Modulation of α-Synuclein Aggregation Amid Diverse Environmental Perturbation

Abdul Wasim[1], Sneha Menon[1], Jagannath Mondal[1*]

[1] *Tata Institute of Fundamental Research, Hyderabad, India-500046*

E-mail: jmondal@tifrh.res.in, phone:+91-40-2020-3091

Abstract

Intrinsically disordered protein α-Synuclein (αS) is implicated in Parkinson's disease due to its aberrant aggregation propensity. In a bid to identify the traits of its aggregation, here we computationally simulate the multi-chain association process of αS in aqueous as well as under diverse environmental perturbations. In particular, the aggregation of αS in aqueous and varied environmental condition led to marked concentration differences within protein aggregates, resembling liquid-liquid phase separation (LLPS). Both saline and crowded settings enhanced the LLPS propensity. However, the surface tension of αS droplet responds differently to crowders (entropy-driven) and salt (enthalpy-driven). Conformational analysis reveals that the IDP chains would adopt extended conformations within aggregates and would maintain mutually perpendicular orientations to minimize inter-chain electrostatic repulsions. The droplet stability is found to stem from a diminished intra-chain interactions in the C-terminal regions of αS, fostering inter-chain residue-residue interactions. Intriguingly, a graph theory analysis identifies *small-world-like networks* within droplets across environmental conditions, suggesting the prevalence of a consensus interaction patterns among the chains. Together these findings suggest a delicate balance between molecular grammar and environment-dependent nuanced aggregation behaviour of αS.

Introduction

In the human body, a significant presence of Intrinsically Disordered Proteins (IDPs) plays diverse and crucial roles.[1–3] These proteins lack a well-defined 3D structure under native conditions, which imparts functional advantages, but also renders them susceptible to irreversible aggregation, especially when affected by mutations. Such aggregates can be pathogenic and are associated with various diseases, including neurodegenerative diseases, cancer, diabetes, and cardiovascular diseases.[4]

Notably, Alzheimer's disease is characterized by the aggregation of the amyloid-β peptide (Aβ), while Parkinson's disease (PD) is linked to α-Synuclein (αS) aggregation. A growing body of evidence has established a connection between IDPs and the phenomenon known as *liquid-liquid phase separation* (LLPS). During LLPS, high and low concentrations of biomolecules coexist without the presence of membranes and exhibit properties similar to phase-separated liquid droplets of two immiscible liquids (Figure 1).[5–9] This intriguing phenomenon has garnered significant attention as it underlies the formation of membrane-less subcellular compartments,[10–12] which, when dysregulated, can lead to incurable pathogenic diseases.

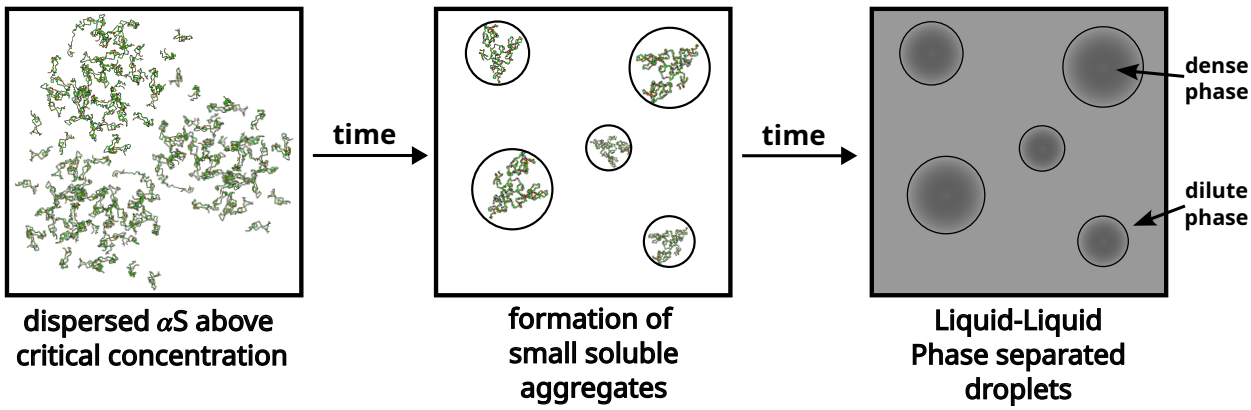

Figure 1: A schematic showcasing the process of liquid-liquid phase separation of αS.

Recent findings have highlighted the capability of αS to undergo LLPS under physiological conditions, specifically when the protein concentration surpasses a critical threshold.[6]

Moreover, it was observed that the aggregation propensity of αS is significantly influenced by various factors, including the presence of molecular crowders, the ionic strength of the protein environment, and pH.[13] Nonetheless, characterizing the interactions and dynamics of these small aggregates poses experimental challenges, leading to limited available reports on the subject.[14–17]

This investigation aims to establish the molecular basis of self-aggregation of αS and underlying process of LLPS under diverse environmental perturbations. In particular, to understand the influence of environmental factors on the inter-protein interactions within a phase separated droplet, we target to computationally simulate the aggregation process of αS under different conditions, emphasizing the roles of crowders and salt. While recent progress in computational forcefields and hardware has enabled the simulation of individual Intrinsically Disordered Proteins (IDPs) especially αS, using All-Atom Molecular Dynamics (AAMD),[18–24] these simulations can be extremely time-consuming and resource-intensive, making multi-chain AAMD simulations, even with cutting-edge software and hardware, impractical. Therefore, to simulate the the aggregation process, we resort to coarse-grained molecular dynamics (CGMD) simulations. Leveraging a tailored Martini 3 Coarse-Grained Force Field (CGFF)[25] for αS, we dissect inter-protein interactions governing stable aggregate formation and LLPS. Optimizing water-protein interactions for αS within the CGFF framework, our multi-chain microsecond-long CGMD simulations yield comprehensive ensembles of substantial protein aggregates across diverse contexts.

As one of the key observations, our simulation unequivocally reveal LLPS-like attributes in the aggregates and show how these get modulated in presence of crowders and salt. The investigation unearths the intricate interplay of mechanical and thermodynamic forces in αS aggregation, achieved through meticulous data analyses. We elucidate the pivotal intra and inter-protein interactions governing LLPS-like protein droplet formation, unveiling the protein's primary sequence's role in aggregation. As would be shown in this article, a graph-based depiction of the droplet's architecture represents the proteins within droplets

as constituting dense networks akin to *small-world networks.*

Results

In this study, we utilized the recently developed Martini 3[25] coarse-grained model to simulate collective interaction of a large number of αS chains in explicit presence of aqueous media at various concentrations commensurate with *in vitro* conditions including the presence of crowders and salt. As Martini 3 was not originally developed for intrinsically disordered proteins (IDPs), we carefully optimized the protein-water interactions against atomistic simulation of monomer and dimer of αS, as detailed in the *Methods* section, to ensure compatibility with αS (see *Methods*).

Initially, we examined the impact of concentration on the protein's aggregation by simulating copies of chains, maintaining a polydispersity of protein conformations of αS. In particular, three different conformations of αS (referred here as $ms1$, $ms2$ and $ms3$) with R_gs (radius of gyration) ranging between collapsed and extended states (1.84-5.72 nm) at different concentrations, with a composition, as estimated in a recent investigation,[23] were employed. First the chains were simulated for extensive period in a set of three protein concentrations, close to previous experiments.

Simulations capture enhanced aggregation beyond a threshold concentrations of αS

We performed simulations of αS at various concentrations, namely 300 μM, 400 μM, 500 μM and 750 μM. We begin by analysing the aggregation behavior of αS. As shown in Figure 2, we observe that most chains do not aggregate at 300 and 400 μM as characterized by the prevalence of high number of free monomers. The respective snapshots of the simulation indicate the presence of greater extent of single chains. Also, the chains that are not free, form very small oligomers of the order of dimer to tetramer (Figure 2).

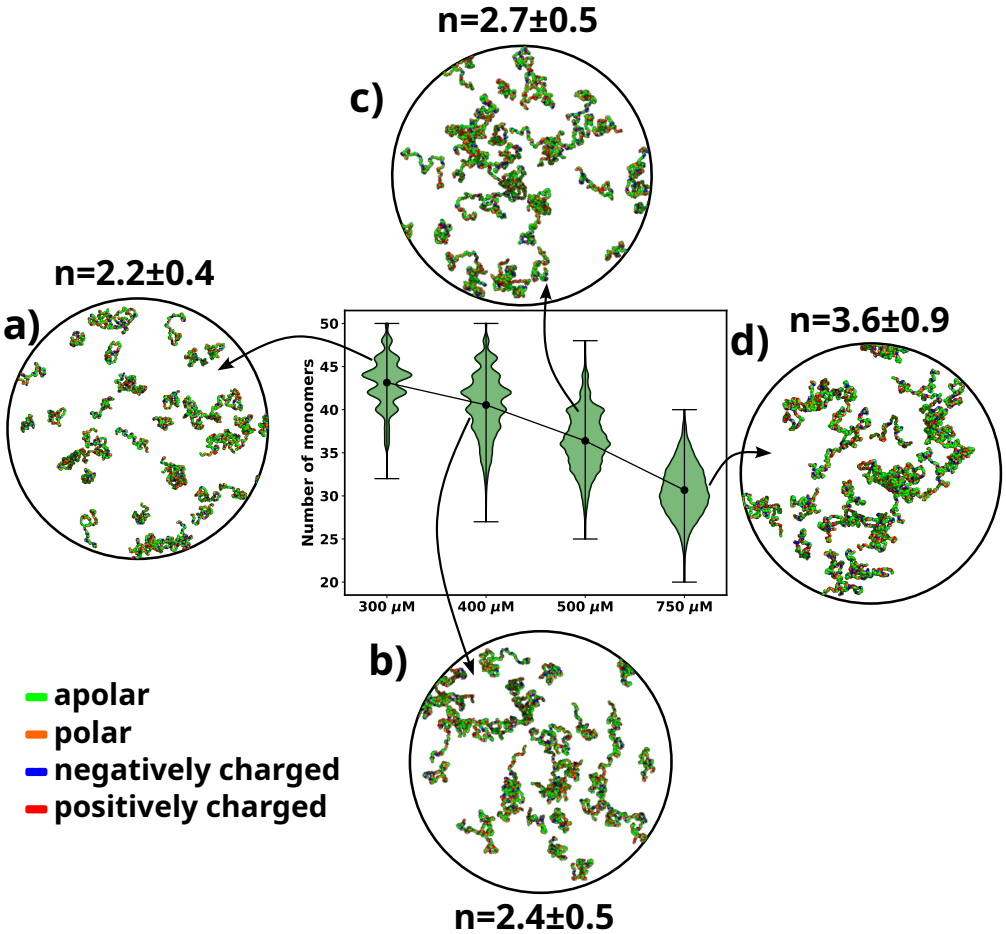

Figure 2: A violinplot showing the distribution of number of monomers present for different concentrations of α-syn. The blue dot at the middle of each distribution represents the mean number of monomers observed for each concentration. For each concentration we show representative snapshots of the system. For each concentration, we also report the statistics of the number of chains in the largest cluster (n). **a)** A snapshot from the simulation at 300 μM α-syn. **b)** A snapshot from the simulation at 400 μM α-syn. **c)** A snapshot from the simulation for 500 μM α-syn. **d)** A snapshot from the simulation at 750 μM α-syn. Each chain in the snapshots has been coloured differently.

However, upon increasing the concentration to 500 μM, which has also been the critical concentration reported for αS to undergo LLPS,[6] we observe a sharp drop in the average number of free monomers in the system (Figure 2). The corresponding representative snapshot of the system also depicts a few higher order aggregates, such as pentamers and hexamers, as well as most chains forming small oligomers. This can be understood from the value of the average number of chains present in the largest clusters, as reported in Figure 2.

The system, being at critical concentration, formation of large aggregates would require longer timescales than the simulation length. Therefore, in order to promote formation of large aggregates (heptamers or more) for finer characterization, we performed a simulation at a higher concentration of 750 μM αS. As shown in Figure 2d, we observe further decrease in the total number of free monomeric chains in the solution. There is simultaneous appearance of a very few droplet like aggregates (hexamer or more) as can be seen from Figure 2 and the adjacent snapshot of the system (Figure 2). However, we note that $\sim$60 % of the protein chains are free and do not participate in aggregation and we think that as such in water, αS does not possess a strong and spontaneous self-aggregation tendency. In the following sections we characterize the aggregation tendency of αS in presence of certain environmental modulator that can shed more light on this hypothesis.

Molecular crowders and salt accelerate αS aggregation

The cellular environment, accommodating numerous biological macromolecules, poses a highly crowded space for proteins to fold and function.[26–29] In *in vitro* studies, inert polymers such as Dextran, Ficoll, and polyethylene glycol (PEG) are commonly employed as macromolecular crowding agents. In the context of αS amyloid aggregation, previous experimental studies have revealed an increased rate of *in vitro* fibrillation in the presence of different crowding agents.[30–32] Notably, a recent experimental study demonstrated the occurrence of phase separation (LLPS) of αS in the presence of PEG molecular crowder.[6] Moreover, considering that *in-vivo* environments also contain various moieties like salts and highly charged ions, a recent *in vitro* study has shown that the ionic strength of the solvent directly influences the aggregation rates of αS,[13] with higher ionic strength enhancing αS aggregation. Given these observations, it becomes crucial to characterize the factors responsible for the enhanced aggregation of αS in the presence of crowders and salt. To address this, we perform two independent sets of simulations: one with αS present at 750 μM in the presence of 10% (v/v) fullerene-based crowders (see *SI Methods*) and the other with the same concentration

of αS but in the presence of 50 mM of NaCl. In this section we characterize the effects of addition of crowders or salt on the aggregation of αS.

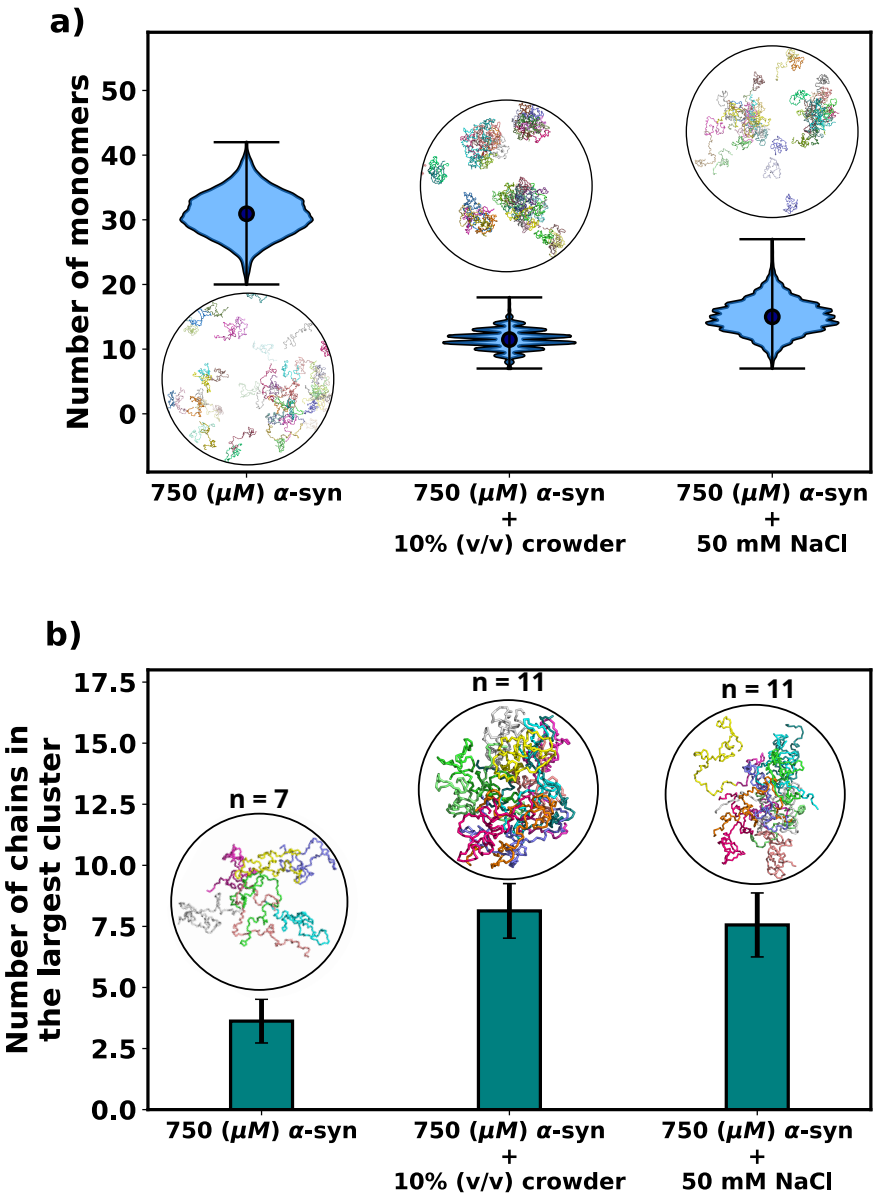

Figure 3: **a)** A violinplot showing the distribution of the number of monomers for α-syn at 750 μM without and with crowder. The blue dots represent the means of each distribution. The snapshots represent the extent aggregation for a visual comparison. **b)** A bar plot showing the number of chain in the largest cluster formed by α-syn at 750 μM without and with crowder. The snapshots show the largest cluster formed for each scenario.

As expected, the addition of crowders leads to an enhancement of αS aggregation due to their volume excluded effects, as depicted in Figure 3a. Notably, the number of monomers

drastically decreases upon the inclusion of crowders. This observation is further supported by the snapshots of the system, which also confirm the reduction in monomer count. Similarly, we observe that the presence of salt also promotes αS aggregation, as illustrated in Figure 3a, where the number of monomers is lower when compared to the case with no salt.

Following this, we conducted an analysis of the number of chains present in the largest clusters that formed. Figure 3b clearly illustrates that the addition of crowder or salt leads to a notable increase in the average number of proteins forming a cluster. This crucial observation points to the fact that the inclusion of accelerators, such as crowder or salt, not only promotes aggregation but also plays a role in stabilizing the formed oligomers. Importantly, we observed that the effect of crowder on aggregation is slightly more pronounced compared to that of the salt. In the subsequent section, we delve into the reasons behind the enhanced aggregation induced by these accelerators, aiming to decipher the underlying mechanisms responsible for their influence on αS aggregation dynamics. As the aggregation is significant enough for performing quantitative analysis only when the concentration of αS is 750 μM, we perform all analysis on scenarios at 750 μM of αS.

Crowders and salt differentially modulate surface tension for promoting LLPS-like αS droplets

The preceding sections underscore our simulation-based observation that, influenced by crowders and salt, αS aggregates into higher-order oligomers (hexamers and beyond) at a significantly accelerated propensity compared to the scenario without these influences. Here, we delve into the investigation of the energetic aspects underlying this aggregation phenomenon. An important contributor to the energetics is the surface tension, arising from the creation of interfaces between the dense and dilute phases of the protein upon droplet formation. This presence of interfaces is accompanied by surface tension and surface energy. The surface energy of a system is directly proportional to its surface area; systems with higher surface energy tend to minimize their surface area. Consequently, systems comprising multiple

smaller droplets exhibit a larger surface area, and hence a higher surface energy. Conversely, systems characterized by fewer, larger droplets possess a comparatively reduced surface area and correspondingly lower surface energy. This insight leads us to conjecture that surface tension could play a pivotal role in driving liquid-liquid phase separation (LLPS) and the formation of larger αS droplets. To explore this hypothesis, we calculate the surface tension of the resultant droplets, as per Eqs-1, 2 and as described *SI Methods* and Ref.[33]

$$\gamma_{20} = \frac{5k_BT}{16\pi\langle(\delta a + \delta b)^2\rangle} \tag{1}$$

$$\gamma_{22} = \frac{15k_BT}{16\pi\langle(\delta a - \delta b)^2\rangle} \tag{2}$$

where $\delta a = a - R$ and $\delta b = b - R$ is the perturbation of the droplet shape from a perfect sphere with a radius R along any two pairs of principle axes of general ellipsoid estimating the shape of the droplet. The surface tension (γ) is thus estimated using $\gamma \approx \gamma_{20} \approx \gamma_{22}$. Please see *SI Methods* and Ref.[33] for more details.

Figure 4a provides a comparison of the surface tension (γ), for three different scenarios involving αS: i) αS in solution, ii) αS in the presence of crowders, and iii) αS in the presence of salt. Notably, in each case, the surface tension is considerably lower (0.0035-0.0075 mN/m) than the surface tension for FUS droplets in water ($\sim$0.05 mN/m).[33] As stated earlier, the magnitude of surface tension is an estimate of the aggregation tendency of any liquid-liquid mixture. Since we find that $\gamma_{\alpha S}$ is much lower than γ_{FUS}, we assert that the propensity with which αS aggregates should be much lower than that of FUS.

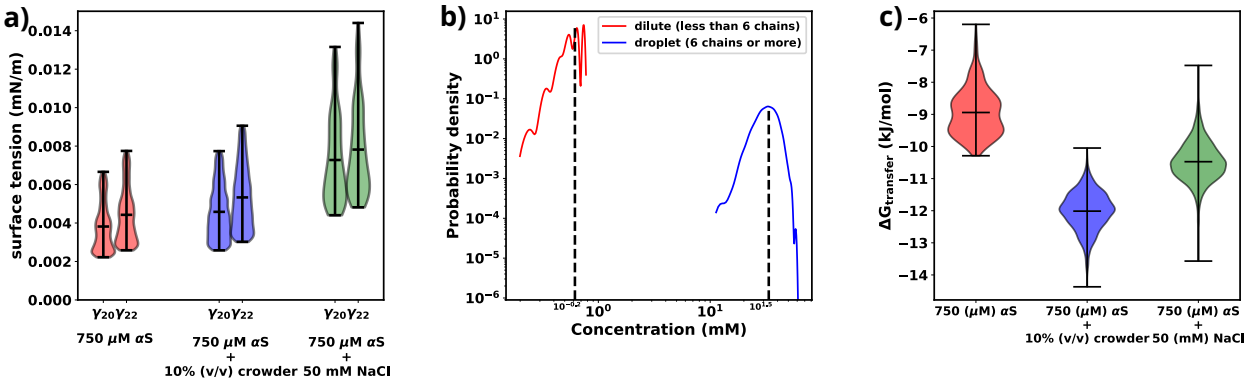

Figure 4: **a)** surface tensions of droplets, estimated from γ_{20} and γ_{22}, for three cases has been shown. Both γ_{20} and γ_{22} provide almost similar estimates of the value of surface. **b)** Comparison of protein concentrations for the dilute(red) and the droplet (blue) phases. **c)** Excess free energy of transfer comparison for three cases.

Next, we conduct a comparison of the three different scenarios to understand the effects of crowders and salt on the aggregation of αS. From Figure 4a, it is evident that the surface tensions are very similar for cases (i) and (ii), while it has increased for case (iii). This implies that the addition of crowders does not significantly impact the surface tension of the aggregates, although it renders the protein more prone to aggregation. On the other hand, the addition of salt causes an increase in surface tension. Given the relationship between surface area and volume, where a higher surface-to-volume ratio signifies numerous smaller droplets, the surface energy is concurrently elevated. In the presence of salt, a tendency is observed for these smaller aggregates to coalesce, giving rise to larger aggregates, albeit in reduced numbers. This behavior is an endeavor to curtail the surface-to-volume ratio and thus mitigate the associated surface energy. Therefore, the larger the surface tension, the higher is tendency of the protein to form aggregates, as seen from the surface tension values of αS and FUS, as mentioned earlier.

To minimize the surface energy, fusion of aggregates, either via merging of two or more droplets into one is seen for liquid-like phase separated droplets in experiments.[6] Although droplet fusion was not observed in our simulations due to the limited system size, it was shown that if a protein undergoes LLPS, a significant difference in protein concentration occurs between the droplet and the dilute phase.[34] To verify whether the aggregates observed

in our simulations exhibit characteristics of LLPS, we calculated the protein concentrations in the dilute and concentrated phases. For the droplet phase, the concentration of the protein was calculated using Eq-3.

$$c_{phase} = \frac{N_{phase}}{N_A \times V_{phase}} \tag{3}$$

where N_{phase} is the number of protein chains in the phase (here dilute or concentrated), N_A is Avogadro's number and V_{phase} is the volume occupied by the phase. For the dilute phase, we estimated the volume of the concentrated/dense phase (V_{dense}) using Eq-4.[34]

$$V^i_{dense} = 4\pi\sqrt{3}\lambda_1\lambda_2\lambda_3 \tag{4}$$

where V^i_{dense} is the volume of the i-th droplet, λ_1, λ_2 and λ_3 are the eigenvalues of the gyration tensor for the aggregate. The volume of the dilute phase is the remainder volume of the system given by Eq-5.

$$V_{dilute} = V - \sum_i V^i_{dense} \tag{5}$$

where V is the total volume of the system.

As shown in Figure 4b and Figure S1, there is an almost two orders of magnitude difference between the concentration of αS in the dilute and droplet phases for all scenarios. Such a pronounced difference is a hallmark of LLPS, leading us to assert that the aggregates formed in our simulations possess LLPS-like properties. Consequently, we use the term "droplet" interchangeably with "aggregates" for the remainder of our investigation.

$$\Delta G_{transfer} = RT \ln\left(\frac{c_{dilute}}{c_{dense}}\right) \tag{6}$$

Finally, utilizing the calculated concentrations, we proceed to estimate the excess free energy of monomer transfer ($\Delta G_{transfer}$), from Eqn-6, between the dilute and droplet phases,

where c_{dilute} is the concentration of αS in the dilute phase, c_{dense} is the concentration of αS in the dense/droplet phase, R is the universal gas constant and T is the temperature of the system (=310.15 K). As illustrated in Figure 4c, both crowder and salt scenarios demonstrate lower $\Delta G_{transfer}$ values compared to the case without their presence. However, the thermodynamic origins behind this pronounced aggregation differ for crowders and salt. Crowders enhance aggregation primarily through excluded volume interactions, which are of an entropic nature. On the other hand, salt enhances aggregation by increasing the droplet's surface tension, thus contributing to the enthalpy of the system. As a result, apart from the already known fact that macromolecular crowding decreases $\Delta G_{transfer}$ via entropic means, we also infer that salt decreases $\Delta G_{transfer}$ via enthalpic means by increasing the surface tension of the formed droplets.

Aggregation results in chain expansion and chain reorientation in αS

An indicative trait of molecules undergoing Liquid-Liquid Phase Separation (LLPS) is the adoption of extended conformations upon integration into a droplet structure. Given that the aggregates observed in our simulations exhibit a concentration disparity reminiscent of LLPS between the dilute and dense phases, we endeavored to validate the presence of a comparable chain extension phenomenon within our simulations.[34] To address this, we quantified the radius of gyration (R_g) for individual chains and classified them based on whether they were situated in the dilute or dense phase. The distribution of R_g values for each category is illustrated in Figure 5a and Figure S2. Remarkably, the distribution associated with the dense phase distinctly indicates that the protein assumes an extended conformation within this context. As elucidated earlier, this marked propensity for extended conformations aligns with a characteristic hallmark of LLPS as previously seen in experiments.[35]

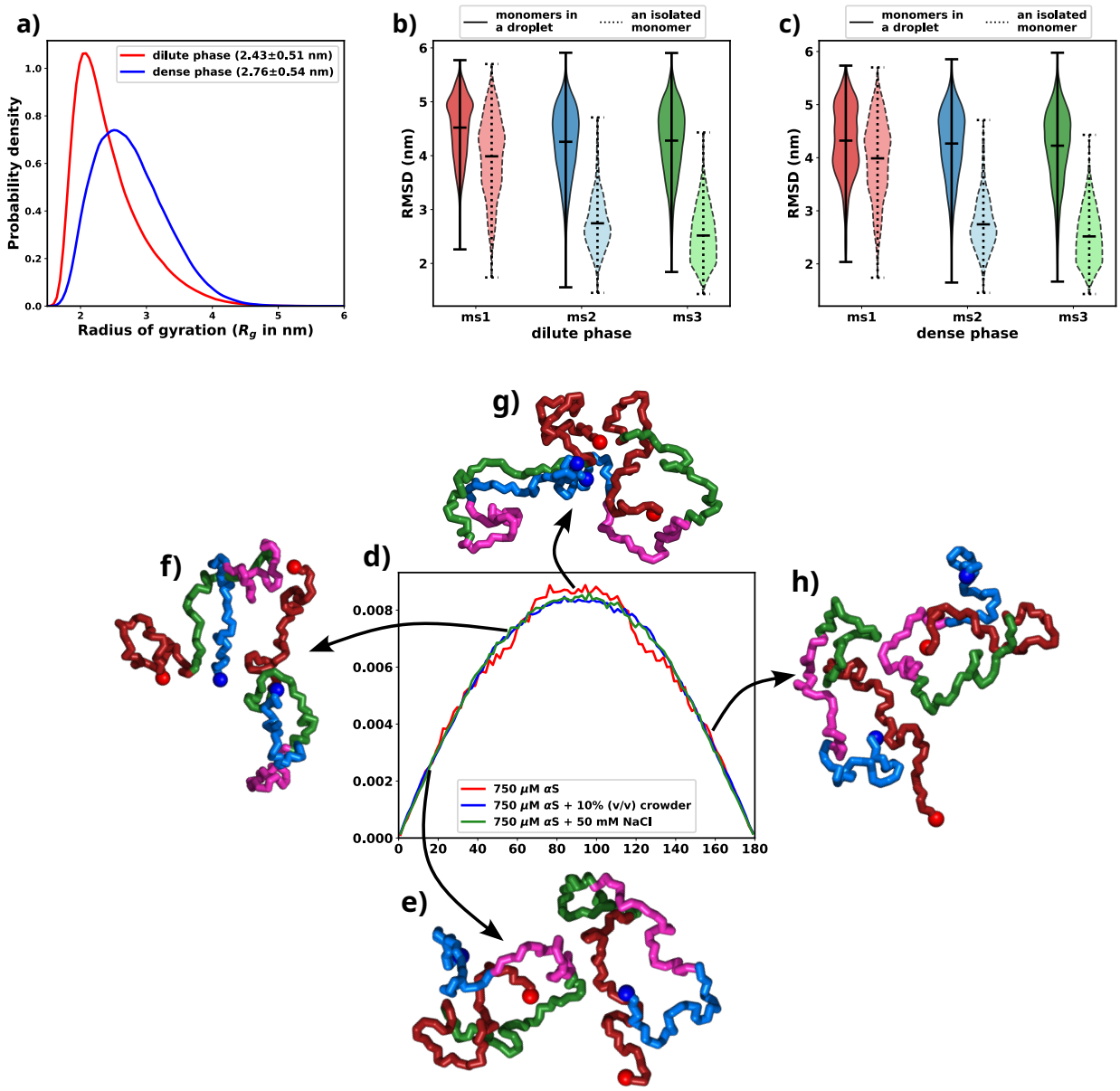

Figure 5: All the figures are for 750 μM αS + 50 mM NaCl. **a)** Distribution of R_g for proteins present in the dense or the dilute phases. **b)** Comparison of RMSD for protein chains present in the dilute phase, with single chain RMSDs as the reference (dotted edges). **c)** Comparison of RMSD for protein chains present in the dense phase, with single chain RMSDs as the reference (dotted edges). **d)** Distribution of the angle of orientation of two chains inside the droplet for the three different scenarios. **e)** representative snapshot for angle between 0 degree and 20 degree. **f)** representative snapshot for angle between 50 degree and 70 degree. **g)** representative snapshot for angle between 80 degree and 120 degree. **h)** representative snapshot for angle between 150 degree and 180 degree.

Having observed the conformational alterations of αS during LLPS, our subsequent aim was to quantify the extent of these conformational changes in relation to their initial states (referred as '$ms1$', '$ms2$' or '$ms3$' in decreasing order of R_g[23]). To achieve this, we computed the Root Mean Square Deviation (RMSD) of each protein relative to its starting conformation. The resulting distributions were visually depicted using violin plots, featuring bold edges in Figures 5b and c. The protein ensemble was segregated into two categories: i) those from the dilute phase (Figure 5b), and ii) those from the dense phase (Figure 5c).

Surprisingly, regardless of their initial configurations, the observed RMSD values were notably high. To facilitate a comparative analysis, we also included distributions of RMSDs for single chains simulated in the presence of 50 mM of salt, depicted using violin plots with broken edges. Intriguingly, the conformational state labeled as $ms1$, exhibited the least RMSD, a characteristic attributed to its notably extended conformation. This phenomenon aligns with the preference of droplets for extended conformations, implying that ms1 required the least conformational perturbation and thus exhibited a lower RMSD.

For both ms2 and ms3, a conspicuous increase in RMSD values was observed across all proteins monomers, irrespective of their respective phases. This phenomenon can potentially be attributed to the pronounced conformational shift experienced by the protein during aggregation. Building on these observations, we put forward a hypothesis: LLPS engenders significant modifications in the native protein conformations, ultimately favouring the adoption of extended states.

As discussed in the previous paragraph that the αS monomers inside the droplets must undergo conformational expansion and we hypothesized that they adopt orientation so as to minimize the inter-chain electrostatic repulsions. To this end, we try to decipher the orientations of the chains via defining their axes of orientations and subsequently calculating the angles between the major axis of two monomers. We calculate the major axis of gyration, given by the eigenvector corresponding to the largest eigenvalue of the gyration tensor, for each monomer inside a droplet. We next find the nearest neighbour (minimum distance of

approach < 8 Å) for each monomer, carefully taking care of over-counting.

The angle between two monomers is defined as the angle between the major axes of gyration between chain i and its nearest neighbour j. We plot the distributions of the angles for all scenarios and all droplets in Figure 5d. We observe that irrespective of the conditions, the distribution peaks at right angles. The representative snapshots (Figure 5e-5h) showcase their mode of orientation. Interestingly the distribution is the same for all the three scenarios, again stressing upon the fact the αS droplets share similar features in terms of interactions and orientations irrespective of their environments.

Characterization of molecular interactions in aggregation-prone conditions

As established in preceding sections, both crowders and salt have been observed to augment the aggregation of αS while concurrently stabilizing the resultant aggregates. This phenomenon leads to the protein adopting extended conformations within a notably heterogeneous ensemble. Shifting our attention, we now delve into a residue-level investigation to unravel the specific interactions responsible for stabilizing these aggregates and, consequently, facilitating the aggregation process.

To compute the differential contact maps, our approach involved initial calculations of average intra-protein residue-wise contact maps, termed as intra-protein contact probability maps, for monomers present in both the dilute and dense phases (refer to Figures S3 and S4). Subsequently, we derived the difference by subtracting the contact probabilities of monomers within the dilute phase from those within the dense phase. As evident from Figure 6a, a discernible reduction in intra-chain Nter-Cter interactions is observed for monomers within the droplet phase, depicted by the presence of blue regions along the off-diagonals. Such a reduction in such interactions has also been observed via experiments[35] and it is similarly noticeable in the two other cases, as evident in Figures 6b and 6c.

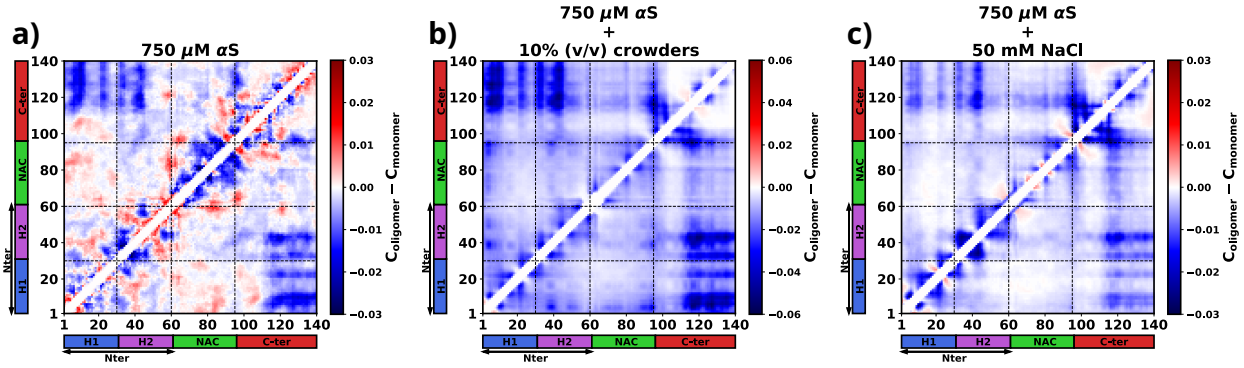

Figure 6: The figure presents the residue-wise, intra-protein difference contact maps where the average contact probability of monomers in the dilute phase were subtracted from the average contact probability of monomers in the dense/droplet phase for three cases: **a)** 750 μM αS in water. **b)** 750 μM αS in the presence of 10% (v/v) crowders. **c)** 750 μM αS in presence of 50 mM NaCl.

Furthermore, a significant decline in intra-protein interactions, especially the NAC-NAC interactions, is predominantly observed at shorter ranges, indicated by deep-blue regions concentrated near the diagonals. Notably, these diminished intra-chain interactions potentially facilitate the formation of inter-chain interactions, as depicted in Figure S5. Building on these observations, we posit that these interactions play a pivotal role in stabilizing the aggregates that have formed.

Moreover, from the difference heatmaps in Figure S6, it can be observed that the residues 95-110 (VVKKKKKKDDQQLLGKKKDDEEEEGAAPPQQEE) have reduced contact probabilities upon introduction of crowders/salt, whereas the rest of the contacts have slightly increased. These residues are highly charged and we think that upon introduction of crowders/salt, the proteins inside the droplet needed to be spatially oriented to facilitate the formation of largest aggregates. This re-orientation occurs to minimize the electrostatic repulsions among these residues belonging to different chains. These analyses provide hints that these residues are present in the protein so as to avoid the formation of aggregation prone conformations, which is why their interactions had to be minimized to form more stable and larger aggregates.

Phase separated αS monomer form *small-world* networks

The investigations so far suggest that irrespective of the factors that cause the aggregation of αS, the interactions that drive the formation of droplet remain essentially the same. However the conformations of the monomers vary depending on their environment. In presence of crowders they adapt to form much more compact aggregates. Therefore here we characterize whether the environment influences the connectivity among different chains of the protein in a droplet.

Figure 7a, 7b and 7c show molecular representations of the largest cluster formed by αS at 750 μM in water, αS at 750 μM in presence of 10% (v/v) crowders and αS at 750 μM in presence of 50 mM NaCl respectively. From the molecular representations for aggregates, it can be seen that irrespective of the system, they form a dense network whose characterization is not possible directly. Therefore we represent each aggregate as a graph with multiple nodes (vertices) and connections (edges), as can be seen from Figure 7d, 7e and 7f. Each node (in blue) represents a monomer in the droplet. Two nodes have an edge (line connecting two nodes) if the minimum distance of approach of the monomers corresponding to the pair of nodes is at least 8 Å. We can see from the graph that not all chains are in contact with each other. They rather form a relay where a few monomers connect (interact) with most of the other protein chains. The rest of the chains have indirect connections via those. Since inter-chain connections/interactions have been denoted by edges and the chains themselves as nodes, such form of inter-chain interactions inside a droplet lead to only a few nodes having a lot of edges, for example node 1 in Figure 7f. The rest of them have only a few (3-5) edges. This is a signature of *small-world networks*[36–39] and we assert that αS inside the droplet(s) form small-world-like networks.

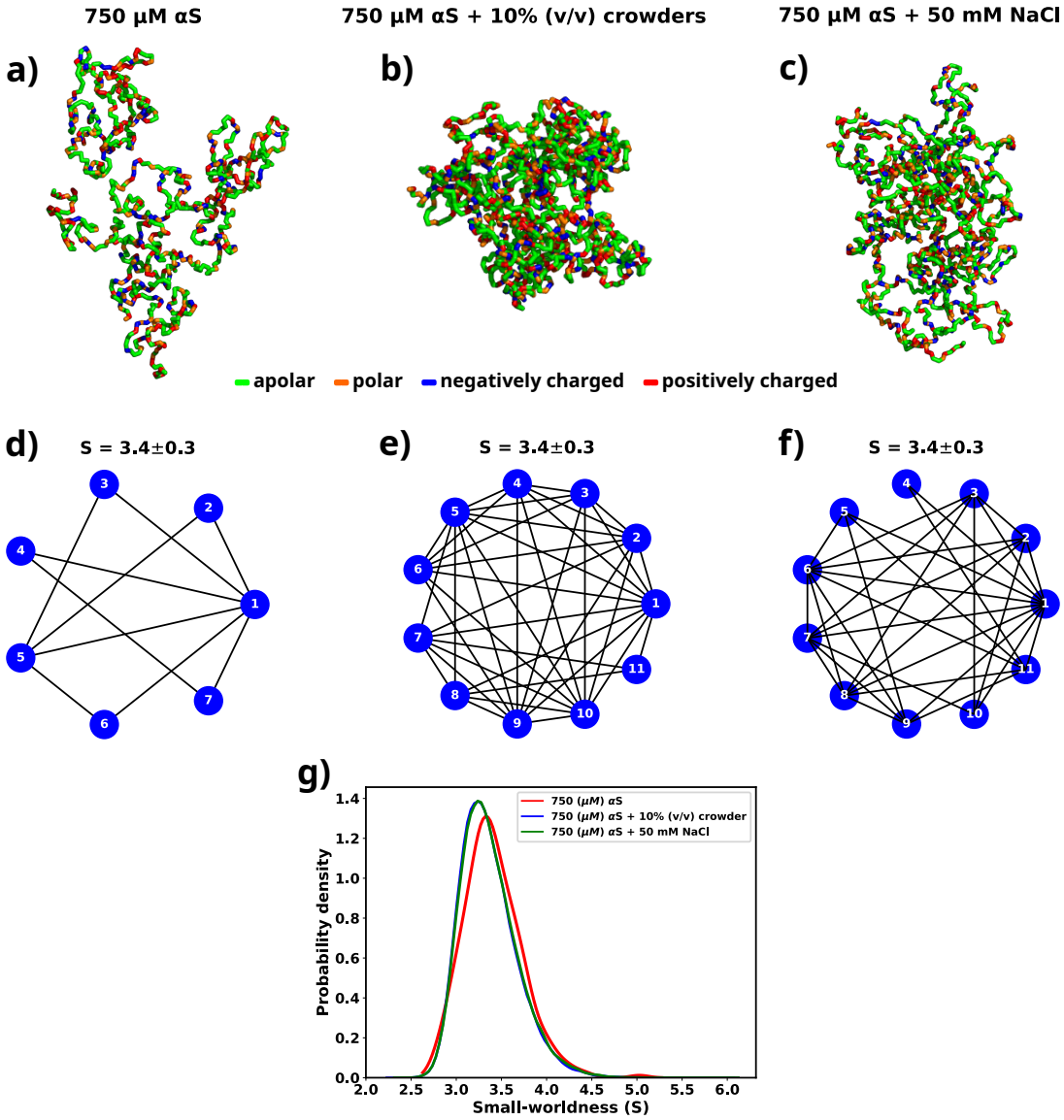

Figure 7: **a)** The largest cluster formed by αS at 750 μM. **b)** The largest cluster formed by αS at 750 μM in the presence of 10% (v/v) crowder. **c)** The largest cluster formed by αS at 750 μM in the presence of 50 mM salt. Different residues have been color coded as per the figure legend. **d)** A graph showing the contacts among different chains constituting the largest cluster formed by αS at 750 μM. **e)** A graph showing the contacts among different chains constituting the largest cluster formed by αS at 750 μM in the presence of 10% (v/v) crowder. **f)** A graph showing the contacts among different chains constituting the largest cluster formed by αS at 750 μM in the presence of 50 mM NaCl. The mean small-worldness (S) of all droplet has been reported above the graph. **g)** Distribution of small-worldness (S) for all scenarios.

A network can be classified as a small-world network by calculating the *Clustering coefficient* and the average shortest path length for the network and comparing those to an equivalent *Erdos-Renyi* network.[40] The *clustering coefficient* (C) is a measure of the "con-

nectedness" of a graph, indicating the extent to which nodes tend to cluster together. It quantifies the likelihood that two nodes with a common neighbor are also connected. On the other hand, the *average shortest path length* (L) is a metric that calculates the average number of steps required to traverse from one node to another within a network. It provides a measure of the efficiency of information or influence propagation across the graph. To estimate the *small-worldness* of a graph, we calculate a parameter (S) defined by Eq-7.

$$S = \frac{\frac{C}{C_r}}{\frac{L}{L_r}} \tag{7}$$

where C and L are the clustering coefficient and average shortest path length for the graph generated for a droplet while C_r and L_r clustering coefficient and average shortest path length for an equivalent Erdos–Renyi network. Small-world networks exhibit the characteristic property of having $C >> C_r$, while $L \approx L_r$. In light of this, for every scenario (solely αS, αS in the presence of crowder, and αS in the presence of salt), we generate an ensemble of graphs that correspond to the droplets formed during the simulation.

For each graph, we calculate the small-worldness coefficient (S)[38] and illustrate the distribution in Figure 7g. We observe a narrow distribution of S with a mean of 3.4 for all cases. In a previous report of RNA-LLPS, a value of $S \approx 4$ was used to classify the droplets small-world networks.[34] Therefore, $S = 3.4$ would suggest that the droplets formed during the simulations are small-world like. Moreover, we observe that the distribution of S is invariant with respect to the environment of the droplet.

Therefore we establish that the modes of interactions, orientations and even connectivities among αS monomers inside a droplet remain same even when their environments are extremely different. We think that this occurs since the residue-level interactions among different monomers inside the droplet are similar irrespective of the environment, as shown in a previous section. This puts forth a very interesting way of viewing αS LLPS. We think that if these residue-level interactions can be disturbed then the stability of the formed droplets might be affected in such a way that they might dissolve spontaneously.

Discussion

We used simulations to investigate the molecular basis of αS monomeric aggregation into soluble oligomers resembling micro-LLPS. The WT protein demonstrated limited aggregation, suggesting a low inherent propensity for LLPS dictated by its primary sequence. IDPs, like αS, often share primary sequence characteristics associated with phase separation. Charged residues distributed with uncharged amino acids, resembling the "sticker and spacer" model, contribute to this molecular grammar. This observation aligns with a general trend in IDPs.[41–43] To assess αS LLPS propensity from its primary sequence, we calculated normalized Shannon entropy (S)[44] (Table S2), Kyte-Doolittle hydrophobicity[45] (Table S3), Normalized, maximum of the sum of PLAAC Log-Likelihood Ratios (NLLR)[46] (Table S4), and LLPS propensity scores obtained from catGranules web-server[47] (Table S5). Comparative analysis with three datasets,[48] namely LLPS+: a dataset of high propensity IDPs whose critical concentrations are 100 μM or below, LLPS- : a dataset of low propensity IDPs whose critical concentrations are greater than 100 μM, and PDB*: a dataset of folded proteins that do not undergo LLPS under normal conditions, revealed αS's distinctive features (Table S6).

We note a significant difference in the Shannon entropy value of αS compared to proteins that do not undergo phase separation, as illustrated in Figure 8a. This deviation suggests a notable inclination of αS to undergo phase separation.[48] Additionally, the hydrophobicity of αS (Figure 8b) is lower than that of the PDB* dataset, aligning more closely with the upper extremes of the LLPS- dataset. This indicates that while αS exhibits a tendency to undergo phase separation, the propensity should be low. Consistent with this, NLLR scores obtained from PLAAC and LLPS propensity scores (Figures 8c and d) reinforce this observation. These collective comparisons, coupled with simulations and experimental data on its critical concentration,[6] conclusively establish that αS does not possess a high LLPS-forming propensity. Instead, this behavior is inherent to its primary structure. In hindsights, this analysis also justifies the requirements of environmental factors for enhancing the proclivity of αS for LLPS, as demonstrated in both our simulations and experimental

findings.[6,13]

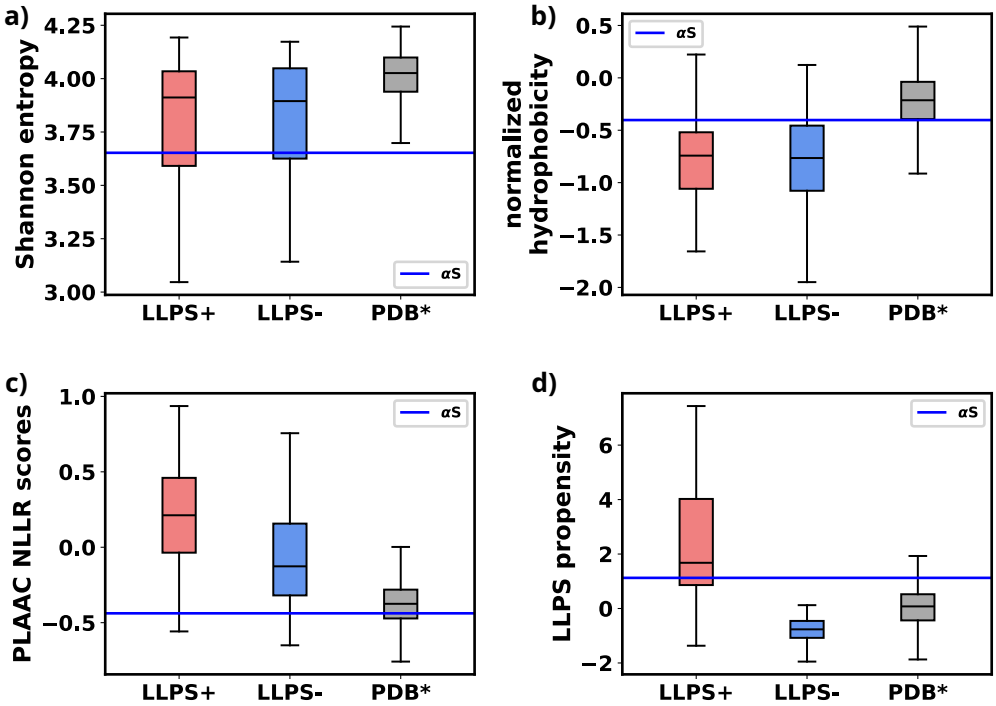

Figure 8: **a)** Comparison of Shannon Entropy of different datasets with αS. **b)** Comparison of Kyte-Doolittle hydrophobicity of different datasets with αS. **c)** Comparison of LLR scores, obtained from PLAAC, of different datasets with αS. **d)** Comparison of LLPS propensity scores, obtained from catGRANULE websever, of different datasets with αS. The values have been summarized in Table S6.

For characterizing αS's aggregation phenomena, we calculated droplet surface tension under varied conditions. We observed that crowders minimally impacted surface tension, while salt increased it; however, both scenarios decreased the relative free energy of the system. Crowders achieved this via entropic means, whereas salt employed enthalpic means. Residue-residue interactions during droplet formation were consistent across environments, with crowder or salt enhancing these interactions. The aggregation pathway involved overall inter-chain interaction enhancement, specifically reducing intra-chain Nter-Cter and Nter-NAC interactions, leading to more extended protein conformations in droplets. Droplet proteins displayed consistent orientation and "small-worldness", a measure of inter-chain connectivity, remained consistent across diverse conditions. Thus, αS aggregates appeared invariant regarding their initial environment in terms of interactions and contacts.

Our study's precision was notably influenced by the careful selection of a simulation force-field. Despite the availability of modern force-fields optimized for multi-chain simulations of IDPs,[49–52] we opted for Martini 3, an explicit water model, due to its emphasis on water's role in aggregation and LLPS, as recently demonstrated in FUS LLPS.[53] Although newer models operate at a faster pace, Martini 3's inclusion of explicit water enhances result accuracy. Additionally, Martini 3 provides a detailed amino acid description and allowing for encoding of protein secondary structures, unlike some newer models that represent amino acids as single beads. Our meticulous choice of the simulation model, combined with a comprehensive analysis, contributes to the accuracy and novelty of this study.

Recent studies have explored the aggregation and LLPS of biopolymers and polyelectrolytes in the presence of membranes, opening a promising avenue for αS research.[54–56] Given that under physiological conditions, αS assumes an oligomeric, membrane-bound form, investigating its interactions with membranes could hold therapeutic potential.[57] Although we exclusively focused on wild-type αS, familial mutations have been reported to exhibit a significantly higher propensity for aggregation.[6] These mutations, involving minor alterations in the primary sequence, highlight the importance of understanding the molecular basis of this distinctive phenotype. Additionally, the observed stability of pre-formed αS droplets[58] poses a challenge in treating Parkinson's Disease (PD). Reversing aggregation/LLPS and understanding associated pathways and mechanisms are crucial. Our study identifies key residues crucial for stable droplet formation, consistent across various environmental conditions.

Methods

Optimizing Martini 3 parameters for αS

Martini 3[25] was trained using DES-Amber[59] that is an atomistic forcefield tuned for single-domain and multi-domain proteins. Therefore, the default parameters of the coarse-grained model is not suited for simulations of disordered proteins and reported to underestimate the

global dimensions of these systems in addition to overestimating protein-protein interactions. Previous attempts to simulate IDPs have modified the Martini force field by tuning the water-protein interactions, specifically, σ and ϵ of Lennard-Jones interactions to render them suitable for modeling a specific IDP or all IDPs.[33,60,61] Here, we follow a similar protocol, however instead of tuning only the ϵ part of the water-protein Lennard-Jones interactions, we refine both the σ and ϵ parameters of the water-protein interactions (Eq-8).

$$V^{'}(r) = 4\epsilon^{'}\left[(\frac{\sigma^{'}}{r})^{12} - (\frac{\sigma^{'}}{r})^{6}\right] \tag{8}$$

where $\epsilon' = \lambda\epsilon$, $\sigma' = \lambda\sigma$ and λ is the scaling parameter that needs to be optimized. Scaling σ tunes the relative radius of the hydration spheres of each residue of a protein while a change in ϵ changes the strength of the water-residue interactions (Figure 9a).

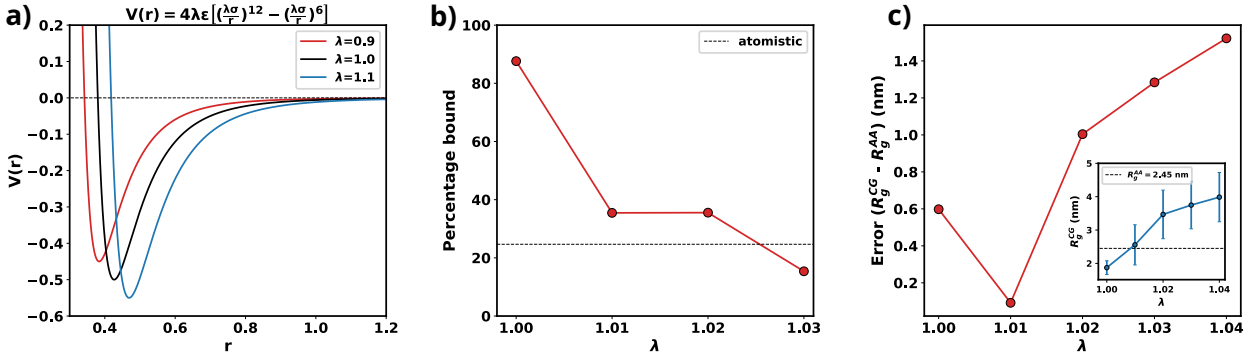

Figure 9: **a)** Plot of LJ potentials with respect to λ. **b)** The percentage bound values between two CG αS chains for different values of λ. The dashed black line represents the percentage bound values for two all-atom chains. **c)** Error between R_g calculated from CG and from all-atom simulations vs λ. The inset plot showcases the average values of R_g obtained from CG along with their respective standard deviations. The dashed line represents the average value from all-atom simulations.

As we are interested in exploiting multi-chain simulations to study the LLPS of αS using Martini 3, we use the percentage of time two all-atom monomers remain bound to each other as the benchmark. To obtain an optimum scaling parameter for the water-protein interactions in Martini 3, specific to αS, we perform CG simulations with two αS chains with different values of λ. We start with two chains, without any secondary structure enforced upon them, randomly placed in a 15.7 nm box making sure that they are apart by at least

0.8 nm which we use as cutoff to classify the chains to be bound. If the minimum distance between any two residues belonging to the different chains are closer than 0.8 nm we consider them to be bound. Using the cutoff defined, we calculate the percentage bound between the two αS monomers for different values of λ in the coarse-grained model. We also calculate the same from atomistic simulations reported in[23] as the reference. From Figure 9b, we can see that for multiple values of λ, we observe a close agreement in percentage bound values between coarse-grained and atomistic simulations.

We conducted additional single-chain coarse-grained (CG) simulations of αS, varying the parameter λ, while refraining from imposing any secondary structure constraints. Subsequently, we compared the mean R_g values derived from these CG simulations with the 73 μs all-atom (AA) trajectory, which replaced the previously published 30 μs all-atom trajectory in[20] and was provided by DE Shaw Research. Figure 9c illustrates that, for $\lambda = 1.01$, the average R_g in the CG simulations closely matches the R_g values obtained from the all-atom data. Consequently, we have chosen $\lambda = 1.01$ for the multi-chain simulations, as it minimizes errors for both single-chain R_g and the observed percentage of time bound in the two-protein chain simulations.

Initial conformation generation for large-scale multi-chain simulations

A recent study used Markov State models to delineate the metastable states based on the extent of compaction (R_g) and identified 3 macrostates and their relative populations.[23] Therefore, in the multi-chain simulations, we maintain similar relative populations of these macrostates. (Figure S7) . The reported percentages of macrostates (labeled as ms1, ms2 and ms3) are 0.06%, 85.9% and 14%, respectively. We added 50 αS monomers consisting of 1 chain of ms1, 45 chains of ms2 and 4 chains of ms3 in a cubic box with their respective secondary structures, determined via DSSP,[62–64] enforced using Martini 3. The size of the box of side a is determined as per Eq-9.

$$a = \sqrt[3]{\frac{N}{N_A \times C}} \tag{9}$$

where N is the number of monomers, N_A is the Avogadro's number and C is the required concentration of αS.

Here, we simulate multiple concentrations of the protein, namely, 300 μM, 400 μM, 500 μM and 750 μM. 300 and 400 μM are below the critical concentrations required to undergo LLPS.[6] We then solvate the system in coarse-grained water. We setup the 50-chain system to simulate 3 conditions; (a) in pure water, (b) in 50 mM NaCl and (c) in presence of 10 % (v/v) crowders. To study effect of salt, we add the required number of Na^+ and Cl^- ions to attain the desired concentration of 50 mM while also adding a few ions to render the system electrically neutral. In the system with crowders, we first add crowders after solvation by replacing a few solvent molecules with the required number of crowder molecules. We next resolvate the system along with the crowders. Finally we render the system electro-neutral by addition of the required number of Na^+ or Cl^- ions. The details of the simulation setup are provided in Table 1.

Table 1: Details of the systems that were explored

summary	conc. of αS(μM)	box size (nm)	# water	# crowders	# Na^+	# Cl^-
300 μM αS in water	300	65.66	2304122	0	450	0
400 μM αS in water	400	59.69	1729213	0	450	0
500 μM αS in water	500	55.52	1389721	0	450	0
750 μM αS in water	750	48.42	920023	0	450	0
750 μM αS + 10%(v/v) crowder	750	48.42	843011	20128	450	0
750 μM αS + 50 mM NaCl	750	48.42	913357	0	3783	3333

*For all simulations, a total of 50 monomeric protein chains have been used which comprise of $1\times MS1$, $45\times MS2$ and $4\times MS3$.

Simulation setup

Upon successful generation of the initial conformation, we first perform an energy minimization using steepest gradient descent using an energy tolerance of 10 kJ mol^{-1} nm^{-1}. We next

perform NVT simulations at 310.15 K using v-rescale thermostat for 5 ns using 0.01 ps as the time step. It is then followed by NPT simulation at 310.15 K and 1 bar using v-rescale thermostat and Berendsen barostat for 5 ns with a time step of 0.02 ps.

Next we perform CG MD simulations using velocity-verlet integrator with a time step of 0.02 ps using v-rescale thermostat at 310.15 K and berendsen barostat at 1 bar. Both Lennard-Jones and electrostatic interactions are cutoff at 1.1 nm. Coulombic interactions are calculated using Reaction-Field algorithm and relative dielectric constant of 15. We perform CG MD for at least 2.5 μs for the systems with 50 αS monomers. The details of the simulation run-times have been provided in Table S1. We use the last 1 μs for further analyses.

List of Software

We have used only open-source software for this study. All simulations have been performed using GROMACS-2021.[65,66] Snapshots were generated using PyMOL 2.5.4.[67–69] Analysis were performed using Python[70] and MDAnalysis.[71,72] Figures were prepared using Matplotlib,[73] Jupyter[74] and Inkscape.[75]

References

(1) Monika Fuxreiter and Peter Tompa. Fuzzy complexes: a more stochastic view of protein function. *Fuzziness: Structural disorder in protein complexes*, pages 1–14, 2012. 2

(2) Julie D Forman-Kay and Tanja Mittag. From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure*, 21(9):1492–1499, 2013.

(3) Alaji Bah and Julie D Forman-Kay. Modulation of intrinsically disordered protein function by post-translational modifications. *Journal of Biological Chemistry*, 291(13):6696–6705, 2016. 2

(4) Vladimir N Uversky, Christopher J Oldfield, and A Keith Dunker. Intrinsically disordered proteins in human diseases: introducing the d2 concept. *Annu. Rev. Biophys.*, 37:215–246, 2008. 2

(5) Yanting Xing, Aparna Nandakumar, Aleksandr Kakinen, Yunxiang Sun, Thomas P Davis, Pu Chun Ke, and Feng Ding. Amyloid aggregation under the lens of liquid–liquid phase separation. *The journal of physical chemistry letters*, 12(1):368–378, 2020. 2

(6) Soumik Ray, Nitu Singh, Rakesh Kumar, Komal Patel, Satyaprakash Pandey, Debalina Datta, Jaladhar Mahato, Rajlaxmi Panigrahi, Ambuja Navalkar, Surabhi Mehra, et al. α-synuclein aggregation nucleates through liquid–liquid phase separation. *Nature chemistry*, 12(8):705–716, 2020. 2, 5, 6, 10, 20, 21, 22, 25

(7) Tong Shu, Jonathan Alberts, Dajun Sang, Sarah Tsao, Matheus Palhares Viana, Meta Heidenreich, Emmanuel Levy, and Liam J Holt. Liquid-liquid phase separation is influenced by the active, crowded cytoplasm. *Biophysical Journal*, 120(3):31a–32a, 2021.

(8) Leandro Cruz Rodríguez, Nahuel N Foressi, and M Soledad Celej. Modulation of α-synuclein phase separation by biomolecules. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1871(2):140885, 2023.

(9) Xinrui Gui, Shuang Feng, Zilong Li, Yanyan Li, Bernd Reif, Bingyang Shi, and Zheng Niu. Liquid-liquid phase separation of amyloid-β oligomers modulates amyloid fibrils formation. *Journal of Biological Chemistry*, page 102926, 2023. 2

(10) Anthony A Hyman, Christoph A Weber, and Frank Jülicher. Liquid-liquid phase separation in biology. *Annual review of cell and developmental biology*, 30:39–58, 2014. 2

(11) Salman F Banani, Hyun O Lee, Anthony A Hyman, and Michael K Rosen. Biomolecular

condensates: organizers of cellular biochemistry. *Nature reviews Molecular cell biology*, 18(5):285–298, 2017.

(12) Yongdae Shin and Clifford P Brangwynne. Liquid phase condensation in cell physiology and disease. *Science*, 357(6357):eaaf4382, 2017. 2

(13) Ajay Singh Sawner, Soumik Ray, Preeti Yadav, Semanti Mukherjee, Rajlaxmi Panigrahi, Manisha Poudyal, Komal Patel, Dhiman Ghosh, Eric Kummerant, Ashutosh Kumar, et al. Modulating α-synuclein liquid–liquid phase separation: Published as part of the biochemistry virtual special issue âĂIJprotein condensatesâĂİ. *Biochemistry*, 60(48):3676–3696, 2021. 3, 6, 21

(14) Mihaela M. Apetri, Nakul C. Maiti, Michael G. Zagorski, Paul R. Carey, and Vernon E. Anderson. Secondary structure of α-synuclein oligomers: Characterization by raman and atomic force microscopy. *Journal of Molecular Biology*, 355(1):63–71, 2006. 3

(15) Dong-Pyo Hong, Shubo Han, Anthony L Fink, and Vladimir N Uversky. Characterization of the non-fibrillar α-synuclein oligomers. *Protein and peptide letters*, 18(3):230âĂŤ240, March 2011.

(16) Serene W. Chen, Srdja Drakulic, Emma Deas, Myriam Ouberai, Francesco A. Aprile, RocÃŋo Arranz, Samuel Ness, Cintia Roodveldt, Tim Guilliams, Erwin J. De-Genst, David Klenerman, Nicholas W. Wood, Tuomas P.J. Knowles, Carlos Alfonso, GermÃąn Rivas, Andrey Y. Abramov, JosÃľ MarÃŋa Valpuesta, Christopher M. Dobson, and Nunilo Cremades. Structural characterization of toxic oligomers that are kinetically trapped during α-synuclein fibril formation. *Proceedings of the National Academy of Sciences*, 112(16):E1994–E2003, 2015.

(17) N. Cremades, S.W. Chen, and C.M. Dobson. Chapter three - structural characteristics of α-synuclein oligomers. In Massimo Sandal, editor, *Early Stage Protein Misfolding*

and Amyloid Aggregation, volume 329 of *International Review of Cell and Molecular Biology*, pages 79–143. Academic Press, 2017. 3

(18) Mustapha Carab Ahmed, Line K Skaanning, Alexander Jussupow, Estella A Newcombe, Birthe B Kragelund, Carlo Camilloni, Annette E Langkilde, and Kresten Lindorff-Larsen. Refinement of α-synuclein ensembles against saxs data: Comparison of force fields and methods. *Frontiers in molecular biosciences*, 8:654333, 2021. 3

(19) Khandekar Jishan Bari and Dube Dheeraj Prakashchand. Fundamental challenges and outlook in simulating liquid–liquid phase separation of intrinsically disordered proteins. *The Journal of Physical Chemistry Letters*, 12(6):1644–1656, 2021.

(20) Paul Robustelli, Stefano Piana, and David E Shaw. Developing a molecular dynamics force field for both folded and disordered protein states. *Proceedings of the National Academy of Sciences*, 115(21):E4758–E4766, 2018. 24

(21) Robert B Best, Wenwei Zheng, and Jeetain Mittal. Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association. *Journal of chemical theory and computation*, 10(11):5113–5124, 2014.

(22) Jing Huang, Sarah Rauscher, Grzegorz Nawrocki, Ting Ran, Michael Feig, Bert L De Groot, Helmut Grubmüller, and Alexander D MacKerell Jr. Charmm36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods*, 14(1):71–73, 2017.

(23) S. Menon and J. Mondal. Conformational plasticity in α-synuclein and how crowded environment modulates it. *J.Phys.Chem. B*, 127:4032–4049, 2023. 4, 14, 24

(24) Sneha Menon and Jagannath Mondal. Small molecule modulates α-synuclein conformation and its oligomerization via entropy expansion. *bioRxiv*, pages 2022–10, 2022. 3

(25) Paulo CT Souza, Riccardo Alessandri, Jonathan Barnoud, Sebastian Thallmair, Ignacio Faustino, Fabian Grünewald, Ilias Patmanidis, Haleh Abdizadeh, Bart MH Bruininks, Tsjerk A Wassenaar, et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature methods*, 18(4):382–388, 2021. 3, 4, 22

(26) R John Ellis and Allen P Minton. Protein aggregation in crowded environments. 2006. 6

(27) Eric J Deeds, Orr Ashenberg, Jaline Gerardin, and Eugene I Shakhnovich. Robust protein–protein interactions in crowded cellular environments. *Proceedings of the National Academy of Sciences*, 104(38):14952–14957, 2007.

(28) Edwin Li, Jesse Placone, Mikhail Merzlyakov, and Kalina Hristova. Quantitative measurements of protein interactions in a crowded cellular environment. *Analytical chemistry*, 80(15):5976–5985, 2008.

(29) Huan-Xiang Zhou. Influence of crowded cellular environments on protein folding, binding, and oligomerization: biological consequences and potentials of atomistic modeling. *FEBS letters*, 587(8):1053–1061, 2013. 6

(30) Vladimir N Uversky, Elisa M Cooper, Kiowa S Bower, Jie Li, and Anthony L Fink. Accelerated α-synuclein fibrillation in crowded milieu. *FEBS letters*, 515(1-3):99–103, 2002. 6

(31) Larissa A Munishkina, Atta Ahmad, Anthony L Fink, and Vladimir N Uversky. Guiding protein aggregation with macromolecular crowding. *Biochemistry*, 47(34):8993–9006, 2008.

(32) Istvan Horvath, Ranjeet Kumar, and Pernilla Wittung-Stafshede. Macromolecular crowding modulates α-synuclein amyloid fiber growth. *Biophysical Journal*, 120(16):3374–3381, 2021. 6

(33) Zakarya Benayad, Soren von Bulow, Lukas S Stelzl, and Gerhard Hummer. Simulation of fus protein condensates with an adapted coarse-grained model. *Journal of Chemical Theory and Computation*, 17(1):525–537, 2020. 9, 23

(34) Hung T Nguyen, Naoto Hori, and D Thirumalai. Condensates in rna repeat sequences are heterogeneously organized and exhibit reptation dynamics. *Nature Chemistry*, pages 1–11, 2022. 10, 11, 12, 19

(35) Daniele Ubbiali, Marta Fratini, Lolita Piersimoni, Christian H Ihling, Marc Kipping, Ingo Heilmann, Claudio Iacobucci, and Andrea Sinz. Direct observation of âĂIJelongatedâĂİ conformational states in α-synuclein upon liquid-liquid phase separation. *Angewandte Chemie*, 134(46):e202205726, 2022. 12, 15

(36) Duncan J Watts and Steven H Strogatz. Collective dynamics of âĂŸsmall-worldâĂŹnetworks. *nature*, 393(6684):440–442, 1998. 17

(37) Alain Barrat and Martin Weigt. On the properties of small-world network models. *The European Physical Journal B-Condensed Matter and Complex Systems*, 13:547–560, 2000.

(38) Mark D Humphries and Kevin Gurney. Network âĂŸsmall-world-nessâĂŹ: a quantitative method for determining canonical network equivalence. *PloS one*, 3(4):e0002051, 2008. 19

(39) Mina Farag, Samuel R Cohen, Wade M Borcherds, Anne Bremer, Tanja Mittag, and Rohit V Pappu. Condensates formed by prion-like low-complexity domains have small-world network structures and interfaces defined by expanded conformations. *Nature communications*, 13(1):7722, 2022. 17

(40) Erdos RENYI. On random graph. *Publicationes Mathematicate*, 6:290–297, 1959. 18

(41) Jeong-Mo Choi, Furqan Dar, and Rohit V Pappu. Lassi: A lattice model for simulating phase transitions of multivalent proteins. *PLoS computational biology*, 15(10):e1007028, 2019. 20

(42) Erik W Martin, Alex S Holehouse, Ivan Peran, Mina Farag, J Jeremias Incicco, Anne Bremer, Christy R Grace, Andrea Soranno, Rohit V Pappu, and Tanja Mittag. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science*, 367(6478):694–699, 2020.

(43) Jeong-Mo Choi, Alex S Holehouse, and Rohit V Pappu. Physical principles underlying the complex biology of intracellular phase transitions. *Annual review of biophysics*, 49:107–133, 2020. 20

(44) Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948. 20

(45) Jack Kyte and Russell F Doolittle. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology*, 157(1):105–132, 1982. 20

(46) Alex K Lancaster, Andrew Nutter-Upham, Susan Lindquist, and Oliver D King. Plaac: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics*, 30(17):2501–2502, 2014. 20

(47) Benedetta Bolognesi, Nieves Lorenzo Gotor, Riddhiman Dhar, Davide Cirillo, Marta Baldrighi, Gian Gaetano Tartaglia, and Ben Lehner. A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression. *Cell reports*, 16(1):222–231, 2016. 20

(48) Kadi L Saar, Alexey S Morgunov, Runzhang Qi, William E Arter, Georg Krainer, Alpha A Lee, and Tuomas PJ Knowles. Learning the molecular grammar of protein condensates from sequence determinants and embeddings. *Proceedings of the National Academy of Sciences*, 118(15):e2019053118, 2021. 20

(49) Gregory L Dignon, Wenwei Zheng, Young C Kim, Robert B Best, and Jeetain Mittal. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS computational biology*, 14(1):e1005941, 2018. 22

(50) Andrew P Latham and Bin Zhang. Maximum entropy optimized force field for intrinsically disordered proteins. *Journal of chemical theory and computation*, 16(1):773–781, 2019.

(51) Roshan Mammen Regy, Jacob Thompson, Young C Kim, and Jeetain Mittal. Improved coarse-grained model for studying sequence dependent phase separation of disordered proteins. *Protein Science*, 30(7):1371–1379, 2021.

(52) Yumeng Zhang, Xiaorong Liu, and Jianhan Chen. Toward accurate coarse-grained simulations of disordered proteins and their dynamic interactions. *Journal of Chemical Information and Modeling*, 62(18):4523–4536, 2022. 22

(53) Saumyak Mukherjee and Lars V Schäfer. Thermodynamic forces from protein and water govern condensate formation of an intrinsically disordered protein domain. *Nature Communications*, 14(1):5892, 2023. 22

(54) Sayantan Mondal and Qiang Cui. Coacervation of poly-electrolytes in the presence of lipid bilayers: mutual alteration of structure and morphology. *Chemical Science*, 13(26):7933–7946, 2022. 22

(55) Zhuang Liu, Arun Yethiraj, and Qiang Cui. Selective and sensitive surface condensation driven by coupled phase behaviors of membrane and biopolymers. *Biophysical Journal*, 122(3):207a–208a, 2023.

(56) Zhuang Liu, Arun Yethiraj, and Qiang Cui. Sensitive and selective polymer condensation at membrane surface driven by positive co-operativity. *Proceedings of the National Academy of Sciences*, 120(15):e2212516120, 2023. 22

(57) André Pineda and Jacqueline Burré. Modulating membrane binding of α-synuclein as a therapeutic strategy. *Proceedings of the National Academy of Sciences*, 114(6):1223–1225, 2017. 22

(58) Vladimir N Uversky, Jie Li, and Anthony L Fink. Evidence for a partially folded intermediate in α-synuclein fibril formation. *Journal of Biological Chemistry*, 276(14):10737–10744, 2001. 22

(59) Stefano Piana, Paul Robustelli, Dazhi Tan, Songela Chen, and David E Shaw. Development of a force field for the simulation of single-chain proteins and protein–protein complexes. *Journal of chemical theory and computation*, 16(4):2494–2507, 2020. 22

(60) F Emil Thomasen, Francesco Pesce, Mette Ahrensback Roesgaard, Giulio Tesei, and Kresten Lindorff-Larsen. Improving martini 3 for disordered and multidomain proteins. *Journal of Chemical Theory and Computation*, 18(4):2033–2041, 2022. 23

(61) Gul H Zerze. Optimizing the martini 3 force field reveals the effects of the intricate balance between protein–water interaction strength and salt concentration on biomolecular condensate formation. *Journal of Chemical Theory and Computation*, 2023. 23

(62) Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers: Original Research on Biomolecules*, 22(12):2577–2637, 1983. 24

(63) Robbie P Joosten, Tim AH Te Beek, Elmar Krieger, Maarten L Hekkelman, Rob WW Hooft, Reinhard Schneider, Chris Sander, and Gert Vriend. A series of pdb related databases for everyday needs. *Nucleic acids research*, 39(suppl_1):D411–D419, 2010.

(64) Wouter G Touw, Coos Baakman, Jon Black, Tim AH Te Beek, Elmar Krieger, Robbie P Joosten, and Gert Vriend. A series of pdb-related databanks for everyday needs. *Nucleic acids research*, 43(D1):D364–D368, 2015. 24

(65) David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E Mark, and Herman JC Berendsen. Gromacs: fast, flexible, and free. *Journal of computational chemistry*, 26(16):1701–1718, 2005. 26

(66) Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015. 26

(67) Schrödinger, LLC. The AxPyMOL molecular graphics plugin for Microsoft PowerPoint, version 1.8. November 2015. 26

(68) Schrödinger, LLC. The JyMOL molecular graphics development component, version 1.8. November 2015.

(69) Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. November 2015. 26

(70) Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009. 26

(71) Naveen Michaud-Agrawal, Elizabeth J Denning, Thomas B Woolf, and Oliver Beckstein. Mdanalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal of computational chemistry*, 32(10):2319–2327, 2011. 26

(72) Richard J Gowers, Max Linke, Jonathan Barnoud, Tyler JE Reddy, Manuel N Melo, Sean L Seyler, Jan Domanski, David L Dotson, Sébastien Buchoux, Ian M Kenney, et al. Mdanalysis: a python package for the rapid analysis of molecular dynamics simulations. In *Proceedings of the 15th python in science conference*, volume 98, page 105. SciPy Austin, TX, 2016. 26

(73) J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. 26

(74) Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica Hamrick, Jason Grout, Sylvain Corlay, Paul Ivanov, Damián Avila, Safia Abdalla, and Carol Willing. Jupyter notebooks – a publishing format for reproducible computational workflows. In F. Loizides and B. Schmidt, editors, *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pages 87 – 90. IOS Press, 2016. 26

(75) Inkscape Project. Inkscape. 26