Intermolecular reorganisation of single-component condensates during ageing promotes multiphase architectures

Adiran Garaizar[1¶], Jorge R. Espinosa[1¶], Jerelle A. Joseph[1,2,3], Georg Krainer[1,2], Yi Shen[1,2], Tuomas P.J. Knowles[1,4,*], and Rosana Collepardo-Guevara[1,2,3,*]

[1]Maxwell Centre, Cavendish Laboratory, Department of Physics, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE
[2]Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK
[3]Department of Genetics, University of Cambridge, Cambridge, CB2 3EH
[4]Centre for Misfolding Diseases, Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK
[*]rc597@cam.ac.uk,tpjk2@cam.ac.uk
[¶]These authors contributed equally to this work

ABSTRACT

Intracellular proteins can undergo phase separation to form liquid-like biomolecular condensates with a multitude of functional roles. Liquid condensates can, however, further age and progressively rigidify. In addition to single-phase systems, multiphase condensates are increasingly identified commonly within multi-component systems, where the different molecular components present sufficient physicochemical diversity to sustain separate phases. Here, we develop a multiscale modelling approach that predicts conditions under which multiphase architectures can arise also within single-component protein condensates. Such single-component condensates are initially homogeneous but become heterogeneous over time due to the gradual enhancement of interprotein interactions. We find that such enhancement could originate, for instance, from intermolecular disorder-to-order transitions within low-complexity aromatic-rich kinked segments in the prion-like domain of FUS. Our model reveals that as increasing numbers of molecules undergo a disorder-to-order transition over time, single-component protein condensates convert into either gel-core/liquid-shell or liquid-core/gel-shell multiphase structures, depending on the relative surface tension of the liquid and gel phases. Despite being formed by proteins that are chemically-identical, the different liquid and gel phases present diverse surface tensions due to their fundamentally different molecular organization. Our study highlights the regulatory role of prion-like domains in tuning condensate behavior and, more generally, suggests a new route by which multilayered compartments or hierarchically organized condensate structures can emerge.

Introduction

Cells compartmentalize their interiors and regulate critical biological functions using both membrane-bound organelles and non-membrane-bound structures, biomolecular condensates[1–3]. Condensates are ubiquitous mesoscopic assemblies of biomolecules that demix from the cytoplasm or nucleoplasm through liquid–liquid phase separation[1–5]. A dominant factor that enables biomolecules to phase separate is their multivalency: their ability to form multiple weak associative interactions[6–8]. Although not always present among phase-separating proteins[9], intrinsically disordered regions (IDRs)—often characterized by (amino acid sequences of low-complexity termed) low complexity domains, LCDs[10,11]—have emerged as important contributors to the multivalency and phase separation capacity of many naturally occurring proteins[2,3,6]. When they have a high aromatic content, such as that found in prion-like domains (PLDs)[12,13], can furthermore promote aggregation (fibrillation) and amyloid formation[10,11,14]. Protein condensates that contain PLDs—for instance, the RNA-binding proteins Fused in Sarcoma (FUS), TAR DNA-binding protein 43 (TDP-43) and heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1)—can undergo a further phase transition from functional liquid-like states to less dynamic reversible hydrogel structures or even irreversible gel-like states sustained by fibrillar aggregates[11,15–23]. The transitions from liquid-like condensates to gel-like structures are implicated in neurodegenerative diseases such as amyotrophic lateral sclerosis and frontotemporal dementia[24–26].

Both within liquid-like and gel-like condensates, proteins interconnect to form percolated networks[27,28]. However, gels can be distinguished from liquids by the prevalence of long-lived inter-protein interactions, which confer local rigidity to the network[27,29]. The loss of the liquid-like character of a condensate over time during ageing[15,25] can be modulated by amino-acid sequence mutations, post-translational modifications, application of mechanical forces, and protein structural transitions, among

many other factors[15–17,20–22,30–32]. These factors are expected to, directly or indirectly, increase the proportion of long-lived protein interactions within the condensate, which gives rise to complex mesoscale properties and network rigidity[27,33].

Besides being highly diverse in terms of their material properties, condensates also vary significantly in their internal architectures[34]. Multi-component condensates can present various internal coexisting phases. The nucleolus[35], paraspeckles[35–37], and stress granules[38,39], are all examples of hierarchically organized condensates with multiple coexisting phases or layers. Recently, intranuclear droplets combining a dense liquid spherical shell of acetylated TDP-43—with decreased RNA-binding affinity—and an internal liquid core rich in HSP70 chaperones were observed[40]. Multicomponent multiphase condensates can also present internal low-density 'bubbles' or a 'hollow' space surrounded by an outer, denser phase[41–43]. Examples of these include the germ granules in Drosophila[41], the condensates formed from intracellular overexpression of TDP-43[42], and in vitro RNA-protein vesicle-like condensates[43].

The multiphase behavior of condensates has been recapitulated in vitro, for instance in the liquid-core-gel-shell condensates formed by the Lge1 and Bre1 proteins[44], different multiphasic complex coacervates[45,46], and multilayered RNA-protein systems[47,48]. In all these cases, the emergence of a multilayered or multiphase organization is connected to the diversity in the properties of the biomolecular components within these condensates[28,44,46,49]. Physicochemical diversity is key as it allows subsets of components to establish preferential interactions with one another leading to segregation into multiple layers inside the condensates, which are typically ordered according to their relative interfacial free energies[50]. Indeed, simulations and mean-field theory have shown that multi-component mixtures of species with sufficiently different valencies and/or binding affinities[51] are likely to segregate into multiple coexisting liquid phases with different compositions[52] or form multilayered architectures[46,49,50]. In agreement with this idea, competing interactions among protein–RNA networks have been shown to drive the formation of multiphase condensates with complex material properties[51]. Experiments and simulations have further demonstrated that the different condensed liquid phases within multiphase condensates are hierarchically organized according to their relative surface tensions, critical parameters, viscosities, and densities[35,46,49,50]. Motivated by these observations, here we explore the fundamental question of whether or not is possible for single-component condensates to sustain multiphasic architectures.

Progressive dynamical arrest of proteins has been observed in vitro in single-component condensates for proteins with PLDs marked by low-complexity aromatic-rich kinked segments (or LARKS)[8,10–20,53]. In these cases, formation of inter-protein β-sheets by LARKS peptides drives gradual fibrillization at the high protein concentrations present within condensates[18,19,23]. In such a situation, rather than establishing only weak, transient attractive interactions, multiple LARKS can assemble into inter-locking structures, which are stabilized by hydrogen bonds and π–π interactions, leading to the formation of reversible, crossed-β-sheet amyloid structures[20]. Such a disorder-to-order transition within condensates highlights how subtle changes in the local behaviour of proteins can result in large-scale transformations of the mesoscale condensate structure.

To probe whether disorder-to-order protein transitions may be sufficient to transform single-component homogeneous condensates into multiphase architectures, we focus on the naturally occurring phase-separating protein FUS. FUS exhibits a liquid to gel transition during ageing[15,30], and contains LARKS regions that can undergo a disorder-to-order transition to form β-sheet rich structures. To capture this behaviour, we develop a multiscale modelling approach—combining atomistic simulations of peptides, with two levels of coarse-grained simulations of protein condensates—that allows us to study key implications of a disorder-to-order transition (i.e., local enhancement of inter-protein interactions and rigidification) on the organization of FUS condensates. Our simulations predict that FUS condensates can transition from homogeneous condensed liquid phases into liquid-core-gel-shell multiphase architectures due to the gradual enhancement in local protein interactions, consistent with an increase in cross-β-sheet content within LARKS regions of FUS proteins. Our simulations also demonstrate that the molecular organisation of the gel-like phase (e.g. whether the arrested state exposes more or less hydrophobic regions of the protein to the surface) significantly influences the relative interfacial free energy of the various protein phases; hence, determining the ordering of the various phases within multiphase condensates. These findings highlight how variations in the binding strengths among proteins, due for instance to secondary structural changes, can modulate mesoscale phase behavior of proteins and form multiphase condensates in single-component protein systems. These results corroborate the idea that PLDs can act as modulators of prion-like protein phase behavior and, thereby, can tune collective interactions among adhesive amino acid motifs that result in condensate structural transformations that are observable at the mesoscale[12,13].

Results

Multiscale model of single-component multiphase condensates

A hallmark of ageing of RNA-binding proteins is the ability to undergo structural transitions within their LARKS. It has been observed[19] that LARKS within FUS (e.g., 37SYSGYS42, 54SYSSYGQS61 and 77STGGYG82) can form pairs of kinked cross-β-sheets, which assemble into ladders and yield reversible fibrils that sustain FUS hydrogels. Peptides at each step of the LARKS ladder form hydrogen bonds with adjacent peptides in the next step of the ladder. In addition, stacking of aromatic sidechains stabilize both the ladder and the individual β-sheets at each step, due to intermolecular PLD–PLD interactions. We

therefore hypothesized that the formation of cross-β-sheets among these critical LARKS in FUS-PLD could introduce sufficient physicochemical diversity into single-component FUS condensates to induce a multiphase organization. To test this hypothesis and provide mechanistic insights into the physical and molecular determinants of the putative multiphase organization, we designed a multiscale simulation approach (see Figure S1).

Molecular simulations are powerful in dissecting the mechanisms, driving forces, and kinetics of phase separation and providing structural details of biomolecules within condensates, which are generally challenging to describe and interrogate using experimental techniques[54,55]. All-atom, coarse-grained, and multi-scale modelling approaches are now well-established tools, used in conjunction with experiments, to investigate biomolecular phase separation[28,56–60]. Our multiscale modelling strategy integrates atomistic molecular dynamics (MD) simulations of FUS LARKS (Step 1; Figure S1), amino-acid resolution coarse-grained simulations of full length FUS proteins in implicit solvent (i.e., using our Mpipi model[61], which describes with near-quantitative accuracy the phase behaviour of proteins; Step 2; Figure S1), and a bespoke minimal coarse-grained model of FUS proteins in explicit solvent that considers the accumulation over time of model disorder-to-order transitions (i.e. mimicked by the gradual emergence of stronger inter-protein bonds and local rigidification of the interacting segments) inside the condensate (Steps 3–4; Figure S1).

A disorder-to-order structural transition diversify the interactions among chemically-identical FUS proteins

First, we use atomistic MD simulations to compute the differences in interaction strength of FUS-PLD LARKS due to disorder-to-order transitions (Step 1; Figure S1). As reported previously for Aβ-peptides[62], we perform atomistic umbrella sampling MD simulations to quantify the changes in the relative binding strengths (or potential of mean force; PMF) among LARKS peptides when they are disordered versus when they stack to form the kinked β-sheet structures resolved crystallographically (Figure 1a)[19]. We compute the intermolecular binding strengths for three different FUS LARKS sequences (SYSGYS, SYSSYGQS, STGGYG). For each case, we calculate the free energy cost of dissociating a single peptide from a system containing four identical peptides. When the four LARKS remain fully disordered, their binding interaction strengths is weak ($\sim$2–5 $k_{\mathrm{B}}T$), which suggests that thermal fluctuations can frequently break these interactions (Figure 1b–d). However, the interaction strength increases significantly, by a factor of $\sim$4–8 (i.e., by 15–40 $k_{\mathrm{B}}T$ in total, depending on the LARKS sequence), when the peptides display canonical stacking, forming kinked β-sheets with two LARKS per step (Figure 1b–d). A summary of the relative interaction strengths for the three LARKS sequences is provided in Figure 1e. Importantly, while this result suggests that LARKS–LARKS interactions within kinked β-sheets are sufficiently strong to sustain thermal fluctuations—and hence can give rise to hydrogels—they remain sufficiently weak for the ensuing hydrogels to dissolve upon salt treatment or heating. This prediction is consistent with the observation that such β-sheets give rise to reversible hydrogels, rather than irreversible amyloid fibrils. This finding also suggests that in aged FUS condensates, where LARKS have undergone a disorder-to-order transition, intermolecular interactions among PLDs can be significantly strengthened, due to ordering and stacking into β-sheet fibrillar-like ladders.

Based on the striking differences in the binding strength between FUS LARKS with β-sheets and those where the peptides remained fully disordered, we next hypothesized that a binary protein mixture composed of two distinct FUS conformational ensembles (i.e., with ordered *versus* with disordered LARKS) within the same condensate might give rise to a multiphase condensate morphology. In other words, we reasoned that two-phase coexistence within a single-component FUS condensate may emerge when a fraction of the FUS proteins have transitioned from having fully disordered PLDs to having PLDs that form inter-protein cross-β-sheets).

Strengthening of PLD interactions dramatically transforms the molecular organization of FUS condensates

Next, we used amino-acid resolution coarse-grained simulations to determine whether strengthening of selected inter-protein bonds due, for instance, to an intermolecular disorder-to-order transition, gives rise to FUS condensates with sufficiently different mesoscale properties from those of FUS condesates with standard interactions.

First, we use our Mpipi residue-resolution coarse-grained model[61] (Step 2; Figure S1), which recapitulates experimental LLPS propensities of FUS mutants[10], to characterize the molecular organization of single-component FUS condensates with standard (i.e., fully disordered and weakly interacting) PLD regions. Next, to probe the differences in the phase behavior of FUS proteins with PLDs that establish stronger inter-protein interactions, we increase the PLD–PLD interactions as suggested by our atomistic PMFs, to approximate one of the key repercussion of inter-protein β-sheet formation. A representation of these residue-resolution coarse grained models for FUS are shown in Figure 2a, with the 526-residue FUS polypeptide chain sectioned into its PLD region (residues 1–165), three disordered RGG rich regions (RGG1: residues 166–267; RGG2: residues 371–421; and RGG3: residues 454–526), and two globular regions (an RNA-recognition motif (RRM): residues 282–371; and a zinc finger (ZF) domain: residues 422–453).

For each of these coarse-grained parameterizations, we conduct residue-resolution direct coexistence simulations[63–65] of tens of interacting full-length FUS molecules and estimate the influence of modulating the intermolecular interactions among

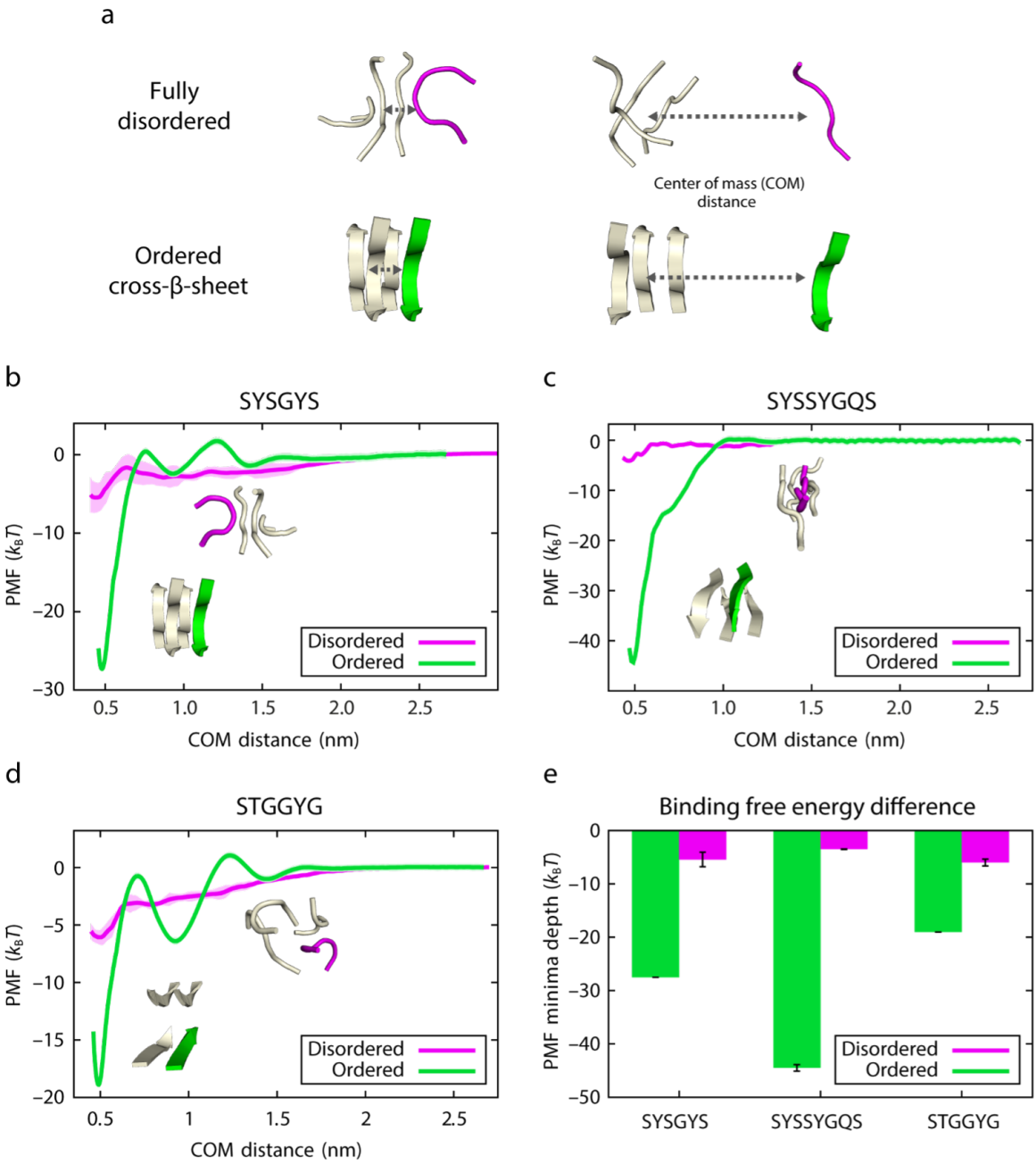

Figure 1. Disorder-to-order transition strengthens intermolecular interactions among FUS LARKS. (a) Schematic of atomistic PMF simulations of LARKS-forming peptides in their disordered (magenta) and ordered states (green). The starting configurations for all simulations consisted of four stacked peptides. PMFs are calculated by estimating the free energy needed to remove one peptide (the 'dissociating' peptide) from the stack as a function of the centre of mass (COM) distance. The example shown is the peptide SYSGYS (PDB code: 6BWZ). (b–d) Plots of PMF *versus* COM for homotypic pairs of FUS LARKS-forming peptides SYSGYS (PDB code: 6BWZ), SYSSYGQS (PDB code: 6BXV) and STGGYG (PDB code: 6BZP), respectively, before (magenta) and after (green) undergoing the disorder-to-order structural transition. (e) Variation in the free energy minimum (as obtained from the profiles in panel b–d)).

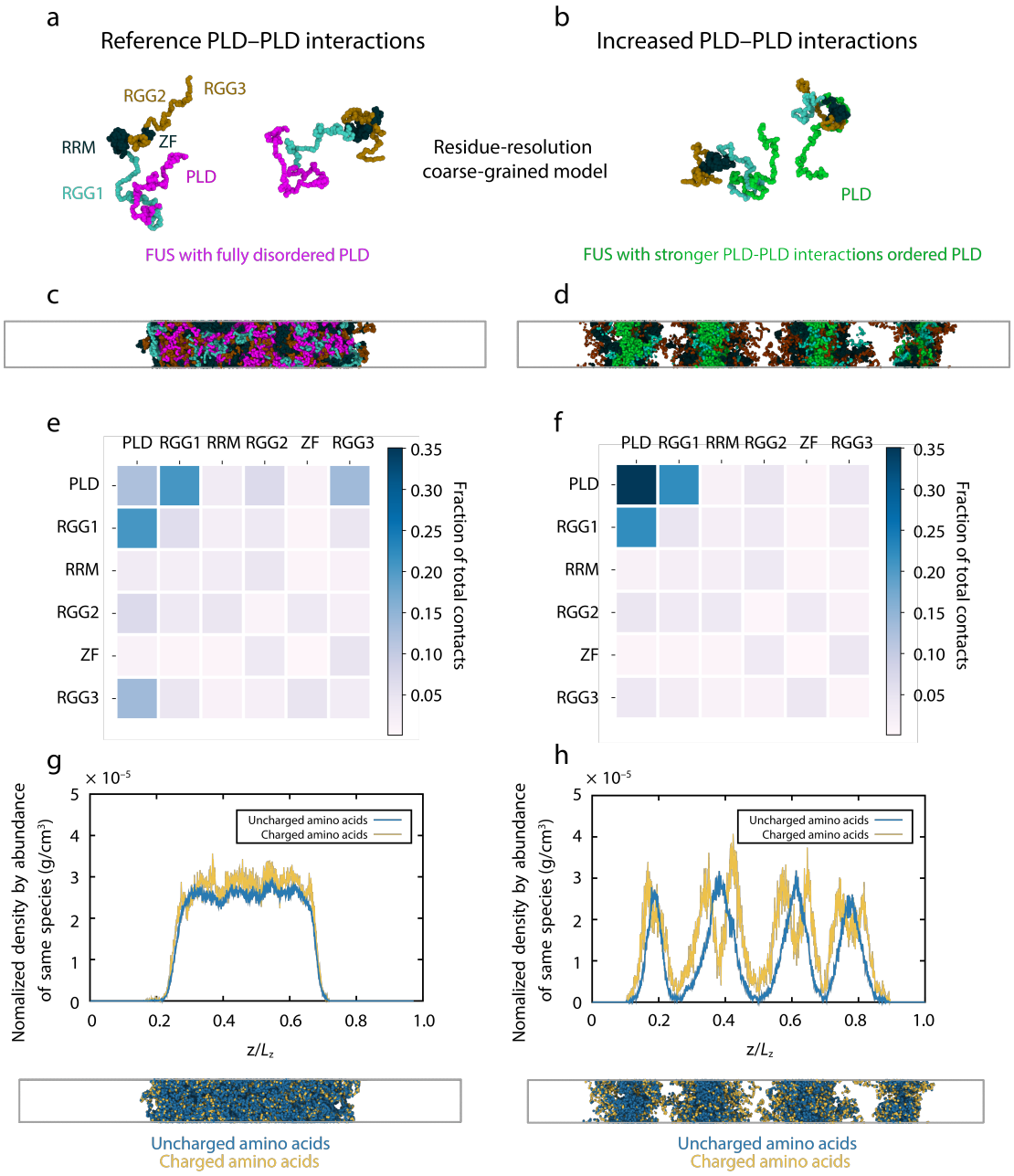

Figure 2. Strengthened PLD interactions give rise to multiphase condensate organization in coarse-grained molecular dynamics simulations. (a, b) Residue-resolution coarse-grained models for FUS with fully disordered PLD (panel a) and with ordered PLD (i.e., with cross-β-sheet elements in the PLDs) (panel b). Representative snapshots of FUS replicas, as obtained via direct coexistence MD simulations, are shown. Amino acid beads are colored according to the domains of FUS, with one bead representing each amino acid: PLD (residues 1–165): magenta (panel a) or green (panel b); RGG1 (residues 166–284): cyan; RGG2 (residues 372–422) and RGG3 (residues 454–526): ochre, RRM (residues 285–371) and ZF (residues 423–453): dark blue. (c, d) Snapshots of direct coexistence simulations with reference interaction strengths among PLDs (panel c) and increased interactions strengths among PLDs (panel d). 48 FUS proteins were included in the simulations. Color code as in panel a and b. (e, f) Frequency of contacts between FUS domains within condensates for simulations with standard interaction strengths among PLDs (panel e) and increased interactions strengths among PLDs (panel f). Heatmaps are color coded and scaled from white to dark blue. (g, h) Normalized density of charged (yellow) and uncharged (blue) species across the long side of the simulation box estimated over the coarse-grained equilibrium ensemble for simulations with standard interaction strengths among PLDs (panel c) and increased interactions strengths among PLDs (panel d). The snapshots from direct coexistence simulations (bottom) are the same as in panel c and d, but now colored-coded according to the charge state of the amino acid residues.

PLDs on FUS phase separation (Step 2; Figure S1). The direct coexistence method enables simulating a protein-enriched condensed liquid phase in contact with a protein-poor diluted-liquid phase in one simulation box and, thus, can determine whether a system phase separates at a given set of conditions.

Consistent with previous observations[66], FUS condensates formed by standard proteins with fully disordered PLDs display a homogeneous molecular architecture (Figure 2c,g) (i.e., all FUS domains are randomly positioned throughout the condensate). In contrast, FUS condensates containing strengthened PLD–PLD interactions (Figure 2d,h) exhibit a micelle-like heterogeneous organization with a PLD-rich hydrophobic core and a charge-rich interface; i.e., the positively charged RGG2 and RGG3 regions are effectively exposed to the solvent, which in turn lowers the droplet surface tension as compared to the homogeneous condensates formed by standard FUS proteins [γ ordered FUS $= 0.06 mJ/m^2$ *vs.* (γ standard FUS $= 0.16 mJ/m^2$)]. The reduction in the surface tension of FUS condensates with stronger PLD–PLD interactions is also evident from the ability of these proteins to stabilize multiple droplets in the same simulation box (Figure 2d,h).

Presence of distinct FUS structural ensembles supports the formation of hollow condensates

Next, we develop a minimal model (Step 3; Figure S1) that significantly reduces the degrees of freedom needed to describe a solution of FUS proteins, while retaining essential physicochemical information, and that allows us to simulate the large number of proteins needed to simulate multicomponent protein system; i.e., a larger number of proteins reduces the finite-size effects emerging in the treatment of multicomponent systems (e.g., those related to the relative number of copies of each component in the system). Our minimal model, therefore, enables the simulation of mixtures containing the two types of FUS proteins we are interested in. In this model, full-length FUS proteins are each represented as a chain of 20 beads (i.e., 6 beads for FUS-PLD, and 14 beads for the RGG1, RRM, RGG2, ZF, and RGG3 regions).

To distinguish between FUS proteins with fully disordered PLDs (termed 'disordered' FUS) and FUS proteins where the PLDs form inter-protein cross-β-sheets (termed 'ordered' FUS), our minimal model considers the following three key physicochemical differences: (1) While disordered FUS proteins are modelled as fully flexible chains, we increased the local rigidity among PLD beads within ordered FUS proteins to mimic the structural effect of cross-β sheet formation. (2) We parameterize the strengths of protein–protein interactions depending on whether they occur between disordered or ordered FUS protein regions. Specifically, for each type of FUS protein (i.e. 'disordered' and 'ordered'), we set the relative interaction strengths among pairs of FUS domains (e.g., PLD–PLD, PLD–RGG1, RGG1–RGG1, etc.) to match the relative frequencies of interactions obtained in their respective residue-resolution simulations; i.e., the disordered FUS minimal model according to the contact maps of the single-component standard FUS simulations (Figure 3a), and the ordered FUS model based on the contact maps of our single-component FUS simulations where the PLD–PLD interactions were increased based on our atomistic results (Figure 3b) . (3) We assign a higher hydrophilicity to the charged-rich FUS regions (i.e., RGG1, RGG2 and RGG3 regions) than to the other FUS domains, and incorporate a compatible minimal explicit solvent model; this allows us to recover the lower surface tension we observed in our residue-resolution simulations for single component ordered FUS condensates.

Using these two minimal FUS protein models, we perform direct coexistence simulations of a mixture of 50 mol % 'disordered' FUS proteins and 50 mol% 'ordered' FUS proteins in explicit water/solvent. By starting from a homogeneous mixed phase (Figure 3c), the system readily demixed into a phase-separated condensate with an inhomogeneous organization (Figure 3d). Evaluation of the density profile of the equilibrated system reveals the formation of a liquid-core-gel-shell (i.e., hollow) multiphase condensate architecture; i.e., condensates are hierarchically organized with a low-density core phase, made up by 'disordered' FUS proteins, which is surrounded by a high-density shell composed of 'ordered' FUS species (Figure 3e). The lower density of the inner FUS phase with disordered PLDs is evident from its higher water content. Our simulations further reveal that FUS proteins in the outer ordered-FUS shell are less mobile than in the liquid core as gauged by their reduced mean squared displacement (Figure S2). These findings predict that increased PLD interactions due to the ordering and stacking into kinked cross-β-sheet fibrillar-like ladders of FUS LARKS can drive multiphase condensate formation in a single-component protein system.

Ageing simulations predict FUS multiphase condensate formation

The highly concentrated environments of biomolecular condensates may serve as seeding grounds for disorder-to-order transitions that occur in FUS-PLD LARKS (Step 4; Figure S1). To gain deeper insights into the mechanism of FUS multiphase condensate formation, we developed a dynamical algorithm that approximates the effects (i.e., strengthening of inter-protien bonds, local protein rigidification, and changes in the molecular organization of the condensed phase) of the gradual accumulation of inter-protein β-sheet structures in a time-dependent manner and as a function of the local protein density. Rather than imposing a fixed concentration of 'disordered' and 'ordered' FUS proteins in the simulation a priori, this algorithm enabled us to study the spontaneous emergence of multiphase FUS condensates during ageing.

We again perform direct-coexistence simulations using our minimal coarse-grained FUS model starting from a system containing 100 mol% 'disordered' FUS proteins (i.e., with interactions mimicking disordered PLDs) and explicit solvent. As time progresses, our algorithm triggers disorder-to-order transitions within the PLD of FUS when high local fluctuations of

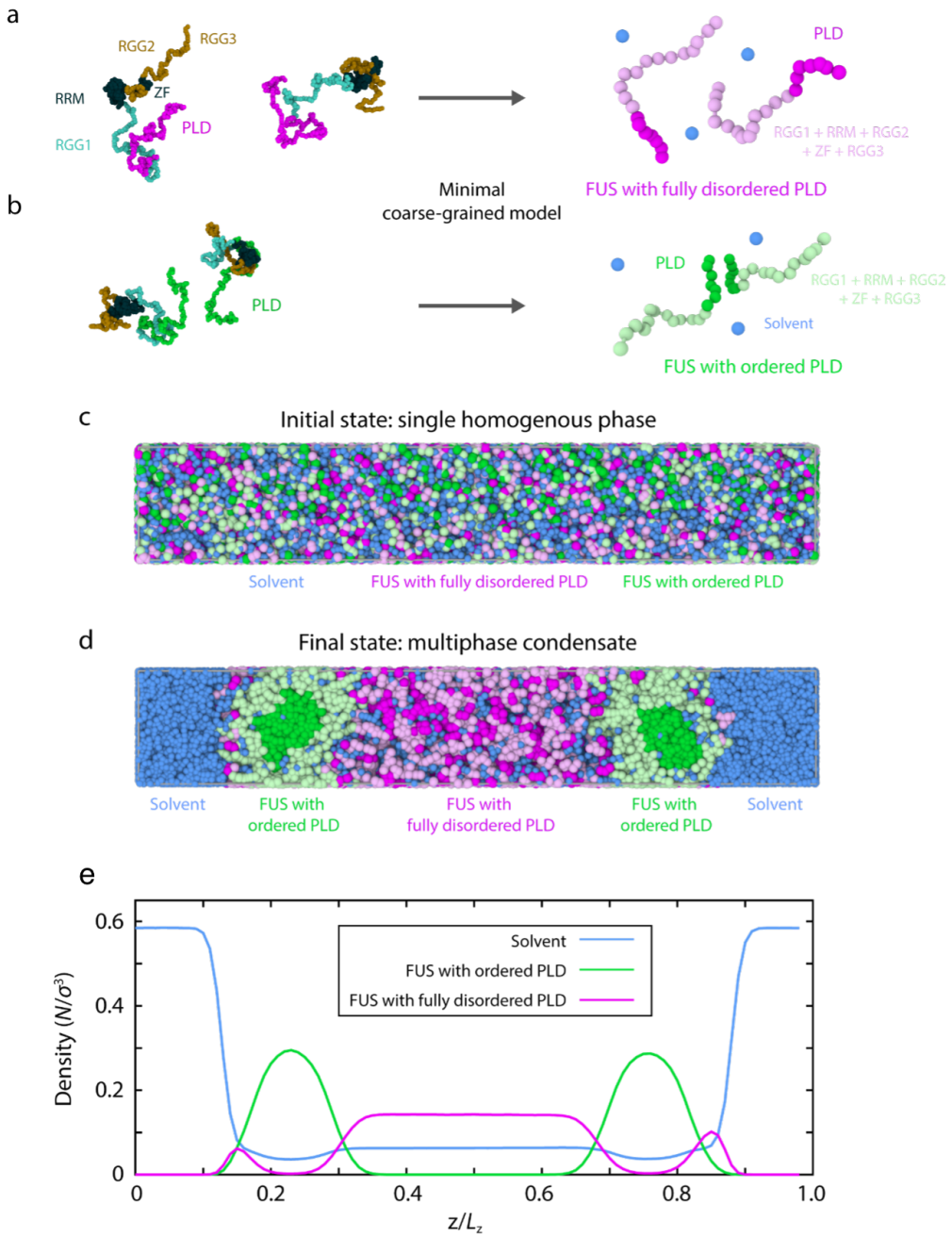

Figure 3. A disorder-to-order transition within PLDs drives multiphase organization in FUS condensates. (a, b) Minimal coarse-grained models for FUS with fully disordered PLDs (panel a) and with ordered PLDs (i.e., with cross-β-sheet elements in the PLDs) (panel b). Representative snapshots of FUS replicas, as obtained via direct coexistence MD simulations, are shown. Here one bead represents 26 amino acids. PLD: magenta (panel a) or green (panel b), RGG1, RRM, RGG2, ZF, RGG3: light magenta (panel a) or light green (panel b). Solvent (water) is depicted by blue beads. (c, d) Snapshots of direct coexistence simulations of a mixture of 50% FUS proteins with fully disordered PLDs and 50% FUS proteins with cross-β-sheet elements in their PLDs and explicit solvent. The initial state is shown in panel c and the final state in panel d. 1088 FUS proteins were included in the simulations. (e) Density profile (in reduced units) of FUS species and explicit solvent across the long side of the simulation box estimated over the coarse-grained equilibrium ensemble (as obtained in panel d). FUS proteins with fully disordered PLDs: magenta; FUS proteins with ordered PLDs (i.e., with kinked cross-β-sheets): green, solvent (water): blue.

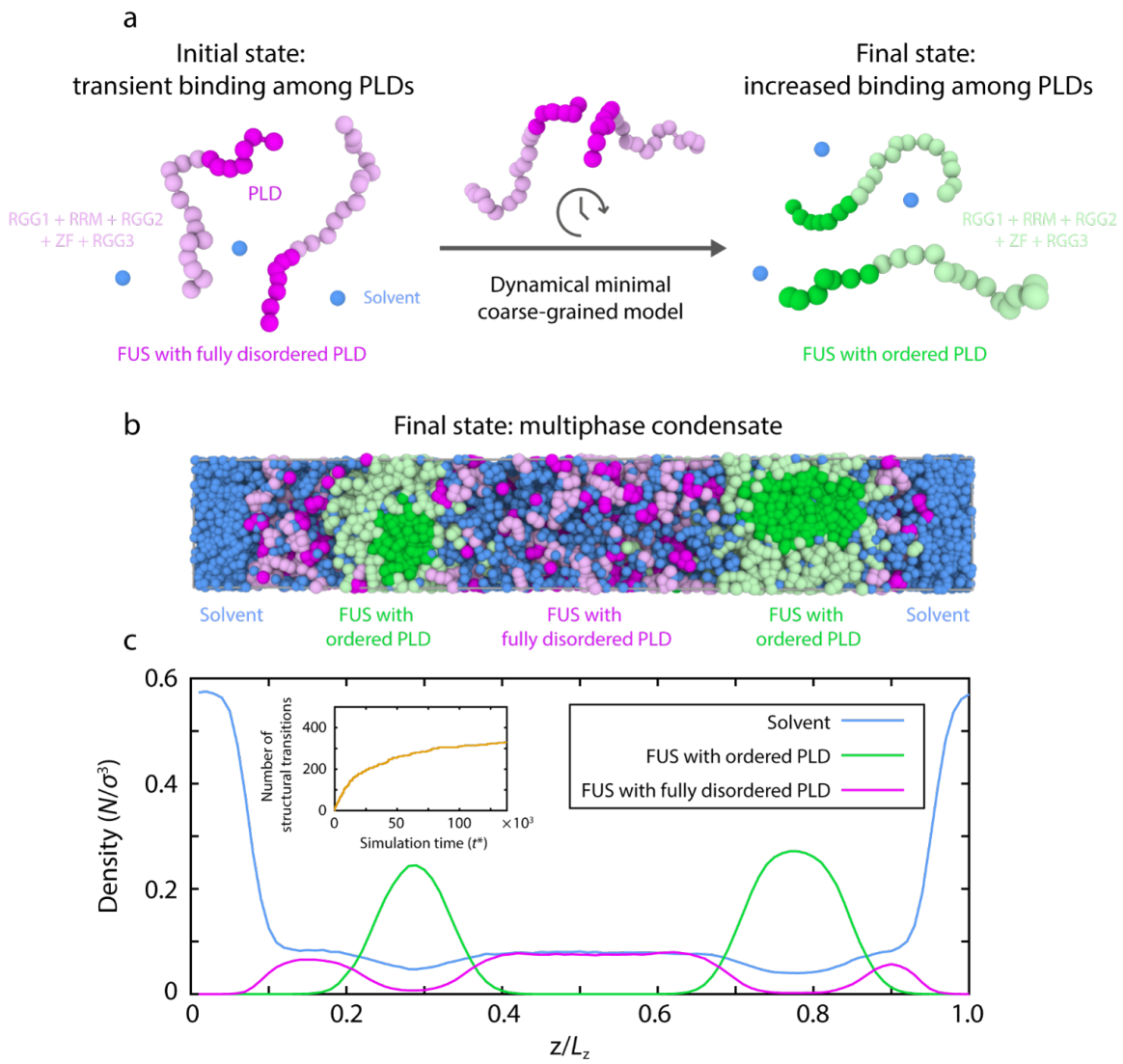

Figure 4. An ageing model predicts multiphase FUS condensates. (a) Schematic illustration of the dynamical algorithm for triggering disorder-to-cross-β-sheet transitions in the minimal coarse-grained model of FUS. This algorithm fosters disorder-to-order transitions by increasing the PLD–PLD interaction strengths by a factor of 8 (according to our atomistic PMF simulations) when FUS-PLDs are in close contact with other four other PLD-FUS domains and still possess enough surrounding free volume to undergo the transition into kinked cross-β-sheets. PLD: Magenta or green; RGG1, RRM, RGG2, ZF, RGG3: light magenta or light green. Solvent (water) is depicted as blue beads. (b) Snapshot of direct coexistence simulations using the dynamical algorithm at a quasi-equilibrium state after structural transitions have saturated. (c) Density profile (in reduced units) of FUS species and explicit solvent across the long side of the simulation box estimated over the coarse-grained quasi-stationary state (as obtained in panel b). Inset: Number of structural transitions in FUS-PLD domains as a function of simulation time (t^*). The plateau region was used to calculate the snapshots and density profiles in panel b and c, respectively. FUS proteins with fully disordered PLDs: magenta; FUS proteins with ordered PLDs (i.e., with kinked cross-β-sheet elements): green; solvent (water): blue.

protein densities are found within the condensate (Figure 4a). Specifically, transitions of LARKS into kinked cross-β-sheets are enforced and recapitulated by modulating the binding interaction strength of PLD–PLD interactions by a factor of 8 (according to our atomistic PMF simulations) when a PLD is in close contact with four other PLDs and still possesses enough free volume around it to undergo a disorder-to-order structural transition (i.e., solvent-mediated)[?,?].

Using our dynamic algorithm, we observed an initial steep increase in the emergence of 'ordered' FUS within the PLDs over time and the formation of multiphase condensate structures with high- and low-density FUS coexisting phases (Figure 4b). Once the multiphase condensate was formed, the rate of emergence of structural transitions in the system decayed rapidly until a quasi-dynamically arrested state was reached (Figure 4c, inset). The density profile of this state is shown in Figure 4c, with the 'ordered' FUS phase located towards the outermost regions of the condensate, and the 'disordered' FUS phase concentrated in the interior. Crucially, the organization of the resulting condensates is equivalent to the architecture obtained after equilibrating condensates with two different populations of FUS determined a priori (c.f., Figure 3).

We next inferred that, besides FUS, other LARKS-containing proteins could also form multiphase gel/liquid architectures as they aged. While FUS forms hollow condensates—with a liquid-core–gel-shell architecture—other proteins may form different multiphase architectures combining gels and liquids. In other words, the specific ordering of the gel phase as the outer layer in FUS emerges from its amino acid sequence and the molecular organization of the various FUS domains inside its gel phase—i.e., as strong LARKS–LARKS bonds form, the charged-rich RGG regions of FUS are pushed towards the surface of the gel, making the gel more hydrophilic than the inner liquid core.

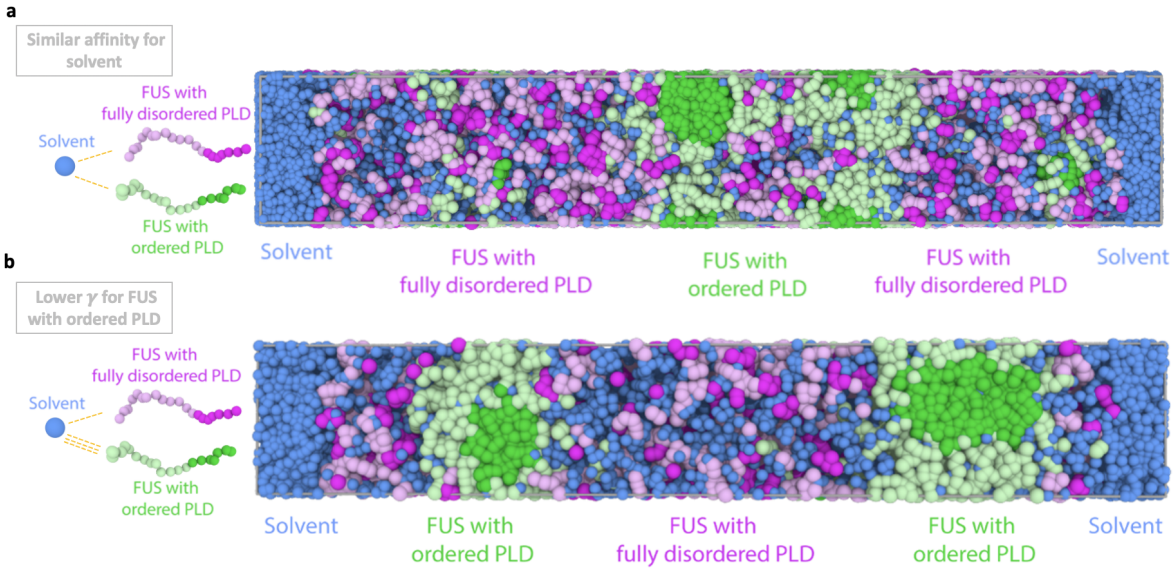

Figure 5. Equal hydrophobicity for structured and disordered PLD FUS proteins causes gel-like phase to relocate to the core of multi-phase condensates. Final state comparison between a control simulation with the dynamical algorithm (a) where we set the same solvent–protein interactions regardless of protein type (for both ordered and disordered PLDs) and (b) the dynamical algorithm simulations from Figure 4 where FUS with ordered PLDs present lower surface tension due to the exposure of their RGG regions to the interface of the droplets. Left: Schematic illustration of solvent–protein interactions. Right: (a) Snapshot of the quasi-equilibrium state of a Direct Coexistence simulation where fully disordered FUS and FUS with ordered PLDs exhibit equal hydrophilicity (i.e. they have the same affinity for water.) (b) Snapshot of the quasi-equilibrium state of a Direct Coexistence simulation where FUS with ordered PLDs have a preferential interaction with water, and thus, a higher hydrophilicity.

To investigate how the layers may organize when the surface of the gel is equally (or less) hydrophilic as that of the liquid-like condensed phase, we devise a set of control simulations where we now define all domains within the protein as equally hydrophilic. Using this model (see parameterization in Table 1 and 2 of the Methods Section), we perform two types of direct coexistence simulations: (1) for a system where we mix a priori 50% of fully 'disordered' proteins with 50% 'ordered' proteins with structured LARKS, and (2) for a system that is formed initially by 100 mol% 'disordered' proteins and where the LARKS can dynamically undergo disorder-to-order transitions according to their local environment (as in Figure 4). Importantly, these new simulations predict that the condensates will now exhibit a multiphase gel-core-liquid-shell arquitecture (Figure 5a); i.e., where the gel-like aged phase is now preferentially located in the core and the liquid forms an outer shell around it. These simulations also strongly suggest that the specific ordering of the gel and liquid phases is determined by

the relative hydrophilicity of the coexisting gel and liquid phase surfaces. Given equal hydrophilicity of both gel and liquid surfaces, the gel, unsurprisingly accumulates at the core due to the stronger protein–protein interactions that sustain it, and the lower surface tension of the liquid-like phase with the surrounding solvent[50].

Conclusions

In this study, we explore the formation of liquid-core-gel-shell and gel-shell-liquid-core multiphase architectures in single-component protein condensates. We show that the disorder-to-order transition of FUS LARKS, which form β-sheet structures, can sufficiently enhance the binding strengths between PLD–PLD domains, and effectively transform a single-component mixture into a multi-component system even in the absence of chemical modifications or external stimuli. Subsequently, such single-component but physicochemically heterogeneous FUS system can form multiphase organizations. This surprising result is underpinned by the structural diversity of each of the coexisting FUS phases. When we examine each of the two coexisting phases, we observe that despite the inner liquid-like core and outer gel-like shell being composed of chemically-identical proteins (only distinguished by the structure of their PLDs), they exhibit strikingly different molecular organisations. That is, the liquid-like low-density phase at the core of FUS multiphase condensates is structurally homogeneous as it is sustained by weak and transient interactions among FUS proteins that can diffuse freely across the whole phase. By contrast, the molecular organization of the gel-like high-density FUS shell is heterogeneous; i.e., PLD regions form a hydrophobic core due to strengthened PLD–PLD interactions that form kinked cross-β-sheets, and RGG2 and RGG3 domains preferentially expose their positively charged side chains to the solvent. These differences in the molecular organization of the liquid *versus* the gel phases of FUS results in significantly different surface tensions, and positions the liquid in the condensate core with a surrounding gel phase. In other words, because the gel phase has a higher charge density at its interface, due to the exposure of positively charged RGG2 and RGG3, it exhibits a lower surface tension than the liquid homogeneous phase. Consistently, in protein systems where the gel phase has equal or lower hydrophilicity than the liquid, we predict the system will form instead a gel-core-liquid-shell architecture, as that organization increases the enthalpic gain for condensate formation and lowers the surface tension of the overall system.

Importantly, our findings suggest that the formation of higher-order protein structures, hence molecular scale processes, can modulate mesoscale phase behavior of protein condensates and lead to the emergence of multiphase architectures in single-component protein systems. This finding is significant as it demonstrates that multiphasic organization can arise not only from multicomponent systems that contain two or more different molecular entities (e.g., two different biomolecules), but also from a single-component system composed of distinct structural ensembles, in this case, FUS-PLD in a fully disordered *versus* an ordered state, can give rise to multiphase behavior. There are several other scenarios, where such structural heterogeneity can arise, for example, through chemical modifications (e.g., post-translational modifications), and it will be interesting to see whether these cases may also lead to single-component multiphase organizations.

The emergence of multiphase organization in a single-component system is interesting from a fundamental point of view as it represents an alternative path to generating molecular heterogeneities needed to drive the formation of hierarchically organized multilayered topologies found in multicomponent systems. We speculate that such transformations in single-component systems may have wide-spread physiological and pathological implications; for example, in the establishment of core–shell structures (e.g., in stress granules)[38], in the formation of multilayered compartments of the FUS-family protein TDP-43 with vacuolated nucleoplasm-filled internal space[42], or other hierarchically organized subcellular condensate structures where molecules may undergo a structural change or experience some other chemical or physical alteration.

More broadly, the multiscale modelling strategies developed in this study can be extended to probe the underlying mechanisms and kinetics of other complex condensate architectures shedding light on the roles of such compartments in cellular function and dysfunction.

Methods

Potential of mean force calculations

We estimated the potential of mean force (PMF) by performing a set of atomistic umbrella sampling MD simulations for three different FUS LARKS tetra-peptide systems, with sequences: 37SYSGYS42, (PDB code: 6BWZ), 54SYSSYGQS61 (PDB code: 6BXV) and , 77STGGYG82 (PDB code: 6BZP)[19]. The simulations were performed with explicit water and ions using the a99SB-*disp* force field for protein/water/ions[67] using the GROMACS 2018 package[68]. The starting configurations for each simulations consisted of four stacked peptides (two pairs of peptides, stacked on top of each other forming a ladder) taken from the cross-β-fibril structures resolved crystallographycally[19]. Acetyl and N-methyl capping groups were added to the termini of each peptide. For each tetra peptide, we performed two sets of simulations. First, we computed the interactions among the tetra peptides when they remain in the cross-β-fibril structures by imposing positional restraints on all the heavy atoms (1000 kJ mol^{-1} nm^{-2}) in all directions (except for the dissociating peptide in the pulling direction); this ensures that the four individual

peptides maintain their secondary structure during the simulation. For the interactions among disordered peptides, we allowed all peptides to freely sample their configurational landscape. In both cases, we kept the positions of the central C_α atoms fixed to control the relative distances between peptides (except in the direction of the reaction coordinate for the dissociating peptide).

As the reaction coordinate, we used the distance between the center-of-mass (COM) of the 'dissociating' peptide and the peptide closest to it in the sampling direction. In the case of the disordered peptides, rather than the COM distance, we took the distance between the central C_αs in the aforementioned chains, to avoid noise due to changes in the COM from fluctuations in the peptide conformations. We spaced windows approximately 0.2 Å from one another along the reaction coordinate, from 4 to 27 Å (where longest range interactions completely vanish). To sample the steep potential among rigid structures, a spring constant of 24000 kJ mol^{-1} nm^{-2} was used.

For the ordered peptides, we ran simulations of about 30 ns per umbrella window. For the disordered peptides, each window was run as an independent simulation 5 times to ensure adequate sampling. Each independent simulation was run for 60 ns (with a different initial velocity distribution)—i.e., an aggregate simulation time of 300 ns. For the integration of the equations of motion, we used the Verlet algorithm with a time step of 1 fs. The error was calculated by estimating the standard deviation of the set of values of the 5 simulations, while in the case of the structured peptides it was computed by a bootstrapping method[69]. To constrain bond lengths and angles, we used LINCS algorithm[70] with an order of 8 and 2 iterations. The cut-off for the Coulomb and van der Waals interactions were chosen at a conservative value of 1.4 nm. For electrostatics, we used Particle Mesh Ewald (PME)[71] of 4th order with a Fourier spacing of 0.12 nm and an Ewald tolerance of $1.5 \cdot 10^{-5}$. The simulations were performed in the NpT ensemble. The temperature and pressure were kept constant using a Nosé–Hoover thermostat[72] at T = 300 K (with 1 ps relaxation time) and a Parrinello–Rahman[?] isotropic barostat at p = 1 bar (with a 1 ps relaxation time), respectively. A NaCl concentration of 0.1 M was used throughout.

Each simulation system comprised a box of approximately 12 x 5.6 x 5.6 nm. We used ∼12700 water molecules and 24 NaCl ion pairs together with 4 peptides. All our systems were electroneutral (i.e., the total charge was zero). After solvation of the peptides, we performed an energy minimization with a force tolerance of 1000 kJ mol^{-1} nm^{-1}, followed by a short, 1000 ps NpT equilibration both with positional restraints (10000 kJ mol^{-1} nm^{-2}) for the heavy atoms of the chains in all three directions of space. The analysis of the simulations was carried out using the WHAM[73] analysis implemented in GROMACS[68]. The first 10% of the simulations was discarded as equilibration time; although we note that including it had negligible effects on the PMF results.

Residue-resolution coarse-grained simulations

For our residue-resolution simulations of FUS proteins, we employ the residue-resolution sequence-dependent coarse-grained model Mpipi[61], which describes almost quantitatively the temperature-dependent LLPS phase behaviour of different protein solutions. In addition, this model correctly predicts the multiphase behaviour of the PolyR/PolyK/PolyU system, and recapitulates experimental LLPS trends for sequence mutations on FUS, DDX4 NTD and LAF-1 RRG domain variants[61]. Within this force field, electrostatic interactions are modelled with a Coulomb term with Debye–Huckel electrostatic screening[74], while hydrophobic, cation-π and π-π interactions are modelled via the Wang–Frenkel potential[75]. Further details on the force field parameters and potentials are provided in Ref.[61] and the Supplementary Information.

To set up these residue-resolution simulations, an initial atomistic model of the full-length FUS protein (Uniprot code: K7DPS7, 526 residues) was first developed in VMD[76] fusing the IDRs with the atomistic structures of the globular domains of FUS (i.e., PDBs 2LCW and 6G99). The complete atomistic model was then coarse-grained at the residue-level by placing one bead in each of the C_α atom positions of FUS. A solution of 48 copies of this coarse-grained FUS proteins was prepared to perform Direct Coexistence simulations[63,77] at constant volume and temperature. We started by running a simulation of the condensate in the NpT ensemble. After equilibration, we extended one of the sides of the box and then equilibrated in the NVT ensemble. The simulation temperatures were chosen to be slightly below the critical temperature (T_c) $T \sim 0.9T_c$ of the model (T_c = 355 K). The production runs were performed for 2 μs, using a Nosé–Hoover thermostat[72] with a relaxation time of 5 ps and a time step of 10 fs. The LAMMPS software molecular dynamics package was used to carry out all our coarse-grained simulations[78].

Minimal coarse-grained simulations

We developed minimal models for FUS proteins integrating our findings from both the atomistic and residue-resolution simulations. We modelled FUS proteins as 20-bead Lennard–Jones (LJ) polymers with one bead representing ca. 26 amino acids; i.e., 6 beads for FUS-PLD, and 14 beads for the RGG1, RRM, RGG2, ZF, and RGG3 regions. A ratio of 6/20 PLD versus total FUS beads recapitulates the ratio of PLD versus total FUS amino acids (i.e., 163/526). The solvent is explicitly modelled using single-bead LJ particles that mimic water-water and water-ion interactions.

Beads that are not directly bonded to each other, establish non-bonded interactions described via the LJ potential:

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

	Solvent	PLD ordered	PLD disordered	Non-PLD ordered	Non-PLD disordered
Solvent	3.50	2.2	1.07	1.15	1.15
PLD (ordered FUS)	2.20 (1.15)	3.8	1.20	1.50	1.20
PLD (disordered FUS)	1.07 (1.15)	1.2	1.45	1.25	1.25
Non-PLD (ordered FUS)	1.15	1.50	1.25	1.50	1.25
Non-PLD (disordered FUS)	1.15	1.2	1.25	1.25	1.25

Table 1. Model parameters for the different ε values in reduced units. Note that the values in parentheses are those for the control simulations in Figure 5.

	Solvent	PLD ordered	PLD disordered	Non-PLD ordered	Non-PLD disordered
Solvent	1.00	1.15	1.15	1.15	1.15
PLD (ordered FUS)	1.15	1.30	1.30	1.30	1.30
PLD (disordered FUS)	1.15	1.30	1.30	1.30	1.30
Non-PLD (ordered FUS)	1.15	1.30	1.30	1.30	1.30
Non-PLD (disordered FUS)	1.15	1.30	1.30	1.30	1.30

Table 2. Model parameters for the different σ values in reduced units.

where ε is the depth of the LJ potential, r the distance between two beads, and σ the molecular diameter of each bead. The mass of every bead was chosen to be $m^* = 1$. The parameters (ε, σ) for the various cross-interactions are shown in Tables 1 and 2. From here on, we employ reduced units, which are defined as: $T^* = k_B T/\varepsilon$, $\rho^* = (N/V)\sigma^3$, $p^* = p\sigma^3/\varepsilon$, and time as $\sigma\sqrt{(m/\varepsilon)}$.

The cut-off of the LJ interactions is set to 3 times the value of σ. Bonds between consecutive beads are modelled with a harmonic potential $V_{\text{har}} = K(r - r_0)^2$ of K = 40 ε/σ^2, and a resting position of $r_0 = 1.3\sigma$. Furthermore, we apply an angular potential of the form, $V_{\text{ang}} = K_\theta\ (\theta - \theta_0)^2$, between consecutive bonds, with an angular constant of K_θ = 0.2 ε/rad^2 and a resting angle of θ_0 = 180° for fully disordered FUS replicas, and with a constant of K_θ = 4 ε/rad^2 and θ_0 = 180° for FUS with ordered PLD, to account for the partial rigidification of the proteins after exhibiting the disorder-to-order fibrillar transition.

Typical system sizes in the minimal model simulations contained 12300 solvent particles and 570 proteins. We used a mixture of 50 mol% fully disordered FUS proteins and 50 mol% with ordered PLD replicas. NVT simulations were run at $T^* = 3.5$ and at density of $p^* = 0.25$. Temperature was kept constant with a Nosé–Hoover thermostat[72] and with a relaxation time of $t^* = 0.4$. The Verlet equations of motion are integrated with a timestep of $t^* = 0.0004$. We employ the Direct Coexistence method[63,77] as described in the previous section and the LAMMPS Molecular Dynamics software[78].

Dynamic minimal coarse-grained model

To dynamically mimic[45] the structural diversity of FUS proteins during ageing, we developed a time-dependent minimal coarse-grained algorithm. Based on the above minimal coarse-gained parametrization, we started with a system composed of a homogeneous single phase of FUS proteins with fully disordered PLDs. Then, FUS-PLD domains can spontaneously transition from their fully disordered state to the ordered state depending on the local environment. Every 100 simulation timesteps, the algorithm evaluates whether the conditions around each fully disordered FUS protein are favorable for undergoing an 'effective' disorder-to-order cross-β-fibrillar transition (the exact conditions are described below), and thus modify the protein parameters given in Tables 1 and 2, to those corresponding to ordered PLD-FUS proteins. Those conditions are evaluated on the central bead of the PLD, which is a good proxy for the average position of the LARKS found in the PLD of FUS. The algorithm changes the identity of two FUS chains from the fully disordered to the ordered parameters once the two following conditions are met: (1) Their two central beads are at a smaller distance than 2.75 σ, and (2) both central beads are surrounded, within a distance of 2.75 σ, by at least 4 other central beads and 6 solvent particles (i.e., characteristic crowded environment of a protein-rich liquid phase). Due to the strong interaction strength observed in our PMF atomistic simulations between structured FUS domains, we set the transition to be irreversible. To carry out these simulations, we used the USER-REACTION[79] package of LAMMPS which allowed us to change the topology of the underlying system components on-the-fly. The rest of the simulation details are the same as those described in the previous section.

Estimation of contact frequencies from residue-resolution coarse-grained simulations

The average number of protein domain contacts within phase-separated condensates were calculated using the MDAnalysis Python library[80]. Amino acids of different protein replicas were considered to be in contact when they were found within a cut-off distance of 0.65 nm (although our conclusions barely depend on the choice of cut-off distance, as long as it is reasonable),

and we averaged these contact numbers along the production time of each simulation. The contacts between two given domains were obtained by summing all amino-acid contacts found in those domains. Then, from the average number of contacts between different regions along the simulation, we calculated the fraction of total contacts per domain by normalization with the total number of contacts.

References

1. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
2. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357** (2017).
3. Banani, S., Lee, H., Hyman, A. & Rosen, M. Biomolecular condensates: organizers of cellular biochemistry. *Nat. reviews. Mol. cell biology* **18** (2017).
4. Boeynaems, S. *et al.* Protein phase separation: A new phase in cell biology. *Trends Cell Biol.* **28** (2018).
5. Brangwynne, C. P., Tompa, P. & Pappu, R. Polymer physics of intracellular phase transitions. *Nat. Phys.* **11**, 899–904 (2015).
6. Gomes, E. & Shorter, J. The molecular language of membraneless organelles. *J. Biol. Chem.* **294**, 7115–7127 (2019).
7. Martin, E. *et al.* Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).
8. Li, P. *et al.* Phase transitions in the assembly of multi-valent signaling proteins. *Nature* **483**, 336–40 (2012).
9. Larson, A. G. *et al.* Liquid droplet formation by hp1α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
10. Wang, J. *et al.* A molecular grammar governing the driving forces for phase separation of prion-like rna binding proteins. *Cell* **174** (2018).
11. Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
12. Franzmann, T. M. & Alberti, S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J. Biol. Chem.* **294**, 7128–7136 (2018).
13. Lorenzo Gotor, N. *et al.* Rna-binding and prion domains: The yin and yang of phase separation. *Nucleic acids research* **48** (2020).
14. March, Z., King, O. & Shorter, J. Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain Res.* **1647** (2016).
15. Patel, A. *et al.* A liquid-to-solid phase transition of the als protein fus accelerated by disease mutation. *Cell* **162**, 1066–1077 (2015).
16. Shen, Y. *et al.* Biomolecular condensates undergo a generic shear-mediated liquid-to-solid transition. *Nat. Nanotechnol.* **15**, 1–7 (2020).
17. 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* (2015).
18. Murray, D. *et al.* Structure of fus protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. *Cell* **171** (2017).
19. Hughes, M. *et al.* Atomic structures of low-complexity protein segments reveal kinked sheets that assemble networks. *Science* **359**, 698–701 (2018).
20. Luo, F. *et al.* Atomic structures of fus lc domain segments reveal bases for reversible amyloid fibril formation. *Nat. structural molecular biology* **25** (2018).
21. Guenther, E. *et al.* Atomic-level evidence for packing and positional amyloid polymorphism by segment from tdp-43 rrm2. *Nat. Struct. Mol. Biol.* **25** (2018).
22. Gui, X. *et al.* Structural basis for reversible amyloids of hnrnpa1 elucidates their role in stress granule assembly. *Nat. Commun.* **10** (2019).
23. Kato, M. *et al.* Cell-free formation of rna granules: Low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–67 (2012).

24. George-Hyslop, P. *et al.* The physiological and pathological biophysics of phase separation and gelation of rna binding proteins in amyotrophic lateral sclerosis and fronto-temporal lobar degeneration. *Brain Res.* **1693** (2018).

25. Alberti, S. & Dormann, D. Liquid–liquid phase separation in disease. *Annu. Rev. Genet.* **53** (2019).

26. Alberti, S. & Hyman, A. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol.* **22**, 1–18 (2021).

27. Harmon, T., Holehouse, A., Rosen, M. & Pappu, R. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *eLife* **6**, e30294 (2017).

28. Espinosa, J. *et al.* Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components. *Proc. Natl. Acad. Sci.* **117**, 201917569 (2020).

29. Jadrich, R. & Schweizer, K. Percolation, phase separation, and gelation in fluids and mixtures of spheres and rods. *The J. chemical physics* **135**, 234902 (2011).

30. Qamar, S. *et al.* Fus phase separation is modulated by a molecular chaperone and methylation of arginine cation-interactions. *Cell* **173**, 720–734.e15 (2018).

31. Bock, A. *et al.* N-terminal acetylation modestly enhances phase separation and reduces aggregation of the low-complexity domain of rna -binding protein fus. *Protein Sci.* **30** (2021).

32. Monahan, Z. *et al.* Phosphorylation of the fus low-complexity domain disrupts phase separation, aggregation, and toxicity. *The EMBO J.* **36**, e201696394 (2017).

33. Jawerth, L. *et al.* Protein condensates as aging maxwell fluids. *Science* **370**, 1317–1323 (2020).

34. Fare, C. M. & Shorter, J. (dis) solving the problem of aberrant protein states. *Dis. Model. & Mech.* **14**, dmm048983 (2021).

35. Feric, M. *et al.* Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165** (2016).

36. 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**, 4334–4339 (2011).

37. West, J. A. *et al.* Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J. cell biology* **214**, 817–830 (2016).

38. Jain, S. *et al.* Atpase-modulated stress granules contain a diverse proteome and substructure. *Cell* **164** (2016).

39. Guillén-Boixet, J. *et al.* Rna-induced conformational switching and clustering of g3bp drive stress granule assembly by condensation. *Cell* **181**, 346–361 (2020).

40. Yu, H. *et al.* Hsp70 chaperones rna-free tdp-43 into anisotropic intranuclear liquid spherical shells. *Science* **371** (2021).

41. Kistler, K. *et al.* Phase transitioned nuclear oskar promotes cell division of drosophila primordial germ cells. *eLife* **7** (2018).

42. Schmidt, B. & Rohatgi, R. In vivo formation of vacuolated multi-phase compartments lacking membranes. *Cell Reports* **16** (2016).

43. Alshareedah, I., Moosa, M., Raju, M., Potoyan, D. & Banerjee, P. Phase transition of rnaprotein complexes into ordered hollow condensates. *Proc. Natl. Acad. Sci.* **117**, 201922365 (2020).

44. Gallego, L. D. *et al.* Phase separation directs ubiquitination of gene-body nucleosomes. *Nature* **579**, 592–597 (2020).

45. Mountain, G. A. & Keating, C. D. Formation of multiphase complex coacervates and partitioning of biomolecules within them. *Biomacromolecules* **21**, 630–640 (2020).

46. Lu, T. & Spruijt, E. Multiphase complex coacervate droplets. *J. Am. Chem. Soc.* **142**, 2905–2914 (2020).

47. Boeynaems, S. *et al.* Spontaneous driving forces give rise to protein- rna condensates with coexisting phases and complex material properties. *Proc. Natl. Acad. Sci.* **116**, 7889–7898 (2019).

48. Kaur, T. *et al.* Sequence-encoded and composition-dependent protein-rna interactions control multiphasic condensate morphologies. *Nat. Commun.* **12** (2021).

49. Fisher, R. & Elbaum-Garfinkle, S. Tunable multiphase dynamics of arginine and lysine liquid condensates. *Nat. communications* **11**, 4628 (2020).

50. Sanchez-Burgos, I., Joseph, J. A., Collepardo-Guevara, R. & Espinosa, J. R. Size conservation emerges spontaneously in biomolecular condensates formed by scaffolds and surfactant clients. *Sci. Reports* **11**, 15241 (2021).

51. Dar, F. & Pappu, R. V. Multidimensional phase diagrams for multicomponent systems comprising multivalent proteins. *Biophys. J.* **118**, 213a (2020).

52. Jacobs, W. & Frenkel, D. Phase transitions in biological systems with many components. *Biophys. J.* **112**, 683–691 (2017).

53. Garaizar, A., Espinosa, J., Joseph, J. & Collepardo, R. Fast protein disorder-to-order transitions are a hallmark of aspherical aged condensates. *To be published* (2021).

54. Dignon, G., Zheng, W. & Mittal, J. Simulation methods for liquid–liquid phase separation of disordered proteins. *Curr. Opin. Chem. Eng.* **23**, 92–98 (2019).

55. Paloni, M., Bailly, R., Ciandrini, L. & Barducci, A. Unraveling molecular interactions in liquid-liquid phase-separation of disordered proteins by atomistic simulations. *The journal physical chemistry. B* **124** (2020).

56. Benayad, Z., von Bulow, S., Stelzl, L. S. & Hummer, G. Simulation of fus protein condensates with an adapted coarse-grained model. *J. Chem. Theory Comput.* **17**, 525–537 (2020).

57. Zheng, W. *et al.* Molecular details of protein condensates probed by microsecond long atomistic simulations. *The J. Phys. Chem. B* **124** (2020).

58. Krainer, G. *et al.* Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun.* **12** (2021).

59. Dignon, G., Zheng, W., Kim, Y., Best, R. & Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLOS Comput. Biol.* **14**, e1005941 (2018).

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

61. Joseph, A. J. *et al.* Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nat. Comput. Sci.* in press (2021).

62. Šarić, A., Chebaro, Y. C., Knowles, T. P. J. & Frenkel, D. Crucial role of nonspecific interactions in amyloid nucleation. *Proc. Natl. Acad. Sci.* **111**, 17869–17874 (2014).

63. Ladd, A. & Woodcock, L. Triple-point coexistence properties of the lennard-jones system. *Chem. Phys. Lett.* **51**, 155–159 (1977).

64. Fernández, R., Abascal, J. & Vega, C. The melting point of ice i for common water models calculated from direct coexistence of the solid-liquid interface. *The J. chemical physics* **124**, 144506 (2006).

65. Espinosa, J., Sanz, E., Valeriani, C. & Vega, C. On fluid-solid direct coexistence simulations: The pseudo-hard sphere model. *The J. chemical physics* **139**, 144502 (2013).

66. Welsh, T. J. *et al.* Surface electrostatics govern the emulsion stability of biomolecular condensates. *bioRxiv* (2020). https://www.biorxiv.org/content/early/2020/10/20/2020.04.20.047910.full.pdf.

67. Robustelli, P., Piana, S. & Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci.* **115**, E4758–E4766 (2018).

68. Berendsen, H., van der Spoel, D. & Drunen, R. Berendsen, h. j. c. , van der spoel, d. van drunen, r. gromacs: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

69. Efron, B. & Tibshirani, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat. Sci.* **1** (1986).

70. Hess, B. P-lincs: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).

71. Darden, T., York, D. & Pedersen, L. Particle mesh ewald - an n.log(n) method for ewald sums in large systems. *J. Comput. Chem.* **18**, 1463–1472 (1993).

72. Nosé, S. A unified formulation of the constant temperature molecular dynamics method. *The J. Chem. Phys.* **81**, 511 (1984).

73. Hub, J., de Groot, B. & van der Spoel, D. g wham-a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* **6**, 3713–3720 (2010).

74. Debye, P. The theory of electrolytes i. the lowering of the freezing point and related occurrences. *Physikalische Zeitschrift* **24**, 185–206 (1923).

75. Wang, X., Ramírez-Hinestrosa, S., Dobnikar, J. & Frenkel, D. The lennard-jones potential: when (not) to use it. *Phys. Chem. Chem. Phys.* **22**, 10624–10633 (2020).

76. Humphrey, W., Dalke, A. & Schulten, K. Vmd: Visual molecular dynamics. *J. molecular graphics* **14**, 33–8, 27 (1996).

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

78. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.* **117**, 1 (1995).

79. Gissinger, J., Jensen, B. & Wise, K. Modeling chemical reactions in classical molecular dynamics simulations. *Polymer* **128** (2017).

80. McGibbon, R. T. *et al.* Mdtraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. journal* **109**, 1528–1532 (2015).

Acknowledgements

The research leading to these results has received funding from the European Reseach Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013) through the ERC grant PhysProt (agreement no. 337969) (T.P.J.K.), under the European Union's Horizon 2020 Framework Programme through the Future and Emerging Technologies (FET) grant NanoPhlow (agreement no. 766972) (T.P.J.K., G.K.), the Marie Skłodowska-Curie grant MicroSPARK (agreement no. 841466) (G.K.), and the ERC grant InsideChromatin (agreement no. 803326) (R.C.-G.). A.G. acknowledges funding from the EPRSC (EP/N509620/)and the Winton programme. We further thank the Newman Foundation (T.P.J.K.), the Biotechnology and Biological Sciences Research Council (T.P.J.K.), the Herchel Smith Funds (G.K.), the Wolfson College Junior Research Fellowship (G.K.), the Winton Advanced Research Fellowship (R.C.-G.), the Oppenheimer Research Fellowship (J.R.E.), Emmanuel College Roger Ekins Fellowship (J.R.E), and King's College Research Fellowship (J.A.J). We also acknowledge funding from the Wellcome Trust Collaborative Award 203249/Z/16/Z (T.P.J.K.). The simulations were 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.

Author contributions statement

Conceptualization J.R.E., A.G., G.K. Y.S., R.C.-G., T.P.J.K.; Formal Analysis, J.R.E., A.G.; Funding acquisition, J.R.E., R.C.-G., T.P.J.K.; Investigation, J.R.E., A.G.,J.A.J ; Methodology, J.R.E., A.G., R.C.-G., J.A.J; Project administration, R.C.-G., T.P.J.K.; Resources, J.R.E., J.A.J, R.C.-G. ; Software, J.R.E., A.G. ; Supervision, J.R.E., R.C.-G., T.P.J.K. ; Visualization J.R.E., A.G.; Writing – original draft, G.K., J.R.E., J.A.J, R.C.-G.; Writing – review & editing, G.K., J.R.E., Y.S., A.G., J.A.J., R.C.-G., T.P.J.K

Competing Interests

The authors declare no competing interests.

Data Availability

Data supporting the findings of this paper are available from the corresponding authors upon reasonable request. A reporting summary for this Article is available as a Supplementary Information file. The sequence of full-length FUS (UniProt accession code: K7DPS7), as well as the resolved structural domains (residues from 285–371 (PDB code: 2LCW) and from 422–453 (PDB code: 6G99)) used in atomistic models for coarse-grained simulations, are available at UniProt (https:// www.uniprot.org/) and the Protein Data Bank (https://www.rcsb.org/), respectively.