! Effects of environmental stress on mRNA metabolism. *FEMS Yeast Research* (2006) doi:10.1111/j.1567-1364.2006.00032.x.
9. Brangwynne, C. P., Mitchison, T. J. & Hyman, A. A. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proc. Natl. Acad. Sci.* **108**, 4334 LP – 4339 (2011).
10. Guillén-Boixet, J. *et al.* RNA-Induced Conformational Switching and Clustering of G3BP Drive Stress Granule Assembly by Condensation. *Cell* **181**, 346-361.e17 (2020).
11. Feric, M. *et al.* Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **165**, 1686–1697 (2016).
12. Molliex, A. *et al.* Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **163**, 123–133 (2015).
13. Maharana, S. *et al.* RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science (80-.).* (2018) doi:10.1126/science.aar7366.
14. Wang, J. *et al.* A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **174**, 688--699.e16 (2018).
15. Protter, D. S. W. & Parker, R. Principles and Properties of Stress Granules. *Trends Cell Biol.* **26**, 668–679 (2017).
16. Harmon, T. S., Holehouse, A. S., Rosen, M. K. & Pappu, R. V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife* **6**, e30294 (2017).
17. Luo, F. *et al.* Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. *Nat. Struct. Mol. Biol.* **25**, 341–346 (2018).
18. Burke, K. A., Janke, A. M., Rhine, C. L. & Fawzi, N. L. Residue-by-Residue View of

In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol. Cell* **60**, 231–241 (2015).
19. Guo, L. & Shorter, J. It's Raining Liquids: RNA Tunes Viscoelasticity and Dynamics of Membraneless Organelles. *Mol. Cell* **60**, 189–192 (2015).
20. Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2017).
21. Benayad, Z., Von Bülow, S., Stelzl, L. S. & Hummer, G. Simulation of FUS Protein Condensates with an Adapted Coarse-Grained Model. *J. Chem. Theory Comput.* **17**, 525–537 (2021).
22. Kar, M., Posey, A. E., Dar, F., Hyman, A. A. & Pappu, R. V. Glycine-Rich Peptides from FUS Have an Intrinsic Ability to Self-Assemble into Fibers and Networked Fibrils. *Biochemistry* (2021) doi:10.1021/acs.biochem.1c00501.
23. Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2015).
24. Dignon, G. L., Zheng, W., Kim, Y. C., Best, R. B. & Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput. Biol.* (2018) doi:10.1371/journal.pcbi.1005941.
25. Welsh, T. J. *et al.* Surface electrostatics govern the emulsion stability of biomolecular condensates. *bioRxiv* 2020.04.20.047910 (2020).
26. Garaizar, A. *et al.* Intermolecular reorganisation of single-component condensates during ageing promotes multiphase architectures. *bioRxiv* 2021.10.09.463670 (2021) doi:10.1101/2021.10.09.463670.
27. Harmon, T. S., Holehouse, A. S., Rosen, M. K. & Pappu, R. V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife* **6**, e30294 (2017).
28. Bellesia, G. & Shea, J.-E. Effect of β-sheet propensity on peptide aggregation. *J. Chem. Phys.* **130**, 145103 (2009).
29. Ranganathan, S., Maji, S. K. & Padinhateeri, R. Defining a Physical Basis for Diversity in Protein Self-Assemblies Using a Minimal Model. *J. Am. Chem. Soc.* **138**, 13911–13922 (2016).
30. Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
31. Laio, A. & Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports Prog. Phys.* **71**, 126601 (2008).
32. Selinger, J. V. *Introduction to the Theory of Soft Matter*. (Springer, Cham, 2016). doi:https://doi.org/10.1007/978-3-319-21054-4.
33. Seeber, M., Cecchini, M., Rao, F., Settanni, G. & Caflisch, A. Wordom: a program for efficient analysis of molecular dynamics simulations. *Bioinformatics* **23**, 2625–7 (2007).
34. Fujiwara, S. & Sato, T. Molecular Dynamics Simulation of Structural Formation of Short

Polymer Chains. *Phys. Rev. Lett.* **80**, 991–994 (1998).
35. Qamar, S. *et al.* FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. *Cell* **173**, 720-734.e15 (2018).
36. Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2017).
37. Monahan, Z. *et al.* Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* **36**, 2951–2967 (2017).
38. Mitrea, D. M. & Kriwacki, R. W. Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal.* **14**, 1 (2016).
39. Xing, W., Muhlrad, D., Parker, R. & Rosen, M. K. A quantitative inventory of yeast P body proteins reveals principles of composition and specificity. *Elife* (2020) doi:10.7554/eLife.56525.
40. Hofweber, M. & Dormann, D. Friend or foe—Post-translational modifications as regulators of phase separation and RNP granule dynamics. *J. Biol. Chem.* **294**, 7137–7150 (2019).
41. Dasmeh, P. & Wagner, A. Natural Selection on the Phase-Separation Properties of FUS during 160 My of Mammalian Evolution. *Mol. Biol. Evol.* **38**, 940–951 (2021).
42. Pestryaev, E. M. Comparison of various correlation times in polymer melts by molecular dynamics simulation. *J. Phys. Conf. Ser.* **324**, (2011).
43. Kalathi, J. T., Kumar, S. K., Rubinstein, M. & Grest, G. S. Rouse mode analysis of chain relaxation in polymer nanocomposites. *Soft Matter* **11**, 4123–4132 (2015).
44. Wang, J. *et al.* A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **174**, 688-699.e16 (2018).
45. Xu, Q., Teplow, D., Lee, T. D. & Abelson, J. Domain structure in yeast tRNA ligase. *Biochemistry* **29**, 6132–6138 (1990).
46. Olmsted, P. D., Poon, W. C. K., McLeish, T. C. B., Terrill, N. J. & Ryan, A. J. Spinodal-Assisted Crystallization in Polymer Melts. *Phys. Rev. Lett.* **81**, 373–376 (1998).
47. Strobl, G. R. *The Physics of Polymers*. (Springer, Berlin, Heidelberg, 1996). doi:https://doi.org/10.1007/978-3-662-03243-5.
48. Murakami, T. *et al.* ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **88**, 678–690 (2015).

Supplementary Figures

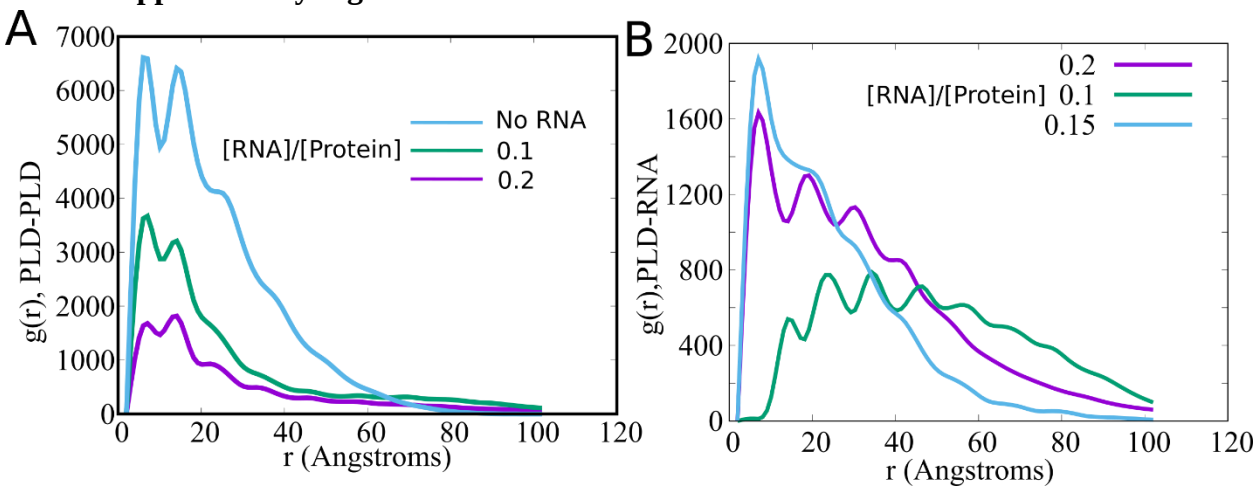

Figure S1. Evolution of pair distribution function **g(r),** for increasing concentrations of RNA. (A) g(r) curves for PLD-PLD contacts, and (B) for PLD-RNA contacts.

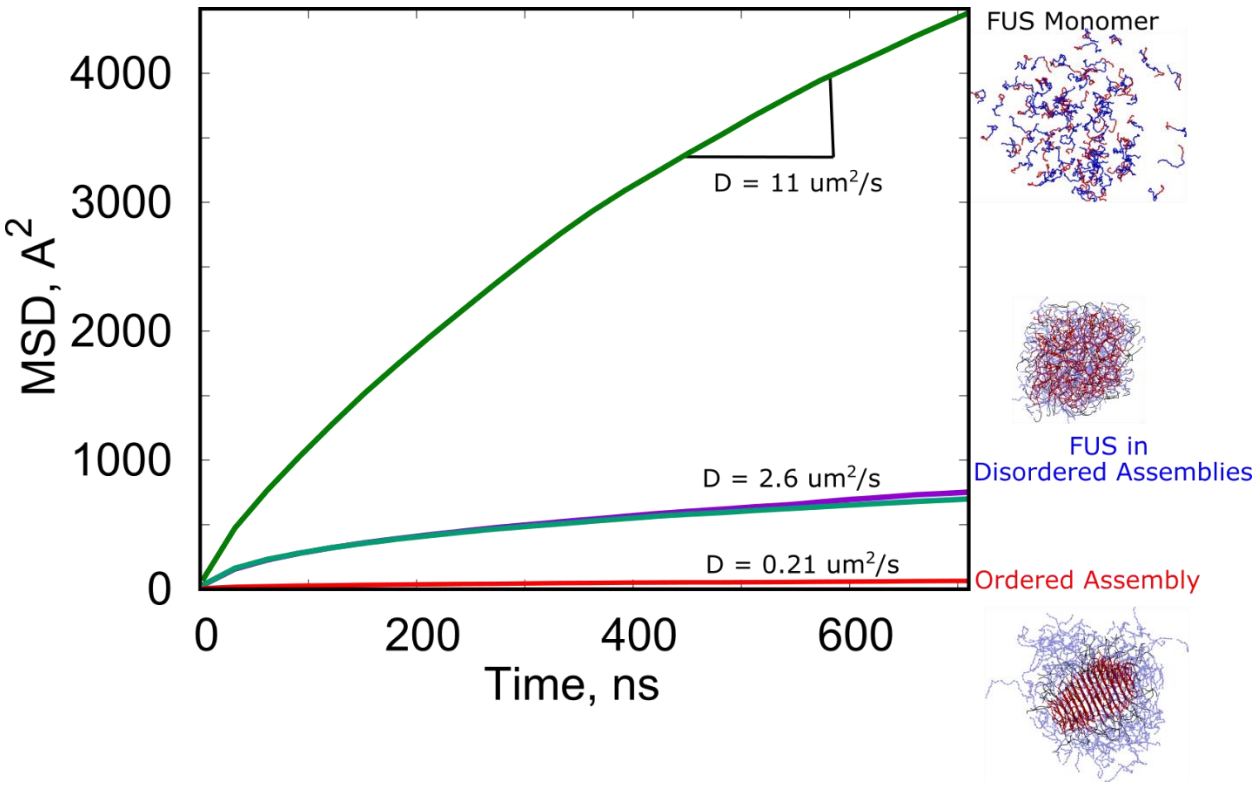

Figure S2. Mean square displacement profiles for FUS chains in different scenarios – in the monomeric state (green curve), within single and multi-phase disordered droplets (blue curve) and ordered assemblies (red curve).

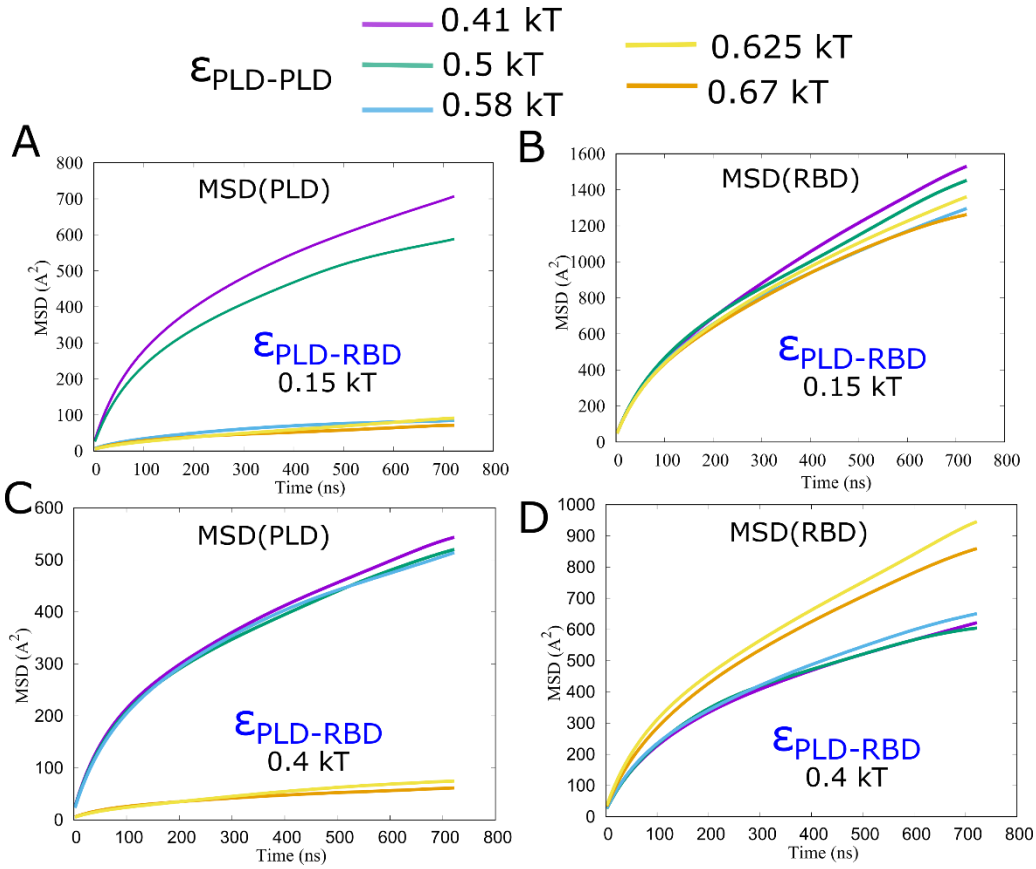

Figure S3. Mean Square Displacement profiles for the PLD (A and C) and RBD (B and D) domains, for different strengths of PLD-RBD (0.15 kT in panel A and B) and PLD-PLD interactions (different colored curves in panels A-D).

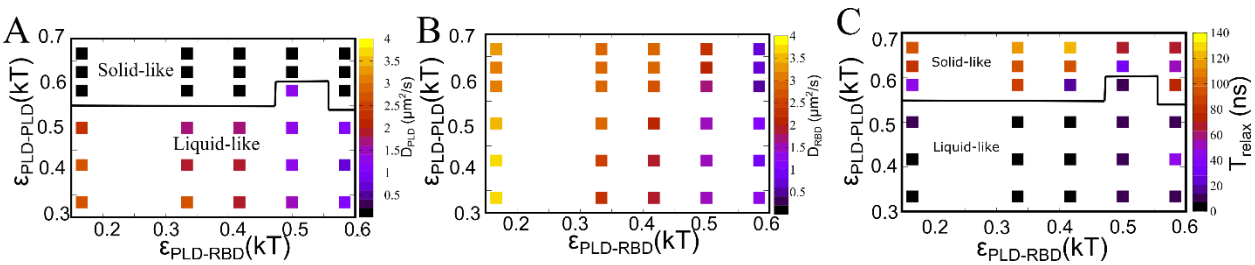

Figure S4. Phase diagrams. A) Diffusion coefficient for PLD domains showing dramatically lower values at higher values of PLD-PLD interactions. B) The RBD regions remain liquid like in the whole range of parameters studied. C) Relaxation times for polymer chains within different structural phases, measured based on the normalized autocorrelation functions of the radii of gyration of polymers.