ARTICLE

Universal predictive scaling laws of phase separation of prion-like low complexity domains

M. Julia Maristany,† Anne Aguirre Gonzalez,‡ Rosana Collepardo-Guevara,‡ and Jerelle A. Joseph*§

†Department of Physics, University of Cambridge, Cambridge, UK
‡Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, UK
¶Department of Genetics, University of Cambridge, Cambridge, UK
§Department of Chemical and Biological Engineering, Princeton University, US
*Corresponding author. Email: jerellejoseph@princeton.edu

Abstract

Proteins containing prion-like low complexity domains (PLDs) are common drivers of the formation of biomolecular condensates and are prone to misregulation due to amino acid mutations. Here, we exploit the accuracy of our residue-resolution coarse-grained model, Mpipi, to quantify the impact of amino acid mutations on the stability of an unprecedented set of 140 PLD mutants from six proteins (hnRNPA1, TDP43, FUS, EWSR1, RBM14, and TIA1). Our simulations reveal the existence of predictive rules that quantify the range of change in the critical solution temperature of PLDs as a function of the number and type of amino acid sequence mutations. Remarkably, these rules are consistent with the physicochemical properties of the mutations and extend across the entire family tested, suggesting universal scaling laws govern PLD phase behaviour. Our work offers a quantitative lens into how the emergent behaviour of PLD solutions varies in response to physicochemical changes of single PLD molecules.

Keywords: phase separation; biomolecular condensates; prion domains; TDP43; FUS; hnRNPA1; scaling laws; TIA1; EWSR1; RBM14; low complexity domains

1. Main

Biomolecular condensates are highly multicomponent systems with remarkable specificity: they exhibit non-random molecular compositions and, concomitantly, contain distinct physicochemical environments defined by the collective behaviour of their components [1–3]. They are thought to form via phase separation: a thermodynamic process by which macromolecules, primarily proteins and nucleic acids, demix into condensed phases suspended in the cellular cytoplasm or nucleoplasm. Condensates contribute to the spatiotemporal organisation of the myriad of molecular components inside cells [4, 5], and have been linked to diverse cellular functions including regulation of reaction kinetics [6–15], localization of components [1, 4], macromolecular transport [16, 17], and signalling [18, 19] and stress response [20, 21].

Phase separation is primarily driven by multivalent biomolecules; i.e., those that can bind to multiple (three or more) partners simultaneously [3]. Theory and simulations have helped explained that such multivalency is dominant because it allows molecules to interconnect with one another sufficiently to form the percolated network that sustains the condensed liquid at the molecular level [3, 22–26]. While multivalency is, to a great extent, dictated by the chemical-makeup of key biomolecules known as scaffolds, it is not all about the scaffolds. Indeed, the valency of scaffolds is regulated by the composition of the medium (e.g the presence of other biomolecules, metabolites,

© Cambridge University Press . This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

and counterions [27, 28]), and the physicochemical parameters of the environment (e.g. temperature [21, 29], pH [21, 30, 31], and pressure [32]). This realisation has gradually expanded the so-called "molecular grammar" of phase separation [33–36]. The consensus is now shifting towards the idea that multiple, context-dependent, molecular grammars actually exist [33, 35]. Nonetheless, a body of experimental, theoretical and simulation work agree that cation–π and π–π interactions are the key drivers of biomolecular phase separation, with charge–charge, dipole–dipole, and non-polar interactions being crucial in some systems too [27, 33, 37–43].

Intrinsically disordered proteins/regions (IDP/Rs) are strong contributors to the multivalency and phase separation capacity of many naturally occurring proteins [1, 33, 38, 44–50], and more complex systems, like chromatin [28, 51]. IDRs characterized by amino acid sequences of low complexity and enriched in aromatic residues, such as the prion-like low-complexity domains (PLDs) of RNA–binding proteins [33, 52–55], are not only prone to phase separation but also to aggregation (fibrillation) and amyloid formation [33, 55]. PLDs are prevalent in the human proteome, with over 70 different RNA-binding proteins containing this type of domain, and have been observed to undergo phase separation in vitro [56–59]. PLDs have been shown to play a critical role in a variety of cellular processes, including epigenetic regulation, aggregation, and intra-cellular organization [60–62]. It is hypothesized that PLDs play a role in RNA-binding protein function by driving their phase separation [63]. However, the factors that promote the ability of PLDs to drive phase separation are conjectured to also facilitate their aggregation [64]. Phase-separating RNA-binding proteins with PLDs, like Fused in Sarcoma (FUS), TAR DNA-binding protein 43 (TDP43) and heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1), have all been shown to form aggregates [55, 65–75].

Amino-acid sequence mutations on the PLDs of RNA-binding proteins have been associated with a number of neurodegenerative disorders such as ALS and frontotemporal dementia [76, 77]. Aberrant phase separation has been hypothesized to be involved in the onset of such diseases [61, 65, 77–79]. Unequivocally relating the phase behaviour of PLDs to cellular function and dysfunction is a complex task. Nonetheless, the properties of biomolecular condensates of PLDs have been shown to change in response to amino acid sequence mutations [38, 80]. The quantitative experimental phase diagrams for a set of hnRNPA1-IDR variants by Mittag and colleagues in combination with mean-field simulations by Pappu and colleagues have provided a quantitative explanation of how different types of amino acids mutations shift the phase diagrams in the temperature-vs-concentration plane [40]. Such work demonstrates that, at least for hnRNPA1, there exist quantitative rules that control the thermodynamic parameters of condensates in response to amino-acid sequence mutations.

These rules can be rationalized under the framework of so-called scaling laws—mathematical relationships that describe how the properties of a system scale with respect to a given variable, such as its size or composition [81]. Scaling laws have been useful in understanding the behavior of many physical systems, notably polymer systems [82, 83]. Such laws can provide insights into the underlying principles that govern these systems. Furthermore, scaling laws have been instrumental in explaining polymers' conformational behaviour under various conditions [84], elucidating the relationship between chain length and end-to-end distance [83], as well as in assessing their response to external forces [85]. In the context of protein condensates, scaling laws, if found, would quantify how condensates respond to different intrinsic perturbations—such as amino acid mutations or presence/absence of different post-translational modifications, changes in protein length, increased/decreased flexibility—or external stimuli—such as changes to the pH of the medium or the ionic environment. Moreover, the values of the scaling constants would provide a robust quantitative measure of the strength of response of the physical property to the given perturbation. Thus, exploring the existence of scaling laws for PLDs is important to understand how molecular features, such as the nature, competition, and cooperation of intermolecular forces between PLDs, govern

the phase behavior of these condensates.

Here, we exploit the increasing accuracy and efficiency of a modern sequence-dependent coarse-grained model [39] to quantify the full phase diagrams (Temperature-vs-density plane) of an unprecedented set of 140 mutants spanning PLDs of six different RNA-binding proteins (i.e., hnRNPA1, TDP43, FUS, EWSR1, RBM14, and TIA1). Many of these RNA-binding proteins, such as TDP43, FUS, TIA1, EWSR1, and hnRNPA1, have been linked to the etiology of neurodegenerative disorders through pathology and genetics [76, 86–94]. Others, such as RBM14, have not yet been associated with disease, but are known to play a key role in the formation of stress granules, whose dynamics are altered in neurodegenerative disorders [95].

With such a vast exploration of protein sequence space via a massive set of simulations, we uncover universal rules—i.e. conserved across the family of PLDs—that quantify the modulation of critical parameters of PLDs solutions as a function of the number and type of amino acid sequence mutations. These rules are based on the specific amino acid sequence of the PLD and take into account the physical properties of individual amino acids, such as their hydrophobicity and charge. Additionally, by examining the impact of different amino acid substitutions on the critical solution temperature of these protein systems, we have gained insights into the underlying molecular mechanisms of protein condensation, thus providing a more comprehensive understanding of the factors that govern the phase behaviour of proteins.

By investigating the phase behaviour of the family of PLD proteins, we aim to contribute to the understanding of the biophysical characterisation of these aggregates. Deciphering the biophyiscal properties of PLD condensates is needed to eventually grasp RNP granule formation and misregulation, as well as to design protein disaggregates to counter aberrant phase transitions. Importantly, the large and unprecedented size of our simulation set has enabled us to establish a universal code for PLD phase separation, presented through scaling laws that describe changes in critical parameters according to sequence mutations. This code, or universal laws, not only allows us to test the effects of composition bias on PLD phase separation, but also serves as a predictive tool.

2. Results and Discussion

The phase behaviour of PLD proteins and their modulation with amino acid sequence mutations has been extensively investigated with a variety of experimental techniques and under different solution conditions for systems including the hnNRPA1 PLD [40, 96], the FUS PLD [33, 65, 68, 73, 74, 91, 93, 97–101], and the TDP-43 PLD [67, 87, 88, 92, 102], among many others. To complement such a wide body of work, here we set to perform a quantitive comparison of the critical parameters of a total of 140 mutants of the PLDs of TDP-43, FUS, EWSR1, hnRNPA1, RBM14, and TIA1 (Fig. 1) under consistent solution conditions. Our aim is to assess if there exists a universal code—i.e., a simple set of scaling laws—that can predict how the critical parameters of the family of PLDs will scale in response to the change of a given number and type of amino acid sequence mutations. To achieve that, we perform direct coexistence (DC) Molecular Dynamics (MD) simulations (as detailed in the Methods section) of solutions of the 140 different PLDs represented with our residue-resolution coarse-grained model. From our DC simulations, we compute phase diagrams in the plane of temperature-vs-density and estimate the critical solution temperature for each system.

Our residue-resolution model Mpipi (see Method 4.2.2) is ideally suited to investigate the phase behaviour of PLD proteins because it achieves the right balance of the dominant π–π and cation–π interactions, which drive the phase behaviour of PLDs. Indeed, the Mpipi model intrinsically recapitulates the experimental phase diagrams of hnRNPA1 variants by Bremer and colleagues [40] with quantitative accuracy [39]. It is worth noting, however, that the Mpipi model was not parameterized based on the experimental dataset of Bremer and colleagues, but rather the model parameters were developed independently by combining atomistic simulations of amino acid pairs with the bioinformatics data of Vernon and colleagues for π-driven contacts [42].

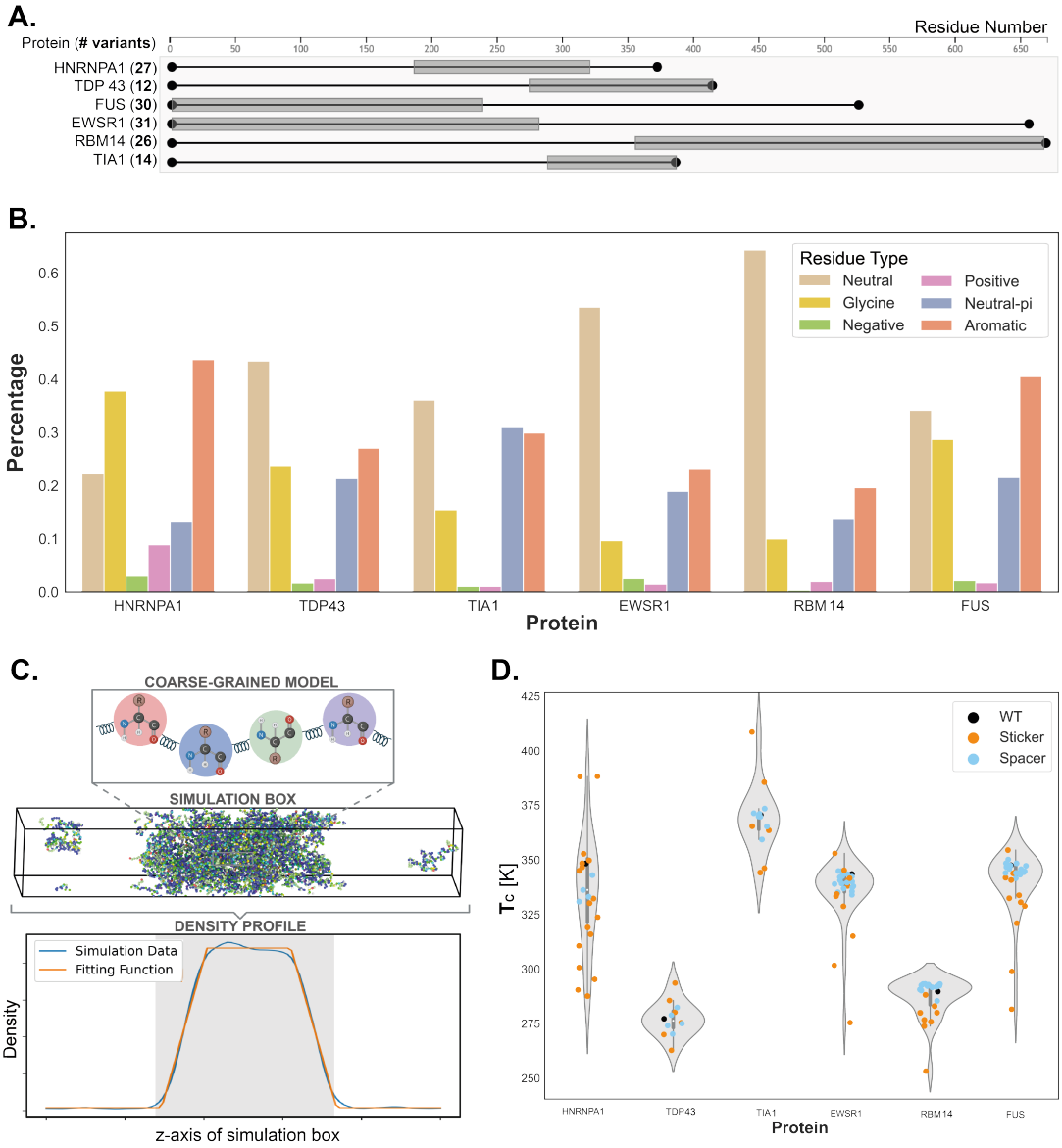

Figure 1. **(A)** Representation of the PLDs of each considered protein, alongside the associated number of variants simulated, totaling a set of 140 sequences. **(B)** Composition of the wild type of the PLDs simulated, in terms of the percentage of neutral amino acids (A, C, I, L, M, P, S, T, and V; no net charge at pH 7 and no π electrons in the side chain), glycine (G), negative (D, E), positive (K, H, R), neutral-π (N and Q; no net charge at pH 7 with π electrons in the side chain) and aromatic (F, W, Y) residues. **(C)** From top to bottom, representation of the simulation model: each amino acid of the protein is represented by a single bead, bonded interactions are modeled with springs, and unbounded interactions are modeled with a combination of the Wang Frenkel potential and Debye-Hückel. The slab simulation method is employed, where both the protein-rich and protein-depleted phases are simulated at the same time. Finally, to obtain the data point in the concentration–temperature phase space, a density profile is computed for each temperature and the concentrations of the rich and depleted phases are obtained via a suitable fitting (Methods). **(D)** Representation of the entire data set. Orange data points represent variants where charged or aromatic residues were mutated (i.e., stickers), while cyan data points represent other types of mutations (i.e., spacers).

The amino acid sequences of all the proteins we investigate, which can be found in the Supplementary Information (SI), were designed based on composition and patterning biases—i.e., attempting to keep the overall amino acid composition percentages and distributing mutations in a homogeneous manner across the entire sequence—in order to probe the effect of each particular type of mutation while avoiding competing and cooperative effects.

2.1 Model validation via the hnRNPA1 PLD

The availability of quantitative experimental binodals for the hnRNPA1-IDR and key amino acid sequence variants [40] has allowed us to test the accuracy of the Mpipi model to describe the phase behaviour of PLD solutions [39]. Here, we make such test more stringent by extending the set of hnRNPA1-IDR variants we use for validation from 8 PLD variants to 21, selecting mutations spanning changes in aromatic, charged, polar and neutral residues. The amino acid sequences of all the hnRNPA1-IDR variants considered are described in the Supplementary Information and are denoted using the intuitive nomenclature established by Bremer and colleagues: i.e., variants that mutate n residues of type X are labeled as $-nX$, and so forth. For more details on the nomenclature, see Methods.

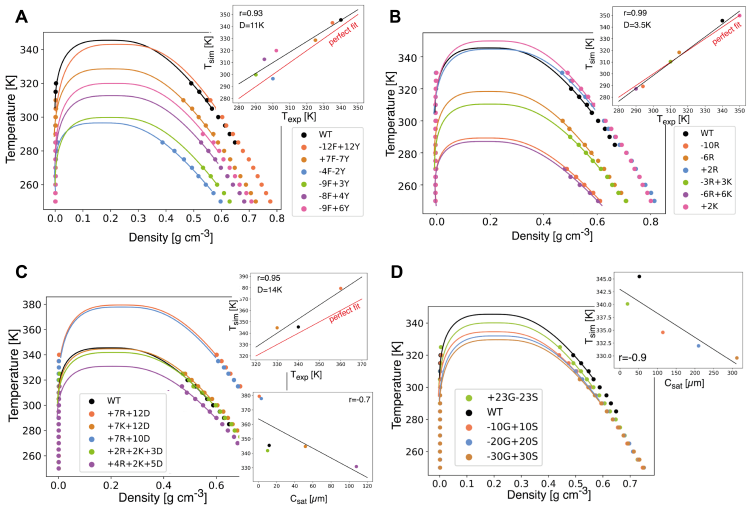

Figure 2. Binodals in the density–temperature phase space of the **A** aromatic, **B** Arg/Lys, **C** charged and **D** Gly/Ser sequence variants of the PLD of the protein hnRNPA1, and comparison with experimental results. The hnRNPA1 PLD was used to validate the accuracy of the model, since experimental critical temperatures are available for a large of these variants [40]. To assess how well the Mpipi model performed for PLDs, we compared critical temperatures obtained in-silico with their experimental counterparts for those variants that were available (depicted in the top right of each binodal, where r is the Pearson coefficient and D is the root mean square deviation between simulated and experimental values), while for the variants where the experimental critical temperature could not be determined, we make a comparison between the critical temperature with the saturation concentration, since for these systems we expect them to be highly correlated.

For each hnRNPA1-IDR variant chosen, we perform direct coexistence MD simulations, compute critical temperatures from the simulations, and compare them with the experimental critical temperatures (Fig. 2). To estimate the critical solution temperatures from the experimental data, we use the fitting methodology outlined in Methods.

Our extended validation confirms that the Mpipi potential can effectively capture the thermodynamic behaviour of a wide range of hnRNPA1-PLD variants, and suggests that it is adequate for proteins with similar sequence compositions.

2.2 Universal Scaling Laws

Building on our previous analysis, we expanded the scope of our study to include the phase behaviour of PLDs, with the goal of investigating whether universal laws govern their behaviour. To accomplish this, we employed a systematic approach to varying the amino acid sequence of the proteins considered, to quantify how the phase behaviour of PLDs changes with compositional perturbations.

2.2.1 Impact of aromatic mutations in phase-separating propensity

Aromatic amino acids, namely tyrosine (Tyr, Y), phenylalanine (Phe, F), and tryptophan (Trp, W), are widely regarded as essential contributors to phase separation of PLDs, due to their ability to form π–π and cation–π contacts [27, 37, 38, 40, 42, 43, 103–105]. Among aromatics, experiments and simulations agree that Tyr is a stronger contributor to the stability of condensates than Phe [27, 33, 40, 100], and that compositional and patterning biases of aromatic residues in PLDs have an impact on the condensate critical parameters and its material state [106, 107].

Aromatic residues are highly abundant in PLDs, with the fraction of aromatic residues (i.e., # of aromatic residues / # of total residues) in the PLD variants analyzed in this study ranged from 0.19 to 0.43 (Fig. 1B). Therefore, as these interactions are so prevalent in PLDs, we begin with analyzing the effect of amino acid presence and identity for the case of aromatic residues.

In order to quantify the impact of swapping Tyr for Phe in shifting the critical solution temperatures of our set of PLDs and searching for the associated scaling law, we designed seventeen protein mutants that varied the Tyr to Phe ratio in the family of PLDs and computed their phase diagrams from direct coexistence simulations (Fig. 3B). Our simulations show that Tyr is stronger than Phe as a driving force for phase separation for the entire PLD family considered, a finding that replicates the results of Bremer et al for hnRNPA1 and confirms that it extends to other PLD proteins. By analyzing the change of critical temperature with respect to the critical temperature of the WT of each PLD, and plotting it with respect to the number of mutations performed, our results reveal a quantitative, linear, relationship between the number of Tyr to Phe mutations and the change in the critical temperature of the system; i.e., our first scaling law, $S_{Y\to F}$, Equation 1. Here, ΔT_c represents the temperature difference between the mutated protein and the wild type (i.e., T_c mutant minus T_c wild type), and $N_{Y\to F}$ represents the number of Tyr residues that were mutated to Phe in the sequence.

$$S_{Y\to F} = \frac{\Delta T_c}{N_{Y\to F}} \tag{1}$$

Note that, by definition $S_{Y\to F} = -S_{F\to Y}$, since $N_{Y\to F}$ is defined as positive for Tyr to Phe mutations and negative for Phe to Tyr mutations, or, in other words, $N_{Y\to F} = -N_{F\to Y}$. For our data set, we find $S_{Y\to F}$ to be -0.40 ± 0.04, where the error given is the standard error. The way to interpret this result is that a single mutation of a Tyr residue to a Phe residue in a given PLD protein would result in a −0.4 K variation in the critical temperature for phase separation.

In addition to the mutations that altered the identity of aromatic residues, we also examined the role of aromatic residues in the phase separation of the PLD as a whole, by designing mutations that replaced aromatic residues (Tyr and Phe) with uncharged, non-aromatic residues—such as Gly, Ser, Gln, Alanine (Ala, A), and Asn—which are prevalent in PLDs. The mutations were carefully designed to preserve amino acid sequence compositional biases by replacing Tyr and Phe with uncharged, non-aromatic residues in the same proportions as they appear in the wild type of each sequence, i.e., maintaining the relative amounts of each type of residue found in the sequence. All amino acid sequences are available in the SI.

The results of our simulations, Fig. 3C, show clearly that aromatic residues are strong modulators of the critical solution temperature of PLD proteins. In particular, deletions of aromatic residues cause a major decrease in the critical solution temperature of PLDs. The scaling law for aromatic deletions suggested by our data is $S_{Y/F\to X}$, where X represents any uncharged, non-aromatic residue.

$$S_{Y/F\to X} = \frac{\Delta T_c}{f_{Y/F\to X}\sqrt{L}} \tag{2}$$

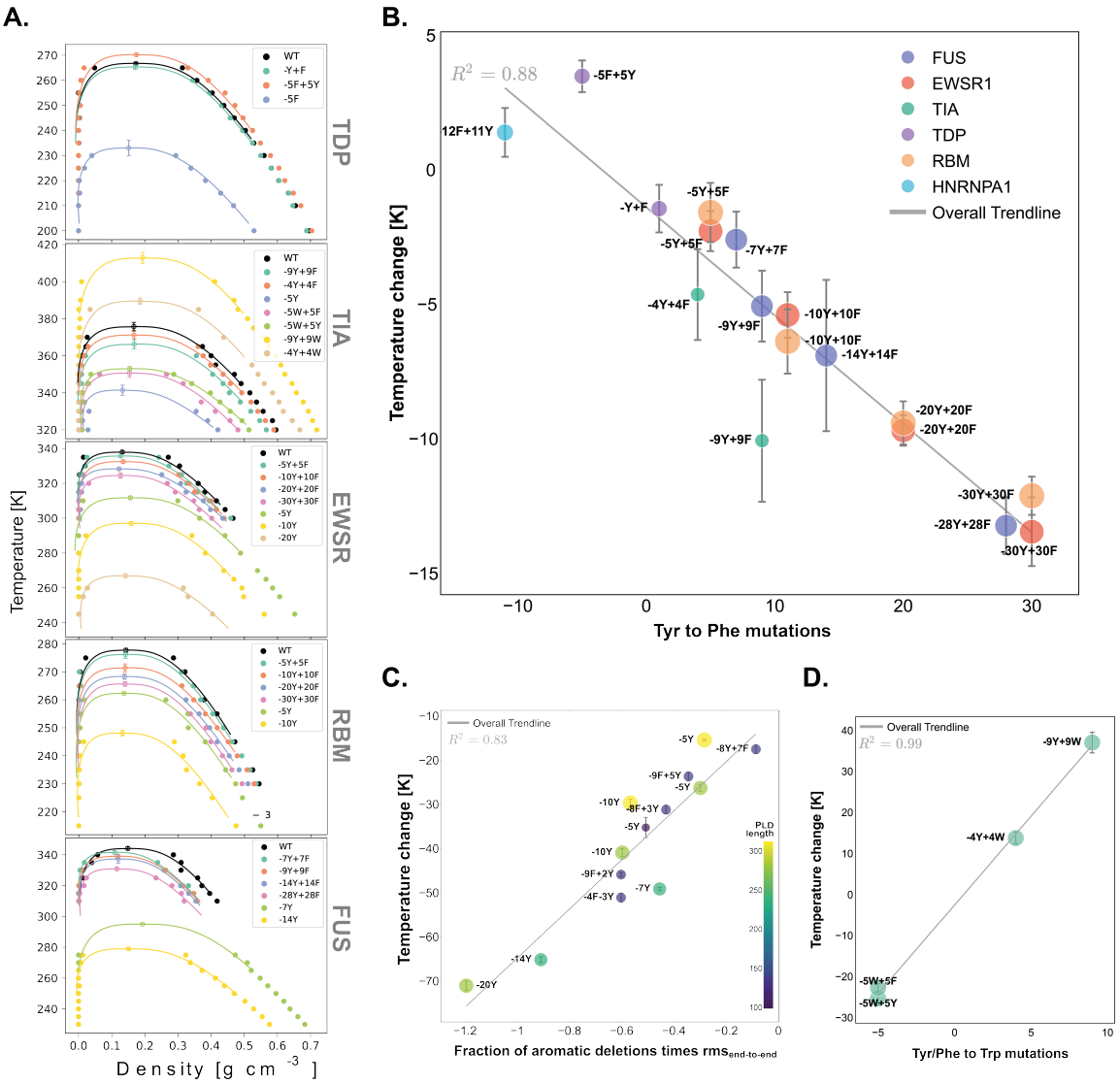

Figure 3. A. Full binodals of each PLD considered and associated variants with aromatic mutations. **B.** Tyr to Phe mutations. The number of Tyr titrated to Phe (x-axis) versus the change in critical temperature (y-axis) from a PLD wild-type condensate to a PLD condensate of the associated variant. The trendline defines the universal scaling law for this mutation. The R^2 value of 0.88 shows the agreement of our data to the universal scaling law defined. **C.** Analysis of the variants involving mutations of aromatic residues to uncharged, non-aromatic amino acids, a.k.a. aromatic deletions. In the y-axis, the variation between the critical temperature in solution of the wild type versus the mutated variant is shown, and in the x-axis, a renormalized measure of mutations: the fraction of aromatic residues mutated times the average root-mean-square end-to-end distance of the PLD, to account for the competing physical interactions between PLDs. **D.** Lys or Phe to Try mutations. In this case, the overall trendline has an R^2 of 0.99.

Here, $f_{Y/F\rightarrow X}$ is the mutation fraction—i.e. the number of mutations divided by protein length—and L is the protein length. The length of the PLDs strongly influences how the critical parameters of PLD solutions change in response to aromatic deletions via two mechanisms. On the one hand, the critical parameters of longer PLDs are more resistant to aromatic deletions simply because for longer PLDs each aromatic deletion represents a smaller percentage of the total strongly interacting amino acids present in the protein. On the other hand, when considered under the framework of heteropolymers, longer PLDs present a higher root-mean-square (rms) end-to-end distance, since for a heteropolymer the distance scales as the square root of the polymer length [108]—longer heteropolymers have higher capture radii, which facilitate the formation of intermolecular contacts and, concomitantly, increase the molecular valencies and critical temperatures.

For our data set, the value of the scaling constant $S_{Y/F\rightarrow X}$ is -56 ± 5. To interpret this result, let us consider as an example a PLD of two hundred amino acids (L = 200). If we mutate one Tyr to an uncharged, non-aromatic amino acid, the change in the critical temperature—computed via the formula $S_{Y/F\rightarrow X} \times \frac{N_{Y\rightarrow X}}{\sqrt{L}}$—is approximately 4 K. Such a value is consistent with the dominance of π-π and cation-π contacts in PLD phase separation reported experimentally [27, 37, 38, 40, 42, 43, 103–105].

Finally, the role of Tyr/Phe to Trp mutations, albeit Trp being rare residue in prion domains, was investigated in TIA1 due to the presence of Trp in the wild type of this PLD, Fig. 3D. The scaling law obtained (Equation 3) exhibits similarities to the law describing Tyr to Phe mutations. However, a notable distinction is that the $S_{F/Y\rightarrow W}$ scaling constant for our data is approximately 4.3 ± 0.1, which is one order of magnitude greater than $S_{Y\rightarrow F}$, consistent with the much stronger π-π contacts established by Trp.

$$S_{F/Y\rightarrow W} = \frac{\Delta T_c}{N_{F/Y\rightarrow W}} \tag{3}$$

2.2.2 Role of arginine in PLD phase separation

The distribution and balance of charged amino acids–i.e. aspartic acid (Asp, D), glutamic acid (Glu, E), histidine (His, H), Arg, and Lys—within an IDR affects its solubility, conformational ensemble, and the range of stability of the ensuing condensates it might form. PLDs are not particularly rich in charged resides; in the wild type PLDs of our data set, the fraction of charged residue ranges from 0.02 to 0.1 (Fig. 1B), and that of only positively charged amino acids is narrowly distributed around 0.015, except for hnRNPA1, which has a value of 0.09 (Figure 1B). Nonetheless, Arg is an important contributor to the stability of a wide-range of biomolecular condensates [37, 40, 109, 110].

While both Arg and Lys are positively charged amino acids, it has been demonstrated that they have unequal roles in the stability of biomolecular condensates [37, 40, 100, 110, 111]. Both residues can form charge–charge and cation-π associative interactions, but the case of Arg is unique. The sp^2-hybridized planar guanidinium group of Arg confers it a special capacity to form strong hybrid π–π/cation–π bonds with the π-electrons of aromatic rings [27]. Thus, to quantify the impact of Arg in the range of stability of PLD condensates, we designed mutants that replaced Arg with Lys and vice-versa (Fig. 4).

The scaling law associated with mutating Arg to Lys is inversely dependent on the fraction of mutations, rather the number of residues mutated. Replacing Arg by Lys does not equate to an equal number of corresponding cation–π and charged contacts. The resulting universal scaling constant for this particular mutation, i.e, $S_{R\rightarrow K}$, has a value of -1300 ± 100, indicating that for a typical PLD of L = 200, the mutation of a single Arg into a Lys results in a reduction of the critical temperature of the system by 6.5 K. Hence, replacing Arg to Lys is a destabilising mutation of PLDs condensates in agreement with the wide-body of experimental evidence [37, 40, 100, 110, 111].

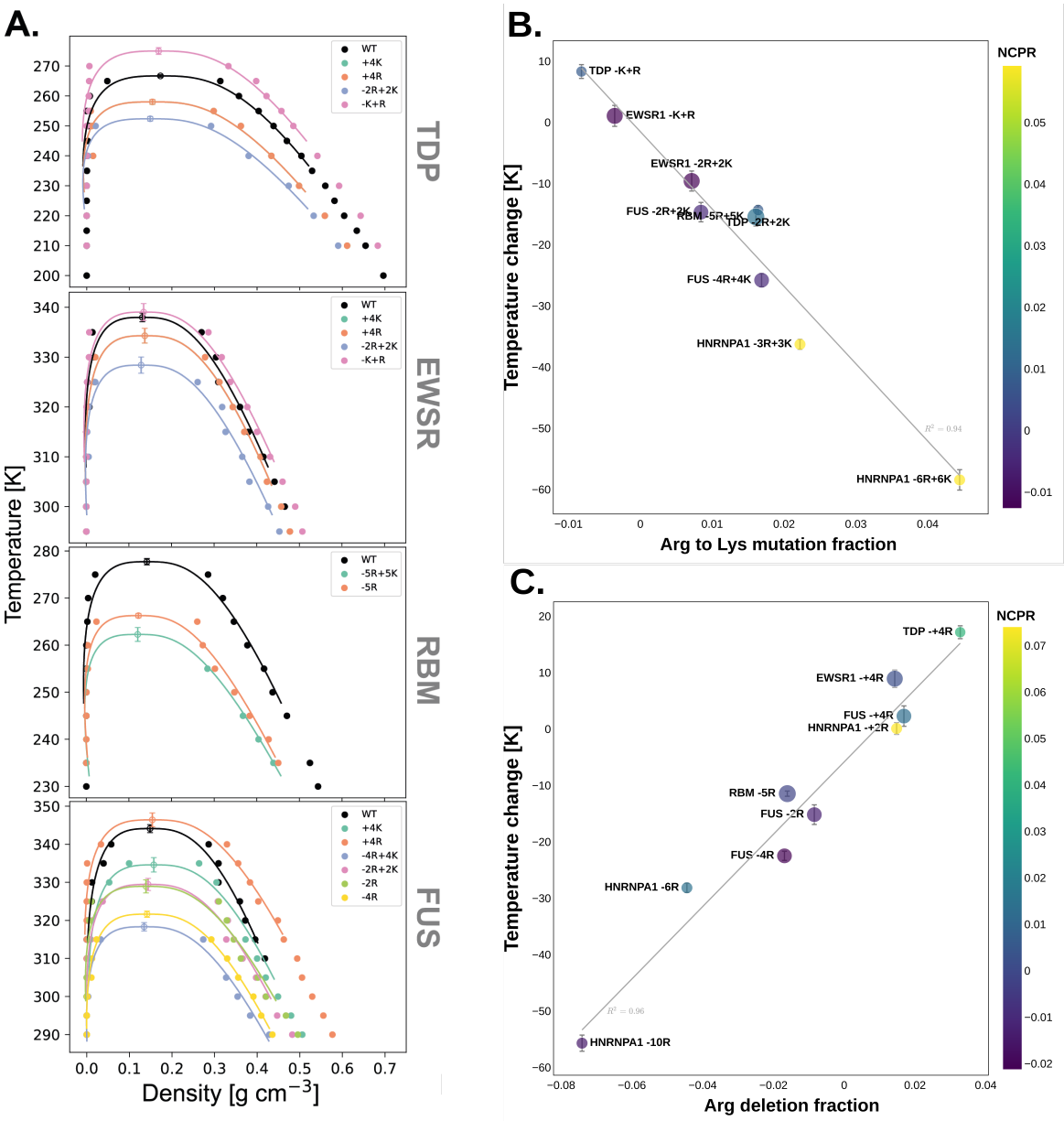

Figure 4. **(A)** Full binodals of each PLD considered and associated variants with Arg mutations. **(B)** Analysis of the critical solution temperature of PLDs with Arg to Lys mutations. The number of Arg to Lys mutations divided by PLD length (x-axis) versus the change in the critical solution temperature from the PLD wild-type condensate to the PLD variant condensate (y-axis). The trendline, which fits the data with an $R^2 = 0.94$, defines the stability measure of this perturbation. **(B)** Analysis of the critical solution temperature of PLDs with Arg mutations to uncharged, non-aromatic amino acids. Arg deletion fraction (x-axis) versus the change in critical solution temperature due to the mutation (y-axis). The trendline, which fits the data with $R^2 = 0.96$, defines the stability measure.

$$S_{R\to K} = \frac{\Delta T_c}{f_{R\to K}} \tag{4}$$

To further analyze the role of Arg, we also tested replacing Arg with uncharged, non-aromatic residues in the same fashion as we did with aromatic sticker deletions. The resulting scaling law depends inversely on the number of amino acids mutated per unit length, with a value of -640 ± 50, amounting to a reduction of 3.2 K in T_c per Arg deletion for a model PLD of length $L = 200$.

$$S_{R\to X} = \frac{\Delta T_c}{f_{R\to X}} \tag{5}$$

It is then almost twice more destabilising to mutate Arg to Lys than to replace Arg with any uncharged, non-aromatic amino acid, namely, Gly, Ser or Ala. Mutating Arg to an uncharged non-aromatic amino acid is less destabilising than mutating it to Lys likely because the charge-reduced variants decrease the cation–cation repulsion of the PLDs, with respect to the Arg-rich or Lys-rich systems. Furthermore, mutating Arg to Lys decreases the enthalpic gain for condensate formation not only via the replacement of stronger Arg-based associative interactions for weaker Lys-based ones but also by increasing the overall solubility of the Lys-rich PLD [43, 112]—the hydration free energy of Arg is less favourable than that of Lys [27, 43, 109].

2.2.3 Contribution of polar amino acids to phase separation.

Uncharged, non-aromatic amino acids, including Gly, Ala, Proline (P, Pro), Ser, and Thr, are commonly found in PLDs. Because they generally form weaker residue–residue interactions than aromatics and present smaller excluded volumes than aromatics, they are classified as spacers that contribute to maintaining the solubility of proteins and material states of condensates [106, 113]. These amino acids can also have other functions in proteins, and the role they play as spacers depends on their chemical makeup but also the specific context in which they are found [40, 80, 114].

Here, we investigate the role of Gln and Asn. Both Gln and Asn are polar amino acids, hence their charge distribution is not symmetric and they present a dipole moment; they can, therefore, establish transient ion–dipole, dipole–dipole and π–dipole interactions with other residues. While the Mpipi model does not account explicitly for polarity, it is worth noting it was parametrized based on bioinformatics data which implicitly incorporates this aspect, and thus, our study allows for the observation of emergent behaviours that stem from this inherent characteristic.

We designed mutations that vary the relative amounts of Gln and Asn, and examined trends in their phase-separating propensity. Mutating the fraction per unit length of Asn (N) to Gln (Q) increases the ability of the system to phase separate, likely due to Gln having a higher dipole moment than Asn, which is approximated in our model via marginally stronger non-ionic interactions formed by Gln over Asn (Figure 5).

$$S_{N\to Q} = \frac{\Delta T_c}{f_{N\to Q}} \tag{6}$$

The value of the universal scaling constant $S_{N\to Q}$ is 50 ± 5, lower than that of mutations of amino acids with a net charge or aromatic side chains, indicating that mutations of Asn to Gln or viceversa have a smaller impact in the critical parameters than changes on the identity of aromatic or charged residues.

2.2.4 Other residues

While Ser, Gly and Thr are also polar amino acids, their dipole moment is significantly smaller than that of Gln and Asp. Consistently, the values of the universal constants involving mutations Ser, Gly and Thr mutations—i.e., $S_{G\to S}$, $S_{G\to T}$ and $S_{S\to T}$—are smaller.

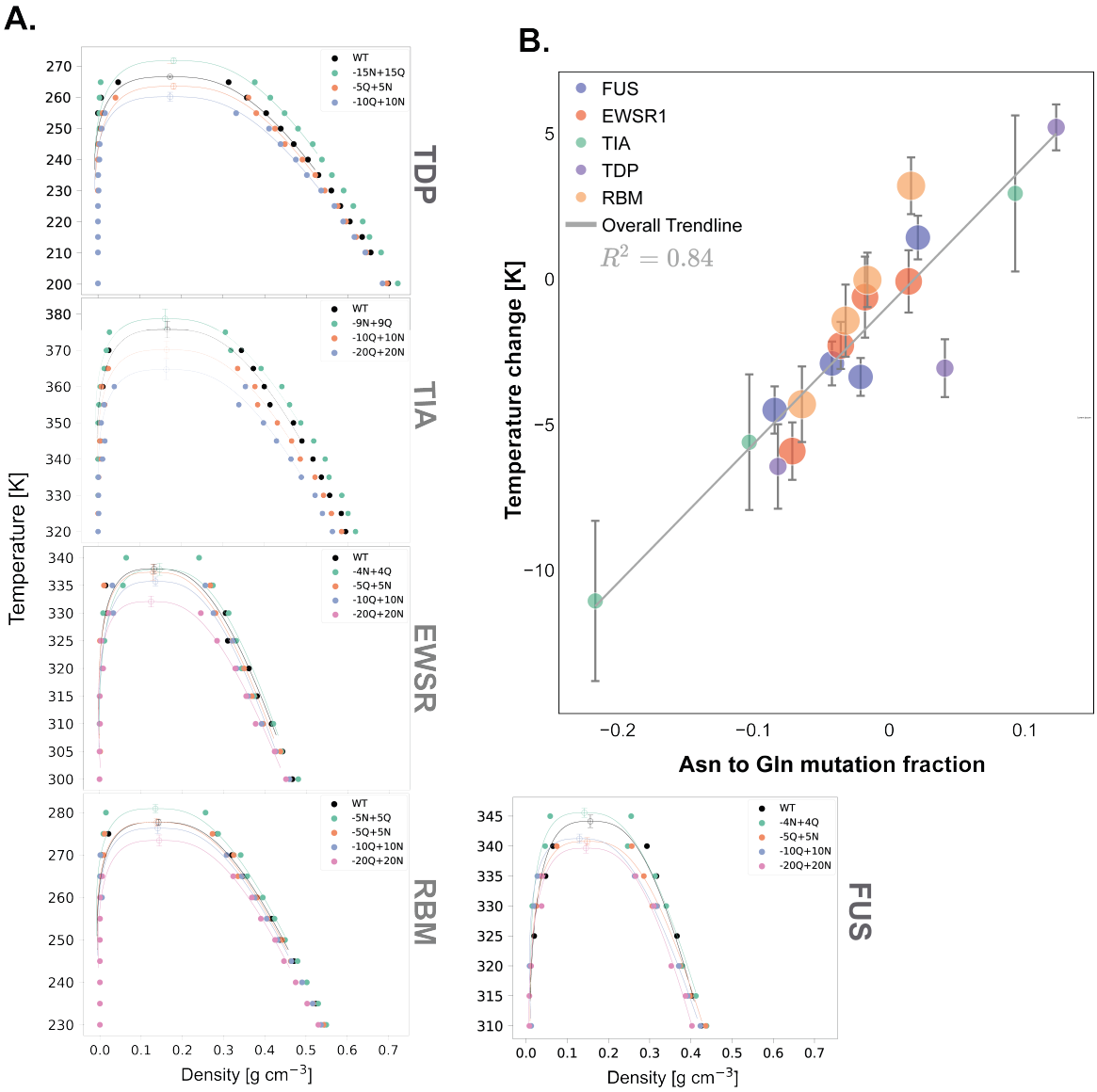

Figure 5. **(A)** Full binodals of each PLD considered and associated variants with Gln and Asn mutations. **(B)** Analysis of the critical solution temperature of PLDs with Asn to Gln mutations. The number of Gln to Asn mutations renormalized by dividing by PLD length (x-axis) versus the change in the critical solution temperature from the PLD wild-type condensate to the PLD variant condensate (y-axis). The trendline, which fits the data with an $R^2 = 0.84$, defines the stability measure of this perturbation.

$$S_{G\to S} = \frac{\Delta T_c}{N_{G\to S}} = -0.010 \pm 0.003 \tag{7}$$

$$S_{S\to T} = \frac{\Delta T_c}{N_{S\to T}} = -0.09 \pm 0.03 \tag{8}$$

$$S_{G\to T} = \frac{\Delta T_c}{N_{G\to T}} = -0.08 \pm 0.02 \tag{9}$$

Regarding Ala, we also tested its relative role with respect to Ser.

$$S_{A\to S} = \frac{\Delta T_c}{f_{A\to S}} = 0.16 \pm 0.05 \tag{10}$$

Amino acids with larger side chains, such as Thr, have a larger excluded volume than those with smaller side chains, such as Gly or Ser. The corresponding value of the scaling constants $S_{S\to T}$ and $S_{G\to T}$ suggests that in the case of Thr mutations, mutations of neutral, non-aromatic amino acids to those with larger excluded volumes inhibit phase separation of the PLD by lowering the corresponding critical temperature, while mutations to those amino acids with smaller excluded volumes have an opposite effect. This is likely due to increased flexibility of the protein backbone, allowing the chain to adopt multiple conformations that promote multivalent contacts.

On the other hand, for mutations involving Ser, Gly, and Ala – all small amino acids with minor side chain size variations – the impact on phase separation seems to stem predominantly from changes in solvation. Solvation plays a key role in phase separation, since it can help balance the driving forces of this process. While hydrophobic regions of the protein, with low values of solvation, tend to form aggregates to reduce interactions with the solvent environment; hydrophilic, more solvating, regions of a protein promote interactions with water. Whether these regions inhibit or promote phase separation is highly context dependent, as on one hand, hydrophilic regions can form energetically favourable contacts with solvent, promoting dispersion, but on the other hand, hydrophilic regions can minimize the entropic cost of sequestering hydrophobic regions within a protein droplet, therefore stabilizing phase separation.

Specifically, in this case, mutations from Gly to Ser increase the size of the side chain while also enhancing solvation, while mutations from Ala to Ser do not significantly change the size of the side chain but significantly increase the system's solvation. In the first case, the critical parameters of phase separation remain virtually unchanged, only presenting a very minor decrease. In the case of Ala to Ser mutations, phase separation is slightly enhanced. Both scaling laws $S_{G\to S}$ and $S_{A\to S}$ reflect such an effect, and underlie the context dependent nature of these mutations.

Thus, mutations of unchanged, non-aromatic amino acids seem to have an effect due to two competing effects: change in side chain size and change in solvation. For these amino acids, the balance between side chain size and solvation appears to govern their phase separation propensity. However, it is important to note that the magnitude of the scaling laws that govern these mutations is extremely small, and thus, likely extremely sensitive to small fluctuations such as changes in the system and solution conditions. Notably, these findings are consistent with recent experiments that report small changes in phase separation propensity for mutations involving neutral, non-aromatic residues.

2.3 Beyond composition: patterning considerations

Beyond composition, the relative positioning of amino acids in the sequence, i.e. residue patterning, also plays a role in protein folding, function and, in this case, phase behaviour [38, 115, 116]. Due to the sequence heterogeneity of PLDs, and, more generally, IDPs, it is particularly complex to define adequate patterning measures [117–120].

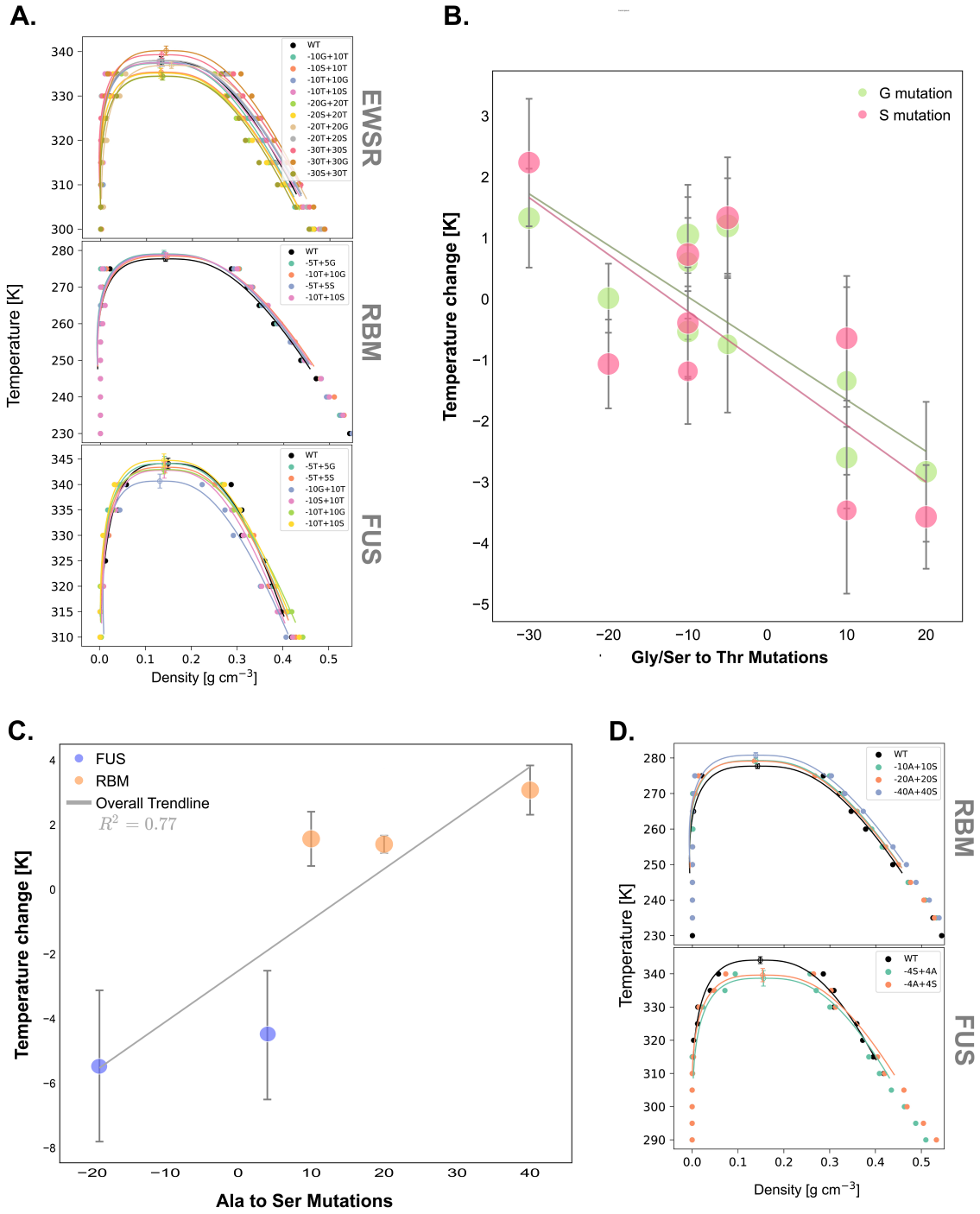

Figure 6. **(A)** Full binodals of each PLD considered and associated variants with Thr, Gly, and Ser mutations. **(B)** Analysis of the critical solution temperature of PLDs with Gly/Ser to Thr mutations (x-axis) versus the change in the critical solution temperature from the wild-type condensate to the variant condensate (y-axis). The trendlines fit the data with an R^2 = 0.7, and define the stability measures of each perturbation. **(C)** Full binodals of each PLD considered and associated variants with Ala and Ser mutations. **(D)** The number of Ala to Ser mutations(x-axis) versus the change in the critical solution temperature from the PLD wild-type condensate to the PLD variant condensate (y-axis). The trendline fits the data with an R^2 = 0.77, and defines the stability measure of this perturbation

In this work, we have taken advantage of our large data set to analyze the impact of amino acid sequence patterning by adopting a patterning order parameter [38, 121, 122], σ_{aro}, that can describe patterning effects of aromatic residues.

To compute σ_{aro}, we divide the amino acid sequence into peptide segments of size l, compute a measure of sequence aromatic asymmetry for each segment, and finally sum the values of such measure over all the segments. Consistent with literature [38, 121, 122], we take l to be 5 and 6. σ_{aro} is then given by:

$$\sigma_{\text{aro}} = \left(\sum_{\text{segments}} \frac{\sigma_i}{\sigma_{max,i}} \right) \frac{N_{stickers}}{L}, \qquad (11)$$

where L is the total number of amino acids in the protein domain, $N_{stickers}$ the number of sticker residues—i.e. charged or aromatic residues—and σ_i has the functional form

$$\sigma_i = \left(\left(2\frac{N_{aro,i}}{L_i} - 1 \right)^2 - \sigma_T \right)^2. \qquad (12)$$

In here, $N_{aro,i}$ is the number of aromatic residues in the i-th segment, L_i the segment length, and σ_T is a function of the total number of aromatic residues in the sequence N_{aro}:

$$\sigma_T = \left(2\frac{N_{aro}}{L} - 1 \right)^2 \qquad (13)$$

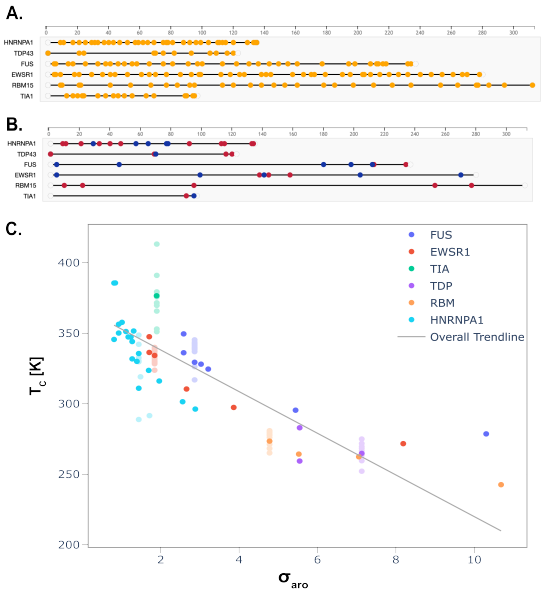

Figure 7. **A** Representation of the aromatic amino acids present in the wild type of each PLD considered **B** Representation of the positively charged (red) and negatively charged (blue) amino acids present in the wild type of each PLD considered **C** Analysis of the critical solution temperature

In turn, $\sigma_{max,i}$ is the measure of aromatic amino acid distribution for an artificial sequence with all the aromatic amino acids concentrated in one extreme.

These results indicate, on the one hand, that the sequences with the most distributed aromatic amino-acids phase separate more readily, due to the increase in multivalency and percolated network structure. On the other hand, the order parameter presented here gives a quantitative tool that we can use to predict the phase-separating propensity of a PLD.

3. Conclusions

Computer simulations have increasingly improved their ability to probe the link between the emergent properties of protein solutions and the chemical makeup of their individual biomolecules, while also considering the physicochemical parameters of the environment. For instance, modern residue-resolution protein coarse-grained models [39, 103, 123–127] can now be used to quantify the modulation of phase diagrams of protein solutions as a result of amino acid sequence mutations of their native proteins with near-quantitative agreement with experiments [39, 125, 126]. In this study, we exploit the accuracy of one of these models—i.e., Mpipi [39]—to compute and analyse the temperature-vs-density phase diagrams of an unprecedentedly large set of PLD proteins—140 variants of the proteins hnRNPA1, TDP43, FUS, EWSR1, RBM14, and TIA1. Collectively, our set

of phase diagrams reveals universal scaling laws that quantitatively predict the change in the critical solution temperatures of single-component PLD solutions as a function of the number and type of amino acid sequence mutations of the PLDs.

Consistent with experimental and computational evidence that interactions involving aromatic residues of the type π–π and cation–π are dominant in stabilising condensates formed by PLDs [33, 39, 40], our findings reveal the largest decrease in the critical solution temperature of a PLD condensate occurs when an aromatic residue is replaced by an uncharged, non-aromatic amino acid. Our scaling law for such aromatic deletions (i.e., $56\,\mathrm{K}/\sqrt{L}$) predicts that the destabilising effect of replacing an aromatic with a non-aromatic residue is, on the one hand, linearly amplified by the number of such mutations and, on the other, screened by the square root of the length of the protein (L). Condensates made of longer PLDs are more robust against aromatic deletions because, by construction, each aromatic residue represents a smaller percentage of the total valency of the PLD in question. In other words, longer PLDs are higher valency molecules than their shorter counterparts and can, thus, form densely connected percolating condensed-liquid network, even when a few of their aromatic stickers have been deleted.

Conversely, our findings also highlight the crucial role Arg has in the critical parameters of PLD condensates. Specifically, our scaling law for Arg deletions (that is, $640\,\mathrm{K}/L$) predicts that the destabilising effect of an Arg mutation to an uncharged, non-aromatic residue is inversely proportional to the length of the protein (L), as well as linear with respect to the number of mutations. While formulaically similar to the scaling law for aromatic deletions, a notable difference is the change in dependence on the PLD length. In the case of Arg deletions, the effect of mutations is screened by L rather than $\sqrt{L}$, indicating a stronger screening dependence on the PLD length. This indicates that for long PLDs, the destabilising effects due to Arg deletions are diminished when compared to aromatic deletions, while in shorter PLDs this scenario is inversed with each Arg deletion exerting a more significant destabilising effect than the aromatic deletions.

These results underscore once again mounting evidence that, for PLDs, cation-π and $\pi - \pi$ interactions are the main drivers of phase separation [27, 37, 38, 40, 42, 43, 103–105]. While it may seem counterintuitive for Arg deletions to have a lesser effect in the critical parameters of longer PLDs than their aromatic counterparts (as Arg residues are known to form strong cation-π interactions [128, 129]) the explanation lies in the type of systems in question, which are particularly rich in aromatic residues. Longer PLDs have, as a result of their higher valency, a larger network of possible pairwise interactions, which can mitigate the destabilizing effect of Arg deletions. In the case of shorter PLDs, each Arg residue represents a larger proportion of the total valency, and thus its deletion has a proportionally larger impact on the critical parameters. In contrast, aromatic residues contribute to the stability of the condensate through both cation-π and π-π interactions, which are predicted to be weaker than their cation-π analogues [27, 39], but more numerous in PLDs. In a longer PLD, the deletion of an aromatic residue results in the loss of more potential interactions, creating a larger destabilizing effect than the deletion of an Arg residue. However, in shorter PLDs, due to their lower total valency, the loss of potential associative interactions due to the deletion of an aromatic residue is lower, and hence its destabilizing effect is comparatively less than for the Arg deletion.

Our scaling laws have also allowed us to quantify the destabilizing effect of Arg to Lys mutations in the phase separation of PLDs—mutations of Arg residues to Lys result in an overall decrease in the critical temperature of phase separation, two-fold in comparison to the impact of mutating Arg into an alternate uncharged, non-aromatic residue. Consequently, the scaling law considerably intensifies for mutations from Arg to Lys, nearly doubling to $1300\,\mathrm{K}\ L$. This implies that each mutation of Arg to Lys exerts a substantially more destabilising effect than mutations of Arg to other uncharged, non-aromatic residues, consistent with experimental evidence [37, 40, 110, 130] and highlighting the significant, modulatory role Lys plays in the phase behaviour of PLDs.

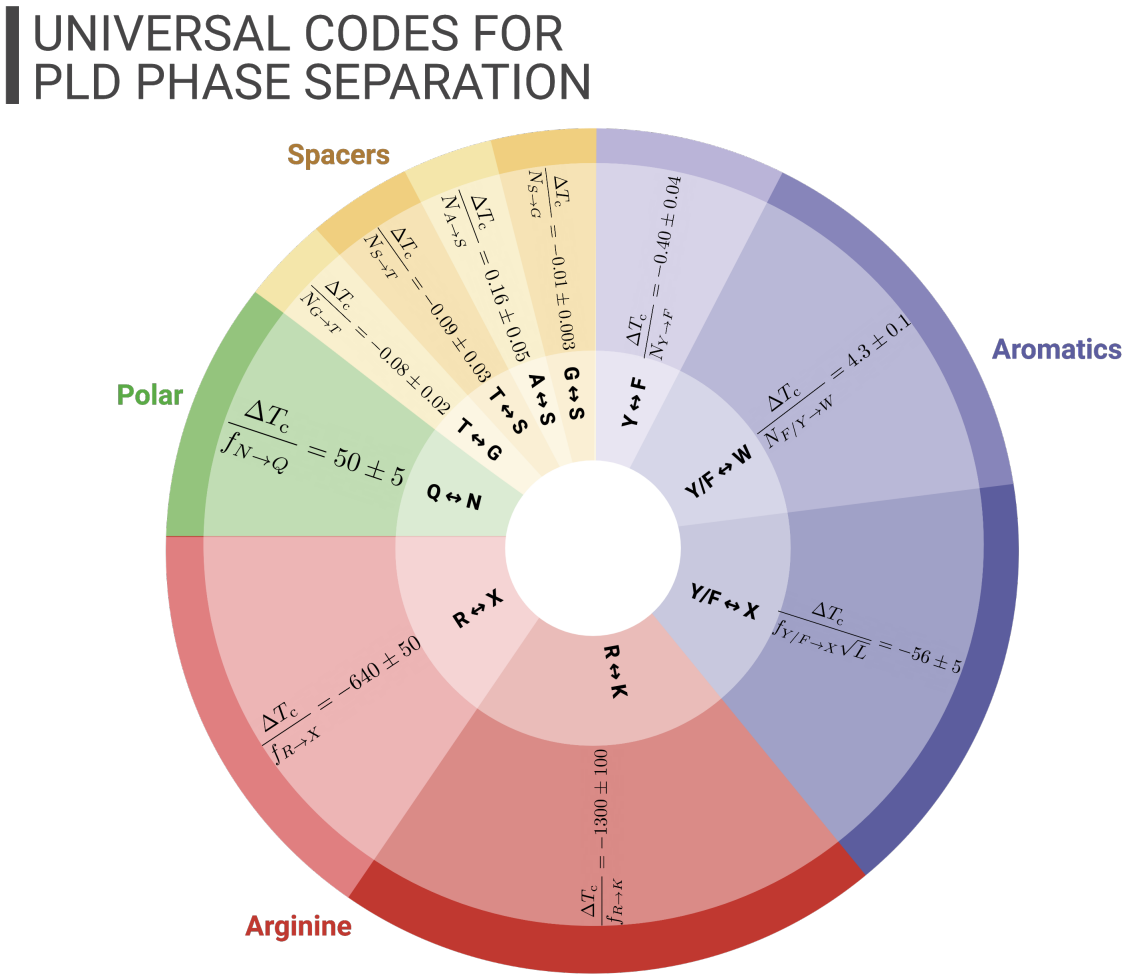

Figure 8. Universal scaling laws for PLD phase separation. Each scaling law quantifies the effects of amino acid sequence mutations in the emergent behavior of PLDs condensation.

Uncharged, non-aromatic residues such as Ser, Gly, Asn, Gln, and Thr are also found to play a role in PLD condensate stability [40, 80, 106, 113]. Notably, PLD condensates are highly stable in the face of mutations of Gly, Ser, Thr and Ala. This behavior is quantified by their respective scaling laws, with values for the scaling constants for mutations from Ser to Gly, Gly to Thr, Ser to Thr and Ala to Ser at –0.01 K, –0.09 K, –0.08 K and 0.16 K respectively. These numerical values indicate that, for a change of one degree in the critical temperature, approximately 10 amino acids of this set would need to undergo mutation. Notably, these scaling laws do not display length dependency. This observation aligns with the idea that these mutations primarily affect the condensate stability through modulation in excluded volume and solvation, rather than by interrupting specific interaction networks.

In contrast, the condensates display heightened sensitivity towards Asn and Gln mutations, underscored by the Asn to Gln scaling of 50 K/L. This law is also inversely length dependent, indicating that mutations on these amino acids affect interaction networks, and that mutations Asn to Gln result in a heightened ability of the system to phase separate. We hypothesize that, in addition to differences in solvation volume, the polarity of each amino acid plays a role by forming charge–dipole and dipole–dipole interactions. Both Gln and Asn are polar amino-acids, with Gln having a longer side chain than Asn, which could potentially interact with a larger number of nearby atoms.

Furthermore, the additional methylene group (-CH2-) may slightly alter the electron distribution around the amide group, which can lead to the formation of more stable dipole–dipole contacts. It is important to note that our computational model, Mpipi, is a charge-fixed model, and hence does not directly capture the effects of polarization. Nonetheless, in this case, the lack of physical details in the Mpipi m odel are compensated by the parameters having being derived from bioinformatics data.

The definition of a universal set of scaling laws acts as a unifying framework for the study of the critical parameters of PLD phase separation, by allowing for direct comparison of individual sequence effects. Such laws are cumulative, meaning that the effects of successive mutations can be predicted by the addition of the appropriate laws, as long as the resulting mutation keeps within PLD composition and patterning biases. While the laws defined in this work are contingent on PLD composition and patterning biases, we hypothesize that they are extendable to other families of IDPs.

We have then provided a universal code for PLD phase behavior, defined through the identification of universal scaling laws that can predict changes in phase behavior based on sequence composition mutations. This code, illustrated in Figure 8, not only allows for testing the effects of composition bias on PLD phase separation but also serves as a predictive tool, allowing for a better selection of mutations to achieve desired changes in different condensates. This study focused on PLDs, but we hypothesize similar scaling behavior in other IDPs, with the laws having a dependence on the protein length that reflects the nature of the driving interaction networks of the system. Our findings have important implications for understanding PLD condensation and for designing modifications to counter their aberrant aggregation, by setting the basis for characterizing and quantifying the molecular determinants of PLD phase separation.

4. Methods

4.1 Design of the dataset

For the design of sequence variants to test the phase behaviour of prion-like domains, we used in-silico targeted mutagenesis to introduce specific changes to the amino acid sequence of the PLD of interest. These changes involved substituting different amino acid types, while carefully maintaining patterning and compositional biases, to test the role of that particular amino acid in the phase-separating propensities of the PLD.

The nomenclature employed is consistent with that of Bremer and colleagues [40]. Namely, if n amino acids of type X are removed from the sequence and replaced with n amino acids of type Y, the variant is labelled $-nX + nY$. If instead, n amino acids of type X are removed from the sequence and replaced with amino acids in proportion to the compositional bias of the PLD, the variant is labelled $-nX$.

As a starting point, we selected a family of PLDs composed of hNRNPA1, RBM14, FUS, TDP43, TIA1 and EWSR1, all PLDs with different composition and patterning biases, in order to assess the hypothesis present in the literature on the impact of sequence features in PLD phase separation holds for the whole family of PLDs and to develop a universal code that can aid us in understand and predicting their phase behaviour.

The full sequences wild types and associated mutations can be found in the SI.

4.2 Binodal determination through MD

Determining the full binodals for PLDs in solution (i.e., the boundaries that separate two coexisting phases in a thermodynamic system) can be achieved through molecular dynamics (MD) by analyzing the phase behaviour of a protein ensemble systematically under different thermodynamic conditions. One approach, and the one we use in this study, is to perform a series of NVT-MD simulations at different temperatures, where NVT stand for canonical-ensemble simulations with a

constant volume, and then analyse the resulting densities in the temperature-density phase space. This method allows for the sampling of the phase space under different thermodynamic conditions to obtain the coexistence line, i.e the binodal.

4.2.1 Direct Coexistence Simulations

The utilization of direct coexistence simulations enables the simultaneous simulation of both high and low-density phases of a mixture, separated by an interface, within a single simulation box. This methodology facilitates the calculation of phase diagrams of protein mixtures within the temperature-concentration space.

To initiate the NVT simulations, we set up a box elongated in one of the Cartesian coordinates, with a specific number of proteins, N, in a compact slab within the box. In order to optimize the computational efficiency and avoid finite-size effects on the system, we chose N=64, a value that provided a balance between computational speed and system size. The box size was also optimized to reduce finite-size effects.

From this starting configuration, the system is maintained at constant temperature and volume, and let to converge to a state in equilibrium. Convergence tests can be found in the SI.

All the simulations were run using the LAMMPS software package [131] and the Mpipi coarse-grained force-field [39]. The Mpipi model was used due to its well-balanced electrostatic, cation–pi and $\pi - -\pi$ interactions and its proven quantitative accuracy with experiments.

4.2.2 Mpipi model

The Mpipi model is a coarse-grained model where each amino acid or nucleic acid is represented by a single bead, with its corresponding physical properties such as mass, molecular diameter (σ), charge (q), and an energy scale that reflects the relative frequency of planar $\pi - \pi$ contacts (ϵ) [39].

$$E_{\text{mpipi}} = E_{\text{bond}} + E_{\text{elec}} + E_{\text{pair}} \tag{14}$$

Bonds are represented via stiff harmonic springs

$$E_{\text{bond}} = \sum_i \frac{1}{2} k (r_i - r_i^0)^2 \tag{15}$$

The electrostatic contribution is modelled via a Coulomb interaction term, with Debye screening

$$E_{\text{elec}} = \sum_{i,j} \frac{q_i q_j}{4\pi\epsilon_r \epsilon_0 r_{ij}} \exp\left(-\kappa r_{ij}\right) \tag{16}$$

Finally, the specific pair interactions are modelled with a Wang-Frenkel potential

$$E_{\text{pair}} = \sum_{i,j} \phi_{i,j} = \sum_{i,j} \epsilon_{i,j} \alpha_{i,j} \left[\left(\frac{\sigma_{i,j}}{r} \right)^{2\mu_{i,j}} - 1 \right] \left[\left(\frac{R_{i,j}}{r} \right)^{2\mu_{i,j}} - 1 \right]^{2\nu_{i,j}} \tag{17}$$

4.2.3 Data fitting

To estimate the critical temperature, we use the law of coexistence densities

$$\left(\rho_{\text{high}}(T) - \rho\text{low}(T) \right)^{3.06} = d \left(1 - \frac{T}{T_c} \right) \tag{18}$$

and critical densities were computed by assuming that the law of rectilinear diameters holds, namely

$$(\rho_{\text{high}}(T) - \rho_{\text{low}}(T))^{3.06} = d(1 - T/T_c) \tag{19}$$

where $\rho_{\text{high}}(T)$, $\rho_{\text{low}}(T)$ and ρ_c are the densities of the high-density and low-density phases and the critical density, respectively, Tc is the critical temperature, and d and A are fitting parameters.

4.3 Finite-size scaling

Finite size effects are important to consider in the context of in-silico simulations since they can have strong effects. To ensure our calculations were robust with respect to system size, we performed several systematic tests.

In the first place, we performed the simulations at a higher value of N—i.e., the number of particles—and confirmed the resulting phase diagrams were equivalent within error bars. We did this test for 91 and 128 proteins for the protein wild types, and since we found no difference in the predicted critical temperatures, we hypothesize that finite-size effects are not negligible.

To study this hypothesis more carefully, we also varied systematically the cross-section and length of the simulation box. More information, including graphs of convergence, can be found in SI 2.

Acknowledgements

We thank Dr. Jan Huertas and Pin Yu Chew for useful comments on the manuscript. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 803326). MJM would like to acknowledge the Winton Programme for Physics of Sustainability for doctoral funding. AAG is funded by the European Research Council (ERC, grant agreement No 803326). This project made use of time on HPC granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim (http://hecbiosim.ac.uk), supported by EPSRC (grant no. EP/R029407/1). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript

References

[1] Salman F. Banani et al. "Biomolecular condensates: Organizers of cellular biochemistry". In: *Nature Reviews Molecular Cell Biology* 18.5 (May 2017), pp. 285–298. ISSN: 14710080.

[2] Sabareesan Ambadi Thody et al. "Small Molecule Properties Define Partitioning into Biomolecular Condensates". In: *bioRxiv* (2022).

[3] Pilong Li et al. "Phase transitions in the assembly of multivalent signalling proteins". In: *Nature* 483.7389 (2012), pp. 336–340.

[4] Clifford P. Brangwynne et al. "Germline P granules are liquid droplets that localize by controlled dissolution/condensation". In: *Science* 324.5935 (June 2009), pp. 1729–1732. ISSN: 00368075.

[5] Clifford P Brangwynne, Timothy J Mitchison, and Anthony A Hyman. "Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes". In: *Proceedings of the National Academy of Sciences* 108.11 (2011), pp. 4334–4339.

[6] Gregory L. Dignon, Robert B. Best, and Jeetain Mittal. "Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties". In: *Annual Review of Physical Chemistry* 71.1 (2020), pp. 53–75.

[7] Bhawna Saini and Tushar Kanti Mukherjee. "Liquid-Liquid Phase Separation Regulates Enzymatic Activity". In: *bioRxiv* (2022).

[8] Guillaume Laflamme and Karim Mekhail. "Biomolecular condensates as arbiters of biochemical reactions inside the nucleus". In: *Nature Communications Biology* 3 (2020).

[9] F Frottin et al. "The nucleolus functions as a phase-separated protein quality control compartment". In: *Science* 365.6451 (2019), pp. 342–347.

[10] Luke M Oltrogge et al. "Multivalent interactions between CsoS2 and Rubisco mediate α-carboxysome formation". In: *Nature structural & molecular biology* 27.3 (2020), pp. 281–287.

[11] Elizabeth S Freeman Rosenzweig et al. "The eukaryotic CO2-concentrating organelle is liquid-like and exhibits dynamic reorganization". In: *Cell* 171.1 (2017), pp. 148–162.

[12] H Wang et al. "Rubisco condensate formation by CcmM in β-carboxysome biogenesis". In: *Nature* 566.7742 (2019), pp. 131–135.

[13] Jessica Sheu-Gruttadauria and Ian J MacRae. "Phase transitions in the assembly and function of human miRISC". In: *Cell* 173.4 (2018), pp. 946–957.

[14] Tetsuro Hirose et al. "NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies". In: *Molecular biology of the cell* 25.1 (2014), pp. 169–183.

[15] Samantha K Powers et al. "Nucleo-cytoplasmic partitioning of ARF proteins controls auxin responses in Arabidopsis thaliana". In: *Molecular cell* 76.1 (2019), pp. 177–190.

[16] Ya-Cheng Liao et al. "RNA granules hitchhike on lysosomes for long-distance transport, using annexin A11 as a molecular tether". In: *Cell* 179.1 (2019), pp. 147–164.

[17] Adam Klosin et al. "Phase separation provides a mechanism to reduce noise in cells". In: *Science* 367.6476 (2020), pp. 464–468.

[18] Mingjian Du and Zhijian J Chen. "DNA-induced liquid phase condensation of cGAS activates innate immune signaling". In: *Science* 361.6403 (2018), pp. 704–709.

[19] Jiaxi Wu et al. "Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA". In: *Science* 339.6121 (2013), pp. 826–830.

[20] Joshua A Riback et al. "Stress-triggered phase separation is an adaptive, evolutionarily tuned response". In: *Cell* 168.6 (2017), pp. 1028–1040.

[21] Christiane Iserman et al. "Condensation of Ded1p promotes a translational switch from housekeeping to stress protein production". In: *Cell* 181.4 (2020), pp. 818–831.

[22] Jorge R. Espinosa et al. "Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components". In: *Proceedings of the National Academy of Sciences* 117.24 (June 2020), pp. 13238–13247. ISSN: 0027-8424.

[23] SF Banani. "Compositional control of phase-separated cell. bodies". In: *Cell* 166 (2016), pp. 651–663.

[24] Jie Wang et al. "A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins". In: *Cell* 174.3 (July 2018), pp. 688–699. ISSN: 10974172.

[25] Jeong Mo Choi, Alex S. Holehouse, and Rohit V. Pappu. "Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions". In: *Annual Review of Biophysics* 49 (May 2020), pp. 107–133. ISSN: 19361238.

[26] Jeong-Mo Choi, Anthony A Hyman, and Rohit V Pappu. "Generalized models for bond percolation transitions of associative polymers". In: *Physical Review E* 102.4 (2020), p. 042403.

[27] Georg Krainer et al. "Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions". In: *Nature Communications* 12.1 (Dec. 2021), pp. 1–14. ISSN: 20411723.

[28] Stephen E Farr et al. "Nucleosome plasticity is a critical element of chromatin liquid–liquid phase separation and multivalent nucleosome interactions". In: *Nature communications* 12.1 (2021), p. 2883.

[29] Gregory L Dignon et al. "Temperature-controlled liquid–liquid phase separation of disordered proteins". In: *ACS central science* 5.5 (2019), pp. 821–830.

[30] Omar Adame-Arana et al. "Liquid phase separation controlled by pH". In: *Biophysical Journal* 119.8 (2020), pp. 1590–1605.

[31] Matthias Christoph Munder et al. "A pH-driven transition of the cytoplasm from a fluid- to a solid-like state promotes entry into dormancy". In: *eLife* 5 (Mar. 2016). Ed. by Robert H Singer, e09347. ISSN: 2050-084X.

[32] Hasan Cinar et al. "Temperature, Hydrostatic Pressure, and Osmolyte Effects on Liquid–Liquid Phase Separation in Protein Condensates: Physical Chemistry and Biological Implications". In: *Chemistry – A European Journal* 25.57 (2019), pp. 13049–13069.

[33] Jie Wang et al. "A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins". In: *Cell* 174.3 (July 2018), pp. 688–699. ISSN: 10974172.

[34] Kadi L Saar et al. "Learning the molecular grammar of protein condensates from sequence determinants and embeddings". In: *Proceedings of the National Academy of Sciences* 118.15 (2021), e2019053118.

[35] Kiersten M Ruff et al. "Sequence grammar underlying the unfolding and phase separation of globular proteins". In: *Molecular cell* 82.17 (2022), pp. 3193–3208.

[36] Matthew R. King et al. "Uncovering molecular grammars of intrinsically disordered regions that organize nucleolar fibrillar centers". In: *bioRxiv* (2022). DOI: 10.1101/2022.11.05.515292.

[37] Rachel S Fisher and Shana Elbaum-Garfinkle. "Tunable multiphase dynamics of arginine and lysine liquid condensates". In: *Nature communications* 11.1 (2020), p. 4628.

[38] Erik W Martin et al. "Valence and patterning of aromatic residues determine the phase behavior of prion-like domains". In: *Science* 367.6478 (2020), pp. 694–699.

[39] Jerelle A. Joseph et al. "Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy". In: *Nature Computational Science* 1.11 (2021), pp. 732–743.

[40] Anne Bremer et al. "Deciphering how naturally occurring sequence features impact the phase behaviors of disordered prion-like domains". In: *Nature chemistry* 14.2 (Feb. 2022), p. 196. ISSN: 17554349.

[41] Jacob P Brady et al. "Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation". In: *Proceedings of the National Academy of Sciences* 114.39 (2017), E8194–E8203.

[42] RM Vernon. "Pi–Pi contacts are an overlooked protein feature relevant to phase separation". In: *eLife* 7 (2018).

[43] Martin J Fossat, Xiangze Zeng, and Rohit V Pappu. "Uncovering differences in hydration free energies and structures for model compound mimics of charged side chains of amino acids". In: *The Journal of Physical Chemistry B* 125.16 (2021), pp. 4148–4161.

[44] Yongdae Shin and Clifford P. Brangwynne. "Liquid phase condensation in cell physiology and disease". In: *Science* 357.6357 (Sept. 2017), eaaf4382. ISSN: 10959203.

[45] Edward Gomes and James Shorter. "The molecular language of membraneless organelles". In: *Journal of Biological Chemistry* 294.18 (2019), pp. 7115–7127.

[46] Timothy J Nott et al. "Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles". In: *Molecular cell* 57.5 (2015), pp. 936–947.

[47] Jarrett Smith et al. "Spatial patterning of P granules by RNA-induced phase separation of the intrinsically-disordered protein MEG-3". In: *Elife* 5 (2016), e21337.

[48] Erik W Martin and Alex S Holehouse. "Intrinsically disordered protein regions and phase separation: sequence determinants of assembly or lack thereof". In: *Emerging topics in life sciences* 4.3 (Dec. 2020), pp. 307–329. ISSN: 2397-8554.

[49] Ammon E Posey, Alex S Holehouse, and Rohit V Pappu. "Phase separation of intrinsically disordered proteins". In: *Methods in Enzymology*. Vol. 611. Elsevier, 2018, pp. 1–30.

[50] Shana Elbaum-Garfinkle et al. "The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics". In: *Proceedings of the National Academy of Sciences* 112.23 (2015), pp. 7189–7194.

[51] Bryan A Gibson et al. "Organization of chromatin by intrinsic and regulated phase separation". In: *Cell* 179.2 (2019), pp. 470–484.

[52] Titus M Franzmann et al. "Phase separation of a yeast prion protein promotes cellular fitness". In: *Science* 359.6371 (2018), eaao5654.

[53] Shovamayee Maharana et al. "RNA buffers the phase separation behavior of prion-like RNA binding proteins". In: *Science* 360.6391 (2018), pp. 918–921.

[54] Nieves Lorenzo Gotor et al. "RNA-binding and prion domains: the Yin and Yang of phase separation". In: *Nucleic acids research* 48.17 (2020), pp. 9491–9504.

[55] Amandine Molliex et al. "Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization". In: *Cell* 163.1 (Sept. 2015), pp. 123–133. ISSN: 10974172.

[56] Mrityunjoy Kar et al. "Phase separating RNA binding proteins form heterogeneous distributions of clusters in subsaturated solutions". In: *bioRxiv* (2022). DOI: 10.1101/2022.02.03.478969.

[57] Steven Boeynaems et al. "Protein Phase Separation: A New Phase in Cell Biology". In: *Trends in Cell Biology* 28.6 (2018), pp. 420–435. ISSN: 0962-8924.

[58] Stephanie Heinrich and Maria Hondele. "Probing liquid-liquid phase separation of RNA-binding proteins in vitro and in vivo". In: *Methods Mol. Biol.* 2537 (2022), pp. 307–333.

[59] Amalia S. Parra and Christopher A. Johnston. "Emerging Roles of RNA-Binding Proteins in Neurodevelopment". In: *Journal of Developmental Biology* 10.2 (June 2022), p. 23. DOI: 10.3390/jdb10020023. URL: https://doi.org/10.3390/jdb10020023.

[60] Zachary M. March, Oliver D. King, and James Shorter. "Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease". In: *Brain Research* 1647 (2016). RNA Metabolism in Neurological Disease, pp. 9–18. ISSN: 0006-8993.

[61] Adriano Aguzzi and Asvin KK Lakkaraju. "Cell biology of prions and prionoids: a status report". In: *Trends in cell biology* 26.1 (2016), pp. 40–51.

[62] Youhna M Ayala et al. "Human, Drosophila, and C. elegans TDP43: nucleic acid binding properties and splicing regulatory function". In: *Journal of molecular biology* 348.3 (2005), pp. 575–588.

[63] Sven Hennig et al. "Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles". In: *Journal of Cell Biology* 210.4 (Aug. 2015), pp. 529–539. DOI: 10.1083/jcb.201504117. URL: https://doi.org/10.1083/jcb.201504117.

[64] Macy L Sprunger and Meredith E Jackrel. "Prion-like proteins in phase separation and their link to disease". In: *Biomolecules* 11.7 (July 2021), p. 1014.

[65] Avinash Patel et al. "A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation". In: *Cell* 162.5 (2015), pp. 1066–1077.

[66] Masato Kato et al. "Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels". In: *Cell* 149.4 (2012), pp. 753–767.

[67] Elizabeth L Guenther et al. "Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation". In: *Nature structural & molecular biology* 25.6 (2018), pp. 463–471.

[68] Tetsuro Murakami et al. "ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function". In: *Neuron* 88.4 (2015), pp. 678–690.

[69] Michael P Hughes et al. "Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks". In: *Science* 359.6376 (2018), pp. 698–701.

[70] Hannah J Wiedner and Jimena Giudice. "It's not just a phase: function and characteristics of RNA-binding proteins in phase separation". en. In: *Nat. Struct. Mol. Biol.* 28.6 (June 2021), pp. 465–473.

[71] Yue Xiao et al. "Regulation of zebrafish dorsoventral patterning by phase separation of RNA-binding protein Rbm14". en. In: *Cell Discov.* 5.1 (July 2019), p. 37.

[72] Yi Shen et al. "Biomolecular condensates undergo a generic shear-mediated liquid-to-solid transition". en. In: *Nat. Nanotechnol.* 15.10 (Oct. 2020), pp. 841–847.

[73] Dylan T Murray et al. "Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains". en. In: *Cell* 171.3 (Oct. 2017), 615–627.e16.

[74] Feng Luo et al. "Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation". en. In: *Nat. Struct. Mol. Biol.* 25.4 (Apr. 2018), pp. 341–346.

[75] Xinrui Gui et al. "Structural basis for reversible amyloids of hnRNPA1 elucidates their role in stress granule assembly". In: *Nat. Commun.* 10.1 (May 2019), p. 2006.

[76] Hong Joo Kim et al. "Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS". In: *Nature* 495.7442 (2013), pp. 467–473.

[77] Fátima Gebauer et al. "RNA-binding proteins in human genetic disease". In: *Nature Reviews Genetics* 22.3 (Nov. 2020), pp. 185–198. DOI: 10.1038/s41576-020-00302-y. URL: https://doi.org/10.1038/s41576-020-00302-y.

[78] Simon Alberti. "Phase separation in biology". In: *Current Biology* 27.20 (2017), R1097–R1102. ISSN: 09609822.

[79] Elke Bogaert et al. "Molecular dissection of FUS points at synergistic effect of low-complexity domains in toxicity". In: *Cell reports* 24.3 (2018), pp. 529–537.

[80] Erik W Martin and Tanja Mittag. "Relationship of sequence and phase separation in protein low-complexity regions". In: *Biochemistry* 57.17 (2018), pp. 2478–2487.

[81] R. C. Ball. "Scaling in Polymer Physics". In: *Scaling Phenomena in Disordered Systems*. Ed. by Roger Pynn and Arne Skjeltorp. Boston, MA: Springer US, 1991, pp. 483–490. ISBN: 978-1-4757-1402-9. DOI: 10.1007/978-1-4757-1402-9_42. URL: https://doi.org/10.1007/978-1-4757-1402-9_42.

[82] *Scaling Concepts in Polymer Physics*. 1979.

[83] Michael Rubinstein, Ralph H Colby, et al. *Polymer physics*. Vol. 23. Oxford university press New York, 2003.

[84] Dave E. Dunstan. "The viscosity-radius relationship for concentrated polymer solutions". In: *Scientific Reports* 9.1 (Jan. 2019). DOI: 10.1038/s41598-018-36596-6. URL: https://doi.org/10.1038/s41598-018-36596-6.

[85] Isaac C. Sanchez. "Critical amplitude scaling laws for polymer solutions". In: *The Journal of Physical Chemistry* 93.19 (Sept. 1989), pp. 6983–6991. DOI: 10.1021/j100356a021. URL: https://doi.org/10.1021/j100356a021.

[86] Joseph P Clarke et al. "A comprehensive analysis of the role of hnRNP A1 function and dysfunction in the pathogenesis of neurodegenerative disease". In: *Frontiers in Molecular Biosciences* 8 (2021), p. 659610.

[87] I-Fan Wang, Lien-Szu Wu, and CK James Shen. "TDP-43: an emerging new player in neurodegenerative diseases". In: *Trends in molecular medicine* 14.11 (2008), pp. 479–485.

[88] Myungjin Jo et al. "The role of TDP-43 propagation in neurodegenerative diseases: integrating insights from clinical and experimental studies". In: *Experimental & molecular medicine* 52.10 (2020), pp. 1652–1662.

[89] Peter EA Ash et al. "TIA1 potentiates tau phase separation and promotes generation of toxic oligomeric tau". In: *Proceedings of the National Academy of Sciences* 118.9 (2021), e2014188118.

[90] Daniel J Apicco et al. "Reducing the RNA binding protein TIA1 protects against tau-mediated neurodegeneration in vivo". In: *Nature neuroscience* 21 (2018), pp. 72–80.

[91] Hao Deng, Kai Gao, and Joseph Jankovic. "The role of FUS gene variants in neurodegenerative diseases". In: *Nature Reviews Neurology* 10.6 (2014), pp. 337–348.

[92] Clotilde Lagier-Tourenne and Don W Cleveland. "Rethinking als: The fus about tdp-43". In: *Cell* 136.6 (2009), pp. 1001–1004.

[93] Julien Couthouis et al. "Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis". In: *Human molecular genetics* 21.13 (2012), pp. 2899–2911.

[94] Junghee Lee et al. "EWSR1, a multifunctional protein, regulates cellular function and aging via genetic and epigenetic pathways". In: *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* 1865.7 (2019), pp. 1938–1945.

[95] Yun R Li et al. "Stress granules as crucibles of ALS pathogenesis". In: *Journal of cell biology* 201.3 (2013), pp. 361–372.

[96] Mina Farag et al. "Phase Separation in Mixtures of Prion-Like Low Complexity Domains is Driven by the Interplay of Homotypic and Heterotypic Interactions". In: *bioRxiv* (2023). DOI: 10.1101/2023.03.15.532828.

[97] Stefan Reber et al. "The phase separation-dependent FUS interactome reveals nuclear and cytoplasmic function of liquid–liquid phase separation". In: *Nucleic Acids Research* 49.13 (July 2021), pp. 7713–7731. DOI: 10.1093/nar/gkab582. URL: https://doi.org/10.1093/nar/gkab582.

[98] Zhe Feng et al. "Formation of biological condensates via phase separation: Characteristics, analytical methods, and physiological implications". In: *Journal of Biological Chemistry* 294.40 (2019), pp. 14823–14835. ISSN: 1083351X.

[99] Timothy J Welsh et al. "Surface electrostatics govern the emulsion stability of biomolecular condensates". In: ().

[100] Seema Qamar et al. "FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interactions". In: *Cell* 173.3 (2018), pp. 720–734.

[101] Anastasia C Murthy et al. "Molecular interactions underlying liquid- liquid phase separation of the FUS low-complexity domain". In: *Nature structural & molecular biology* 26.7 (2019), pp. 637–648.

[102] Auré Lie Zbinden et al. "Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force". In: *Developmental Cell* 55 (2020), pp. 45–68.

[103] Suman Das et al. "Comparative roles of charge, π, and hydrophobic interactions in sequence-dependent phase separation of intrinsically disordered proteins". In: *Proceedings of the National Academy of Sciences of the United States of America* 117.46 (Nov. 2020), pp. 28795–28805. ISSN: 10916490.

[104] H Jane Dyson, Peter E Wright, and Harold A Scheraga. "The role of hydrophobic interactions in initiation and propagation of protein folding". In: *Proceedings of the National Academy of Sciences* 103.35 (2006), pp. 13057–13061.

[105] Charles D Andrew et al. "Stabilizing interactions between aromatic and basic side chains in α-helical peptides and proteins. Tyrosine effects on helix circular dichroism". In: *Journal of the American Chemical Society* 124.43 (2002), pp. 12706–12714.

[106] Alex S. Holehouse et al. "Clustering of Aromatic Residues in Prion-like Domains Can Tune the Formation, State, and Organization of Biomolecular Condensates". In: *Biochemistry* 60.47 (2021), pp. 3566–3581.

[107] Wei Li and Hao Jiang. "Nuclear Protein Condensates and Their Properties in Regulation of Gene Expression". In: *Journal of Molecular Biology* 434.1 (2022). Protein Interactions in Liquid-Liquid Phase Separation, p. 167151. ISSN: 0022-2836. DOI: https://doi.org/10.1016/j.jmb.2021.167151. URL: https://www.sciencedirect.com/science/article/pii/S0022283621003806.

[108] Michael Rubinstein and Andrey V Dobrynin. *Solutions of Associative Polymers*. Tech. rep. 1997.

[109] Yuri Hong et al. "Hydrophobicity of arginine leads to reentrant liquid-liquid phase separation behaviors of arginine-rich proteins". In: *Nature Communications* 13.1 (2022), p. 7326.

[110] Jamie A Greig et al. "Arginine-enriched mixed-charge domains provide cohesion for nuclear speckle condensation". In: *Molecular cell* 77.6 (2020), pp. 1237–1250.

[111] "Polyarginine enters cells more efficiently than other polycationic homopolymers." In: *Journal of Peptide Research* 56.5 (2000), pp. 318–325.

[112] Xiangze Zeng, Kiersten M. Ruff, and Rohit V. Pappu. "Competing interactions give rise to two-state behavior and switch-like transitions in charge-rich intrinsically disordered proteins". In: *Proceedings of the National Academy of Sciences* 119.19 (May 2022). DOI: 10.1073/pnas.2200559119. URL: https://doi.org/10.1073/pnas.2200559119.

[113] Jeong-Mo Choi and Rohit V Pappu. "The stickers and spacers framework for describing phase behavior of multivalent intrinsically disordered proteins". In: *Biophysical Journal* 118.3 (2020), 492a.

[114] David De Sancho. "Phase separation in amino acid mixtures is governed by composition". In: *Biophysical Journal* 121.21 (2022), pp. 4119–4127.

[115] "Hydropathy Patterning Complements Charge Patterning to Describe Conformational Preferences of Disordered Proteins." In: *Journal of Physical Chemistry Letters* 11.9 (2020), pp. 3408–3415.

[116] Benjamin G Weiner et al. "Motif-pattern dependence of biomolecular phase separation driven by specific interactions". In: *PLoS computational biology* 17.12 (2021), e1009748.

[117] Megan C. Cohan et al. "Uncovering Non-random Binary Patterns Within Sequences of Intrinsically Disordered Proteins". In: *Journal of Molecular Biology* 434.2 (2022), p. 167373. ISSN: 0022-2836. DOI: https://doi.org/10.1016/j.jmb.2021.167373. URL: https://www.sciencedirect.com/science/article/pii/S0022283621006100.

[118] Jonathan Huihui and Kingshuk Ghosh. "Intrachain interaction topology can identify functionally similar intrinsically disordered proteins". In: *Biophysical Journal* 120.10 (2021), pp. 1860–1868. ISSN: 0006-3495.

[119] Kingshuk Ghosh et al. "Rules of Physical Mathematics Govern Intrinsically Disordered Proteins". In: *Annual Review of Biophysics* 51.1 (May 2022), pp. 355–376. DOI: 10.1146/annurev-biophys-120221-095357. URL: https://doi.org/10.1146/annurev-biophys-120221-095357.

[120] Ushnish Rana, Clifford P. Brangwynne, and Athanassios Z. Panagiotopoulos. "Phase separation vs aggregation behavior for model disordered proteins". In: *The Journal of Chemical Physics* 155.12 (Sept. 2021), p. 125101. DOI: 10.1063/5.0060046. URL: https://doi.org/10.1063/5.0060046.

[121] Erik W Martin et al. "Sequence determinants of the conformational properties of an intrinsically disordered protein prior to and upon multisite phosphorylation". In: *Journal of the American Chemical Society* 138.47 (2016), pp. 15323–15335.

[122] Rahul K. Das and Rohit V. Pappu. "Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues". In: *Proceedings of the National Academy of Sciences* 110.33 (2013), pp. 13392–13397.

[123] Roshan Mammen Regy et al. "Improved coarse-grained model for studying sequence dependent phase separation of disordered proteins". In: *Protein Science* 30.7 (May 2021), pp. 1371–1379. DOI: 10.1002/pro.4094. URL: https://doi.org/10.1002/pro.4094.

[124] Gregory L. Dignon et al. "Sequence determinants of protein phase behavior from a coarse-grained model". In: *PLoS Computational Biology* 14.1 (Jan. 2018). ISSN: 15537358.

[125] Giulio Tesei et al. "Accurate model of liquid–liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties". In: *Proceedings of the National Academy of Sciences* 118.44 (Oct. 2021). DOI: 10.1073/pnas.2111696118. URL: https://doi.org/10.1073/pnas.2111696118.

[126] Roshan Mammen Regy, Wenwei Zheng, and Jeetain Mittal. "Using a sequence-specific coarse-grained model for studying protein liquid–liquid phase separation". In: *Methods in Enzymology*. Elsevier, 2021, pp. 1–17. DOI: 10.1016/bs.mie.2020.07.009. URL: https://doi.org/10.1016/bs.mie.2020.07.009.

[127] Andrew P. Latham and Bin Zhang. "Maximum Entropy Optimized Force Field for Intrinsically Disordered Proteins". In: *Journal of Chemical Theory and Computation* 16.1 (Nov. 2019), pp. 773–781. DOI: 10.1021/acs.jctc.9b00932. URL: https://doi.org/10.1021/acs.jctc.9b00932.

[128] Sohee Park et al. "Molecular Hydration Tunes the Cation–π Interaction Strength in Aqueous Solution". In: *Advanced Materials Interfaces* 10.4 (Dec. 2022), p. 2201732. DOI: 10.1002/admi.202201732. URL: https://doi.org/10.1002/admi.202201732.

[129] Justin P. Gallivan and Dennis A. Dougherty. "A Computational Study of Cation-π Interactions vs Salt Bridges in Aqueous Media: Implications for Protein Engineering". In: *Journal of the American Chemical Society* 122.5 (Jan. 2000), pp. 870–874. DOI: 10.1021/ja991755c. URL: https://doi.org/10.1021/ja991755c.

[130] Tina Ukmar-Godec et al. "Lysine/RNA-interactions drive and regulate biomolecular condensation". In: *Nature Communications* 10.1 (July 2019). DOI: 10.1038/s41467-019-10792-y. URL: https://doi.org/10.1038/s41467-019-10792-y.

[131] A. P. Thompson et al. "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales". In: *Comp. Phys. Comm.* 271 (2022), p. 108171.

[132] Niels Grønbech-Jensen and Oded Farago. "A simple and effective Verlet-type algorithm for simulating Langevin dynamics". In: *Molecular Physics* 111.8 (2013), pp. 983–991.

[133] H. J. C. Berendsen et al. "Molecular dynamics with coupling to an external bath". In: *The Journal of Chemical Physics* 81.8 (1984), pp. 3684–3690.

[134] John Shipley Rowlinson and Benjamin Widom. *Molecular theory of capillarity*. Courier Corporation, 2013.