Role of non-specific interactions in the phase-separation and maturation of macromolecules

Rakesh Krishnan,[†] Srivastav Ranganathan,[*,‡] Samir K Maji,[†] and Ranjith Padinhateeri[*,†]

†*Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai 400076*

‡*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138*

E-mail: srivastav_ranganathan@fas.harvard.edu; ranjithp@iitb.ac.in

Abstract

Phase separation of biomolecules could be mediated by both specific and non-specific interactions. How the interplay between non-specific and specific interactions along with polymer entropy influences phase separation is an open question. We address this question by simulating self-associating molecules as polymer chains with a short core stretch that forms the specifically interacting functional interface, and longer non-core regions that participate in non-specific/promiscuous interactions. Our results show that the interplay of specific (strength, $\epsilon_{\rm sp}$) and non-specific interactions (strength, $\epsilon_{\rm ns}$) could result in phase separation of polymers and its transition to solid-like aggregates (mature state). In the absence of $\epsilon_{\rm ns}$, the polymer chains do not dwell long enough in the vicinity of each other to undergo phase separation and transition

into a mature state. On the other hand, in the limit of strong ϵ_{ns}, the assemblies cannot transition into the mature state because they get kinetically trapped, suggesting an optimal range of interactions favoring mature multimers. In the scenario where only a fraction (N_{frac}) of the non-core regions participate in attractive interactions, we find that a slight modifications to either ϵ_{ns} or N_{frac} can result in a dramatically altered self-assembled states. Using a combination of heterogeneous and homogeneous mix of polymers, we establish how this interplay between interaction energies dictates the propensity of biomolecules to find the correct binding partner at dilute concentrations in crowded environments.

Introduction

Liquid-liquid phase separation plays a key role in regulating Spatio-temporal organization of biomolecules in the cell. Proteins and/or RNA are known to phase-separate into polymer-dense phases known as membraneless organelles (MLOs), often correlated to external stresses or a certain phase of cell division.[1,2] Interestingly, the local concentration of proteins within these compartments is 50-100 fold higher than their bulk concentrations.[3] Given that the bulk concentration of phase-separating components could be in the low nanomolar range, this gives rise to an interesting hypothesis – could liquid-like phase-separated compartments facilitate biochemical reactions within this 'polymer-rich' phase by increasing the likelihood of key functional interactions? This gains importance in the context of two interaction partners ($\approx$ nM concentrations) meeting in space and locking into place to form functional complexes; this is a non-trivial problem. Not only does the low bulk concentration of interacting components make their diffusing-limited meeting unlikely, these dynamic molecules must then undergo further structural reorganization to enable key 'native' interactions for their functionality.

Phase-separated MLOs such as P-bodies and stress granules are known to be stabilized by a network of interactions between specifically interacting domains which bind to unique interaction partners. Interestingly, in addition to these specifically interacting domains,

protein sequences also bear regions with no known functional interactions. Protein-protein interaction studies suggest that large portions of the interactome are made up of promiscuous or "noisy" interactions with no known function.[4–8] Therefore, could non-specific interactions play a role in assembly of liquid-like polymer-dense phases en route to formation of specific contacts? Is there a regime of interaction strengths that results in non-specifically driven phase-separation leading to formation of specific contacts?

Coarse-grained computational models have been extensively employed to study various aspects of biomolecular self-assembly and phase-separation. Typical approaches in coarse-grained simulations include the modeling of proteins as lattice polymers,[9–11] patchy particles[12,13] and spheres/beads.[14] Deeds et al. simulated lattice polymers and suggested the existence of a narrow range of temperatures, where the formation of specific complexes is favored despite the presence of promiscuous interactions.[9] Saric et al. employed a patchy sphero-cylindrical representation of proteins to establish the key role of non-specific interactions in promoting the nucleation of amyloid fibrils.[13] Osmanovic and Rabin employed phenomenological Monte Carlo simulations to demonstrate how non-specific interactions could assist the formation of specific complexes using a group of interacting particles.[14] Another class of computational models is the sticker-spacer model, where the stickers are regions with high attractive interactions and spacer residues, on the other hand, are beads that have excluded volume and weak attractive interactions with other spacers and stickers without any attraction preferences.[15,16] This model has been effectively employed to study phase separating proteins, and how valency of stickers (specific interactions) affect the phase separation and aggregation.[15] Most of these existing theoretical models, however, focus on the role of various interaction parameters on the equilibrium state of the interacting biopolymers.[9,17–20] In contrast, the process of assembly and the role played by non-specific interactions in the formation of specific polymeric complexes is less clearly understood.

In this work, we study the role of non-specific interactions in driving the phase separation of polymers enabling the specifically interacting domains to form contacts in the

polymer-dense phase. Using a coarse-grained representation of proteins, we explore various parameters – specific interaction strengths, non-specific interaction strengths, a fraction of non-specifically interacting residues and systems of different polymer types to understand their role in phase separation of biological macromolecules. The interacting proteins were modeled as semi-flexible polymer chains consisting of two domains, a small central patch involved in functional interactions (specific interactions) and other regions which participate in weak, promiscuous interactions (non-specific interactions). The results from off-lattice Langevin dynamics simulations show that non-specific assemblies could be an essential precursor to the formation of functional complexes. We systematically explore the parameter regimes where these functional domains find each other upon phase-separation. Our results suggest that non-specific interactions are essential in the formation of phase-separated states stabilized by specific interactions. We further tuned the fraction of non-specific interaction and its strength, which resulted in different assembling states. Using an extension of this model, we demonstrate how the transition between non-specific to specific interaction-dominant assemblies is influenced by the presence of non-interacting polymer chains (crowders) and non-specifically interacting polymers in the system. Overall, our results shed light on how interactions involving noisy regions could be tuned to promote the formation of otherwise key, low-probability functional interactions.[11,21,22]

Table 1: Parameters abbreviations in simulations

Parameters	**Abbreviation**
ϵ_{ns}	Non-specific interaction energy (pairwise)
ϵ_{sp}	Specific interaction energy (pairwise)
N_{ns}	Number of non-specific contacts per bead
N_{sp}	Number of specific contacts per bead
N_{frac}	Fraction of non-specifically interacting residues (given in percentage)
R_g^{red}	Radius of gyration of red beads
N^{lc}	Number of polymers in largest cluster.
R_g^{norm}	Radius of gyration normalised (R_g^{red}/N^{lc})

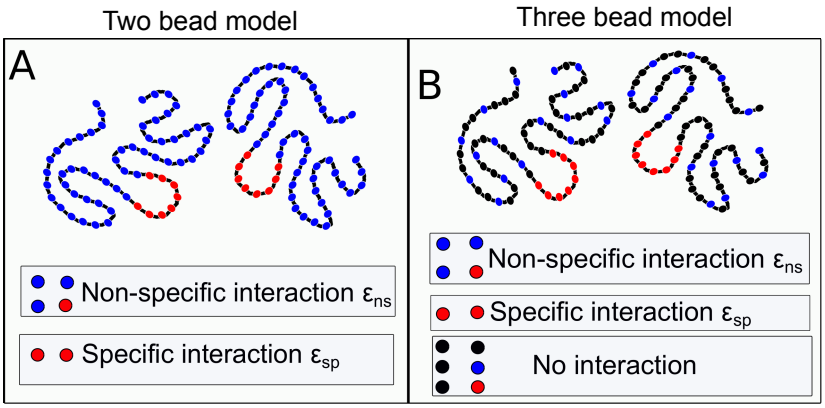

Figure 1: Bead-spring polymer model of protein. A) Chains consist of a central "functional" patch (red) surrounded by other beads which make up non-specifically interacting regions (blue). B) Three bead polymers with black beads representing non-interacting residues, in addition to specifically (red) and non-specifically (blue) interacting residues. Specific and non-specific interaction strengths are represented by $\epsilon_{\rm sp}$ and $\epsilon_{\rm ns}$, respectively, with $\epsilon_{\rm sp} > \epsilon_{\rm ns}$.

Results

As discussed in the model section, we model proteins as bead spring polymer chains, with each bead signifying an effective interaction site per region of one or a few amino acids. We simulate polymer chains that are 80-beads long and each having different types of interactions among the beads as shown in Fig. 1. Apart from the steric hindrance, the beads may have weak non-specific interaction with energy $\epsilon_{\rm ns}$, strong specific interaction with energy $\epsilon_{\rm sp}$ or no interaction as the case may be. We place multiple polymers inside a cubic box at random locations and compute their dynamics using Langevin dynamics and Replica Exchange Langevin Dynamics (RELD) simulations.

Proteins as dynamic polymers: role of non-specific interaction in the formation of functional dimers

Cellular function hinges on biomolecular recognition and the formation of specific protein-protein complexes. Interestingly, interface residues involved in stabilizing biomolecular complexes account for only a tiny fraction of protein surfaces. Here, we first probe how the interplay between specific and non-specific interactions can shape the process of dimeriza-

tion with coarse-grained polymer chains (Fig. 1A). We vary the ϵ_{ns} and ϵ_{sp} systematically and investigate the process of formation of mature dimers – a state where the two central red patches interact with each other. This mature dimeric state is similar to protein-protein complexes stabilized in the correct configurations that enable biomolecular function.

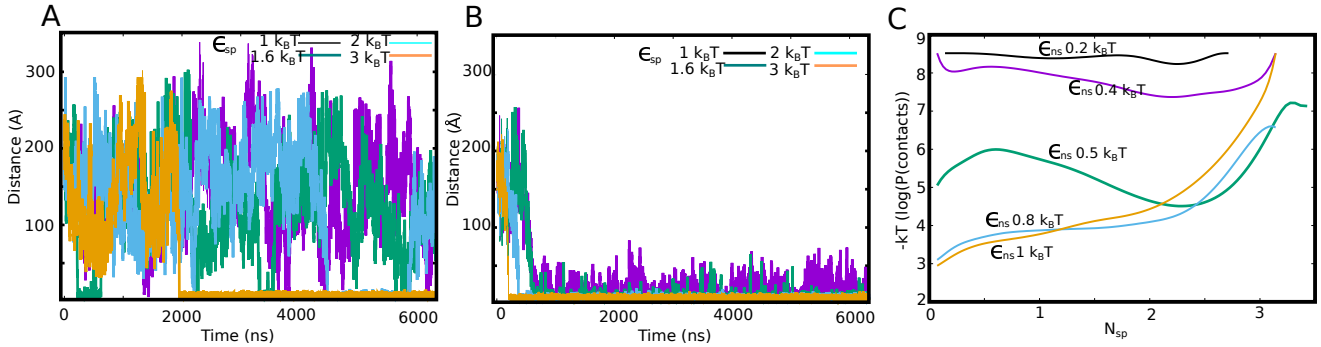

Figure 2: Distance between the center of masses of the two polymeric chains as a function of time. The strength of the ϵ_{ns} in the two panels are A) 0.2 k_BT, B) 0.6 k_BT. C) Free energy as a function of specific contacts per bead computed using replica exchange langevin dynamics simulations. Individual free energy curves represent simulations performed at different values of ϵ_{ns}. For an $\epsilon_{ns} = 0.5k_BT$ (green curve), there is a sharp minima for higher value of specific contacts per bead. For high and low values of ϵ_{ns} this minima is not observed, suggesting that there is an optimal value of ϵ_{ns} at which functional dimers are favored.

In Fig. 2A and Fig. 2B, we plot distance between center of masses of red patches in the two polymers, as a function of time, for varying ϵ_{ns}. At low non-specific interaction strengths ϵ_{ns}=0.2 k_BT in Fig. 2A the two red patches form stable interactions only at extremely large values of specific interaction (ϵ_{sp}=3 k_BT,Orange curve). For values of ϵ_{sp} lower than 3 k_BT, the two 'functional' patches do not meet or stabilize during the timescale of our simulation (6 μs). At lower values of ϵ_{sp}, even when the red patches do come in contact with each other, they do not dwell long enough to make stable contacts as seen green curve of Fig. 2A, same observation is made for ϵ_{ns}=0.4 k_BT (see supplementary Fig. 1A). As we increase the strength of the ϵ_{ns} to 0.6 k_BT (Fig. 2B), the central patches form stable contacts at ϵ_{sp} values as low as 1 k_BT. An increase in ϵ_{ns} from 0.6 k_BT to 1.0 k_BT (supplementary Fig. 1B), results in a slowing down of time it takes for the specific interaction to form — this is due to stronger non-specific interaction (blue-blue) delaying the meeting of red-red patches.

To further understand how ϵ_{ns} could modulate both the equilibrium state as well as the kinetics of this transition, we performed parallel tempering (replica exchange) Langevin dynamics simulations to construct the free energy profiles for the dimeric system. We simulated 24 different replicas with fixed $\epsilon_{sp} = 1.2k_BT$, each spaced 4 K apart for a temperature range of 309-401 K. The free energy curves in Fig. 2C show that for very low values of ϵ_{ns}, we do not observe a clear minima at higher values of N_{sp} (specific contacts per bead). Also, for very large values of ϵ_{ns}, we observe a steep barrier (>3 k_BT) to go from low to high values of specific contacts, suggesting that the system would prefer to remain in a non-specifically bound state. However, at an intermediate value of ϵ_{ns}, we see a defined minima for the higher value of N_{sp}, showing that the system would undergo a transition from the phase-separated non-specific dimeric state to the mature dimeric state in such a scenario. These results suggest that the meeting of functional interfaces (red patch) between two polymer molecules must be facilitated by weak ϵ_{ns} that increases the dwelling time of a non-specifically bound dimer. The individual chains within the dimer can then undergo rearrangement, eventually resulting in mature dimers with the central specifically interacting patches in contact with each other.

Functional biomolecular complexes: non-specific interaction, phase separation and maturation

Phase-separated biological assemblies such as membrane-less organelles can often localize hundreds of molecules within them .[23–27] The concentrations of proteins within these clusters are several-fold higher than their bulk concentrations, facilitating favorable biomolecular interactions within the dense phase. Also, promiscuous interactions have been attributed to stabilizing the phase-separated state while allowing the specific interactions to form within the polymer-rich phase. In this context, we simulated 30 polymers chains, each having two patches with the same behavior as described so far (see Fig. 1B), in a box such that the concentration is $\approx 200\mu$M. We systematically probe the role of the non-functional regions

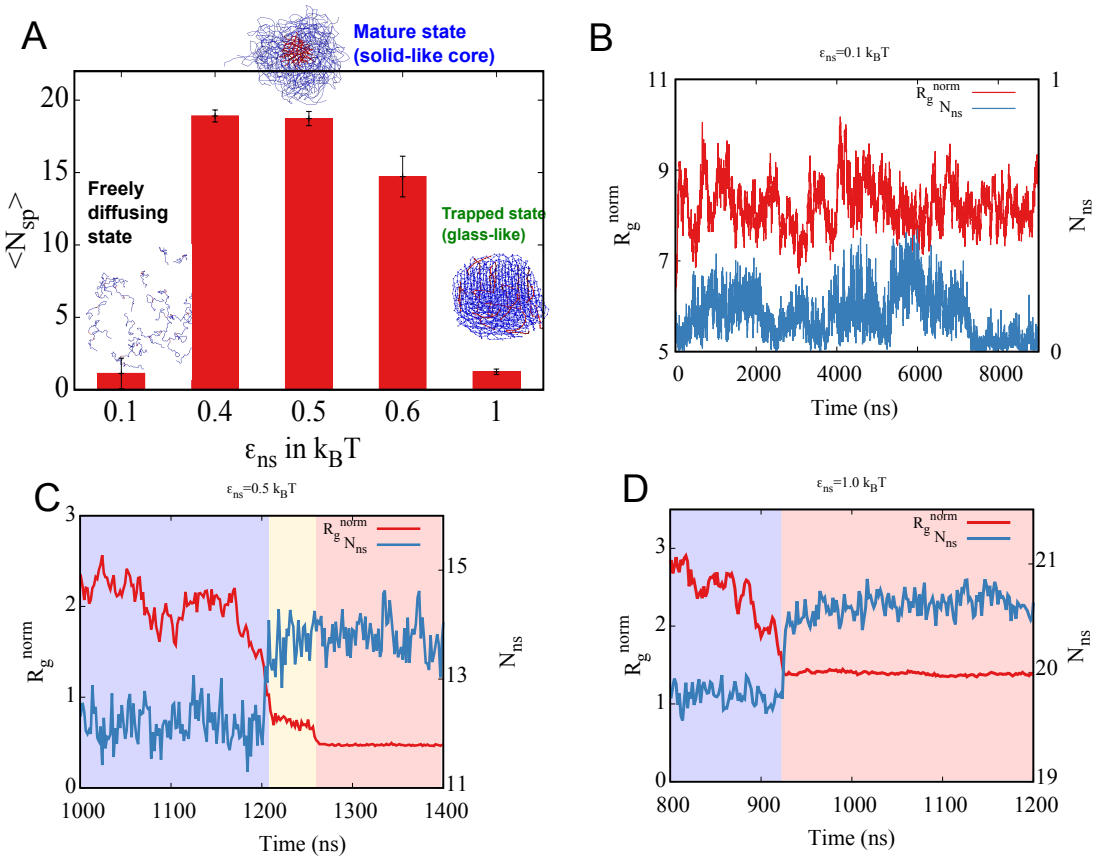

Figure 3: A) Average number of specific interaction per bead (red-red) N_{sp} for different of non-specific interaction strength ϵ_{ns} values. (B,C,D) Radius of gyration (R_g^{norm}) and Number of non-specific interactions per bead (N_{ns}) are plotted as a function of time for $\epsilon_{sp} = 1.2$ k_BT. B) For weak $\epsilon_{ns} = 0.1$ k_BT we observe neglibile non-specific contacts in the system ($N_{ns} \approx 0$). This is accompanied by large values of R_g^{norm} suggesting that the stable multimers are not observed. C)For $\epsilon_{ns} = 0.5$ k_BT, there is an initial regime of low N_{ns} and large R_g^{norm} while the polymer chains are yet to form multimeric clusters. This is followed by a regime (shaded yellow) where there is an increase in N_{ns} suggesting the formation of multimers stabilized by non-specific interactions. This increase in N_{ns} further enables the central functional patches to come in contact, indicated by the further decrease in R_g^{norm} (red regime). A further increase in ϵ_{ns} to 1 kT results in the initial formation of non-specifically stabilized multimers without a further reorganization into a functional complex where the central red-patches do not phase-separate into a separate region. At these high values of ϵ_{ns}, the multimers are trapped in a glass-like state.

of the polymer chain in driving phase separation that enables the meeting of functional patches within the polymer-dense phase. For ϵ_{sp} of 1.2 k_BT, there is negligible development of contacts between specific patches, at low values of ϵ_{ns} of 0.1 k_BT (Fig. 3A). As we increase the ϵ_{ns}= 0.1 k_BT to 0.4/0.5/0.6 k_BT, we see a significant increase in the N_{sp} resulting in

a mature complex, where all red patches come together. However, as the strength of the non-specific interactions approaches that of the specific interactions (ϵ_{sp} = 1.2 k_BT and ϵ_{ns} = 1 k_BT), there is a sudden decrease in the number of specific contacts suggesting a slow down in the search process of the specific interfaces. This suggests that despite the relatively higher strength of the interaction between the red patches, the ability to phase separate and form a mature multimer (solid-like core, see below) is modulated by the 'non-specific' regions of the polymer chain.

For a better understanding of the system, typical snapshots of polymer configurations from the simulation are shown near the graph (Fig. 3A). For low ϵ_{ns} there are only a few N_{sp} (ϵ_{ns} = 0.1 k_BT) and the corresponding polymer configuration shows that the polymers are scattered without much interaction as freely diffusing state. When the ϵ_{ns} is increased to 0.4/0.5/0.6 k_BT, all red patches from different polymers come together and form a mature complex with a solid-like core. The corresponding snapshot of the polymer configuration shows a central core of red beads forming a mature state with a solid-like core of red beads. It is interesting to note that the non-specific interaction leads to a phase separation-like phenomena further transitioned to a mature state. When the value of ϵ_{ns} was increased further to 1.0 k_BT, we find decrease in the N_{sp} . The corresponding polymer configuration shows that even though the polymers are nearby, they are stuck with each other and have much less mobility; the state is like a kinetically trapped polymer in a "glass-like" state.

To further understand the kinetics of red patches finding each other with the help of non-specific blue beads, we tracked two parameters as a function of time: N_{ns} (Number of non-specific contacts per bead) and R_g^{norm}(Radius of gyration of red beads, R_g^{norm} =R_g^{red}/N_{lc}) as in Fig. 3B,C&D. For very weak non-specific interactions (0.1 k_BT, Fig. 3B) there is no development of non-specific contacts (blue curve) as represented in the y_2 axis (the Y-axis on the right side). At the same time, the R_g^{norm} is very large (red curve, marked on the y_1 axis—the major Y-axis on the left side), suggesting that the red beads are far away and are not in contact. However, when ϵ_{ns} = 0.5 k_BT, the non-specific contacts per bead increases

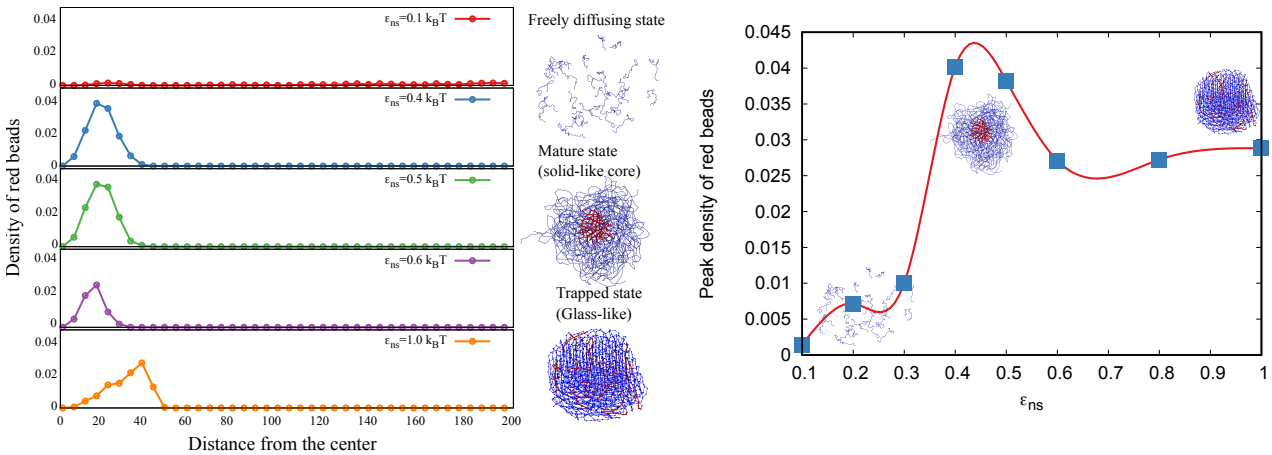

Figure 4: Density profiles. A) Density of red beads within a spherical shell with increasing radial distance from the center of a cluster. Individual panels represent density profiles for increasing values of ϵ_{ns} (top to bottom). B) Peak density of red beads within the single largest cluster shows an initial increase with ϵ_{ns} followed by a decrease at large values of ϵ_{ns}.

substantially Fig. 3C (blue curve) as represented in the y_2 axis. At the same time, the red-beads start to come together and $\mathrm{R_g^{norm}}$ decreases as in the blue region (Fig. 3C red curve represented on the y_1 axis). Even after the $\mathrm{N_{ns}}$ saturate, the $\mathrm{R_g^{norm}}$ keeps decreasing as in the yellow region; shows the kinetics of maturation to form a mature state with solid-like core from the non-specific phase-separated state. In Fig. 3D, where ϵ_{ns} =1.0 $\mathrm{k_BT}$, $\mathrm{N_{ns}}$ increase and $\mathrm{R_g^{norm}}$ decrease initially to form multimeric interaction as in blue region, then both the values saturate showing the formation of kinetically trapped polymers in a glass-like state (red region).

For different values of ϵ_{ns}, there are different plausible states for polymers to form. It can be in three different states, basically freely diffusing state, mature state with a solid-like core of red beads in the center, and glass-like trapped state, where the polymers form a multimer without red beads in the center. To differentiate these states of multimers formed, we plotted the number density of red beads from the center of the cluster towards its periphery (along radial direction) for various ϵ_{ns} in Fig. 4A and the peak density of the red beads in Fig. 4B. For ϵ_{ns} =0.1 $\mathrm{k_BT}$, the density is very low, suggesting polymers in a freely diffusing state. When ϵ_{ns} =0.5 $\mathrm{k_BT}$, the density is at its peak (as in Fig. 4B), indicating the red beads

clustered at the center forming mature state with the solid-like core. For ϵ_{ns} =1 k_BT the density of the red beads is spread across the cluster, showing that the state in which polymers are kinetically trapped in a glass-like state (Fig. 4A).

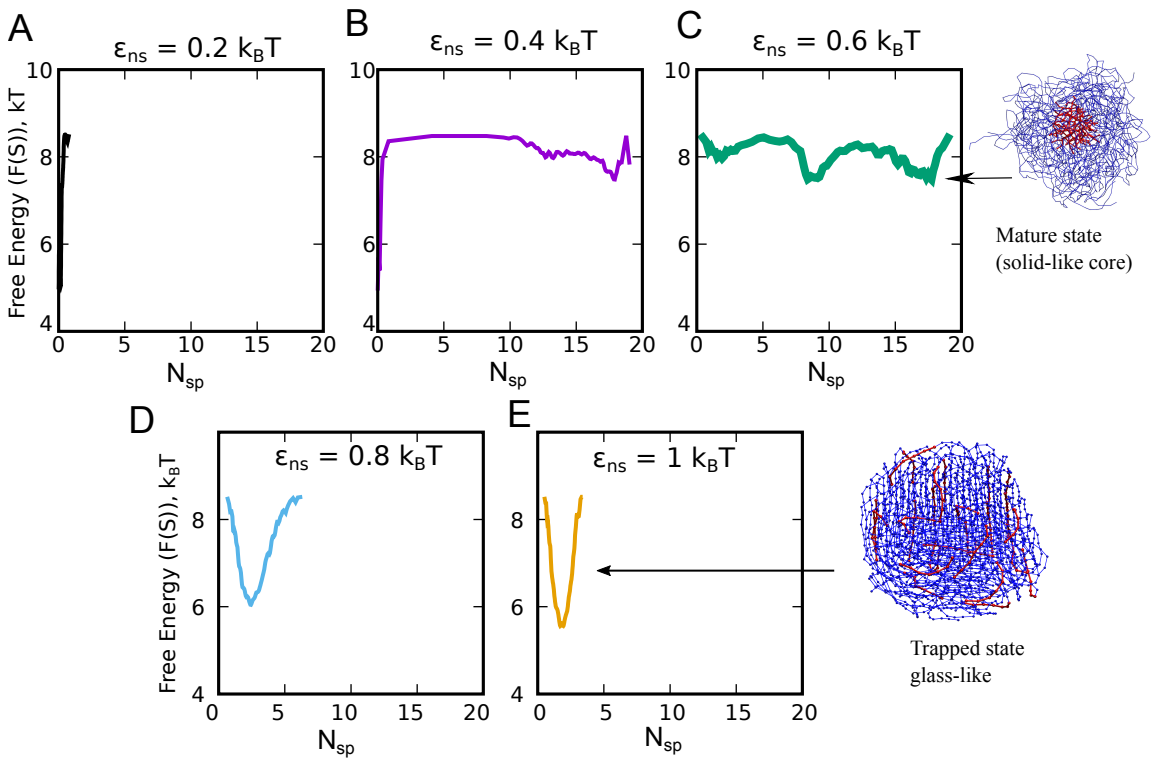

Figure 5: RELD simulations for the multimeric system. Free Energy profiles as a function of specific contacts as the order parameter. The free energy curves in A, B, C, D, E correspond to ϵ_{ns} values of 0.2, 0.4, 0.6, 0.8 and 1 k_BT, respectively. For ϵ_{ns} = 0.2 k_BT, the system remains predominantly in a monomeric state. For an intermediate value of ϵ_{ns} = 0.4 and 0.6 k_BT, we observe a sharp minima at higher values of specific contacts (per red bead), indicating that the system undergoes a transition to the mature state at this value of non-specific interaction strength. For ϵ_{ns} higher than 0.6 k_BT, the system remains trapped in the form of glass-like, non-specifically assembled aggregates, as evident by the sharp minima for low values of specific contacts.

To further establish the narrow range of ϵ_{ns} in which the mature complex is favored, we also performed RELD simulations for the system with 30 polymeric chains with 24 replicas, with temperatures ranging from 309 K to 401 K . We plot the resulting free energy profiles for the 30-mer system at a statistical temperature of 309 K, with N_{sp} as the order parameter used to identify the transition to the mature state with solid-like core (Fig. 5). Consistent with our RELD free energy profiles for the dimeric system, for weak ϵ_{ns} in the multimeric

system, we do not observe a defined minima at high values of N_{sp}. Also, for extremely strong ϵ_{ns} (Fig. 5D and E), we observe a sharp free energy minima for low values of specific contacts, suggesting that the system remains a kinetically trapped glass-like state and does not access the mature state in such a scenario. Interestingly, for an intermediate value of $\epsilon_{ns} = 0.4$ and 0.6 k_BT, we observe a defined minima for large specific contact numbers establishing that the mature state gets facilitated in such a regime of ϵ_{ns} (see Fig. 5B and C). Overall, these results suggest that the phase separation of protein and mature state (solid-like core) is favoured in a very narrow regime of ϵ_{ns}, and a deviation to either side would result in dramatically different equilibrium and dynamic properties.

Table 2: Heterogeneous mix of polymer definition and interaction involved in them and number of polymers in the system.

System names	**Different types of polymers in simulation box**	**Number of polymers**
Default	**type 1** polymer with blue and red beads having the following interactions. Blue-blue: $\epsilon_{ns} = 0.5k_BT$; blue-red: red-red: $\epsilon_{sp} = 1.2k_BT$	30 of type 1
System1	**type 1** as above; **type 2** polymer with black and red beads having following interactions black-black: no interaction, black red: no interaction, red-red: $\epsilon_{sp} = 1.2k_BT$, blue-black: no interaction	15 of type 1 & 15 of type 2
System 2	**type 1** as above; **type 3** polymer with green beads and red beads with following interactions green-green: $\epsilon_{ns} = 0.5k_BT$, green-red: $\epsilon_{ns} = 0.5k_BT$, red-red: $\epsilon_{sp} = 1.2k_BT$ and green-blue: no interaction.	15 of type 1 & 15 of type 3.

Heterogeneous mix of polymers: Phase separation and dynamics

So far, we have simulated multiple copies of the same kind of polymer (two bead polymer model Fig.1). The simulation suggests that critical strength of non-specific interaction is required for the formation of multimer and their subsequent maturation. In this section, we simulate a heterogeneous system composed of a mix of polymer chains and investigate their assembly into multimer for various scenarios described here.

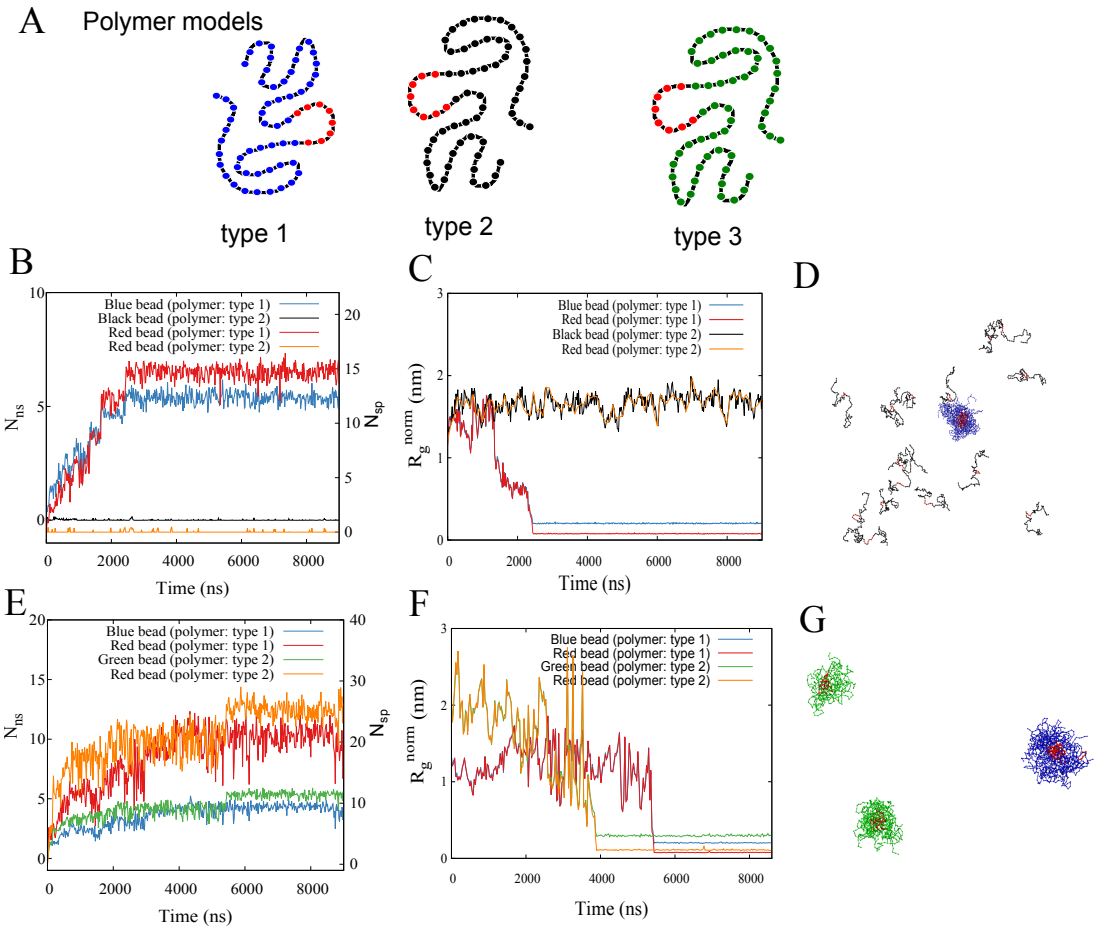

Figure 6: A) Polymer model: Cartoon of the Initial configuration of simulation, polymer **type1** – blue-red polymer with blue beads non-specifically interacting and red beads specifically interacting, **type 2** – black-red polymer, where black beads are non-interacting beads and red beads are specifically interacting, **type 3** – green-red polymer with non-specifically interacting green beads and specifically interacting red beads. B) Non-specific contacts per bead (N_{ns} in y1 axis– major y-axis) and specific contacts per bead (N_{sp}, y2 axis) are plotted over time for different types of beads from two polymers from System 1. C) Radius of gyration R_g^{norm} of different beads from two different polymers is plotted over time from System 1. D) Snapshot of the final trajectory from the simulation showing mature multimer formed by blue-red bead polymer and black-red polymer – freely diffusing in System 1. E) Non-specific contacts per bead (N_{ns} in y1 axis– major y-axis) and specific contacts per bead (N_{sp}, y2 axis) are plotted over time for different types of beads from two polymers from System 2. F) Radius of gyration R_g^{norm} of different beads from two different polymers is plotted over time from System 2. G) Snapshot of the final trajectory from the simulation shows the mature multimers formed in System 2.

In the first, we simulated two types of polymers described as *System 1* in **Table 2**. When these two different types of polymers are mixed and simulated, the question is whether we

will see the polymers with interacting beads (type 1) mixed between the polymers with non-interacting beads (type 2) or complete phase separation between interacting and non-interacting beads consisting of polymers. In Fig. 6B non-specific contacts per bead (as N_{sp} in y1-axis) for blue beads from type 1 and black beads from type 2 polymer and specific contacts per bead (as N_{sp} in y2-axis) for red beads in both polymers are plotted over time. There is a steady increase in contacts per bead (both N_{ns} and N_{sp}) in type 1 polymer, while no contacts are formed in type 2 polymer. In Fig. 6C, the radius of gyration (R_g^{norm}) for each group of beads is plotted over time; for beads in polymer type 1, R_g^{norm} is decreasing as a function of time, suggesting that these polymers are developing contacts. In comparison, the polymer type 2 (black and yellow curve Fig. 6B and C) have high R_g^{norm} and low contacts, while polymer type 1 (blue and red curve Fig. 6B and C) have low R_g^{norm} and high contacts; Red beads in polymer type 1 have even higher contacts and low R_g^{norm} than the blue beads suggesting the formation of mature multimer with a solid-like core by polymer of type 1. This can also be seen from a representative snapshot of a late time configuration from the simulation (see Fig. 6D). In summary, the polymers of two different types are phase separating even though they have a common red core for attracting. This suggests the importance of weak ϵ_{ns} in polymer self-assembly. Note that the polymer of type 2 not having weak ϵ_{ns} are randomly diffusing.

Given that the non-specific interaction is crucial for the self-assembly of polymers, the next question would be the polymers type 1 and 3 from *System 2* forms multimer (See **Table 2**). Essentially both polymers contain the same types of beads, but blue beads from type 1 and green beads from type 3 do not attract each other and have steric repulsion. N_{ns} for green and blue beads (y1 axis, Fig. 6E) along with N_{sp} of red beads (y2 axis, Fig. 6E) in polymer type 1 and type 3 are plotted, and the contacts for all the bead types from both polymers increase with time and equilibrate. While radius of gyration (R_g^{norm}) of the beads plotted in Fig. 6F decreases for all the types of beads from different polymers, showing the formation of the multimer. High contacts per bead and low R_g^{norm} for red beads from both polymers

show the formation of mature multimer in both the polymer types. In the snapshot of the late time configuration of simulation (Fig. 6G), mature multimers of two types of polymers are observed. The disliking nature of blue and green beads from respective polymers makes these distinct multimeric phases emerge.

These simulations are examples of the various types of multimer formation and phase separation that can be achieved in biology by tuning the parameters of non-specifically interacting beads and the interacting partners.

Three bead polymer model: non-specific regions drive mature multimers in a very narrow parameter regime

The simulations so far dealt with the self-organization of a homogeneous system of polymer chains made up of two kinds of beads. The results demonstrate how the interactions between the non-specific region can act as a handle for tuning the timescale of multimer formation and phase separation-like behavior of biopolymers. However, the molecular language of proteins composed of 20 different building blocks (amino acids) with different interaction propensities enables a more subtle tuning of the self-assembly/phase-separation phenomenon. We extend our two-bead model to three bead polymer model with a fraction of non-interacting beads. As in the current model, the polymer has a central patch of red beads having high specific interaction energy and a fraction of the remaining beads has non-specific interaction, and other beads are completely non-interacting (only steric effects) as shown Fig. 1C. This allows us to systematically study the interplay between the strength and concentration of non-specific interaction.

In this section, we demonstrate the amenability of the self-assembly/phase-separation landscape to be fine-tuned in a scenario where the ϵ_{ns} is dispersed across the non-functional regions of the polymer chain. In Fig. 7E and F, we monitor the self-organization of multiple 80-bead long hetero-polymer chains. For a fixed value of interaction strength of ϵ_{sp} =1.2 $k_B T$ between the core residues, we systematically vary the strength of non-specific interac-

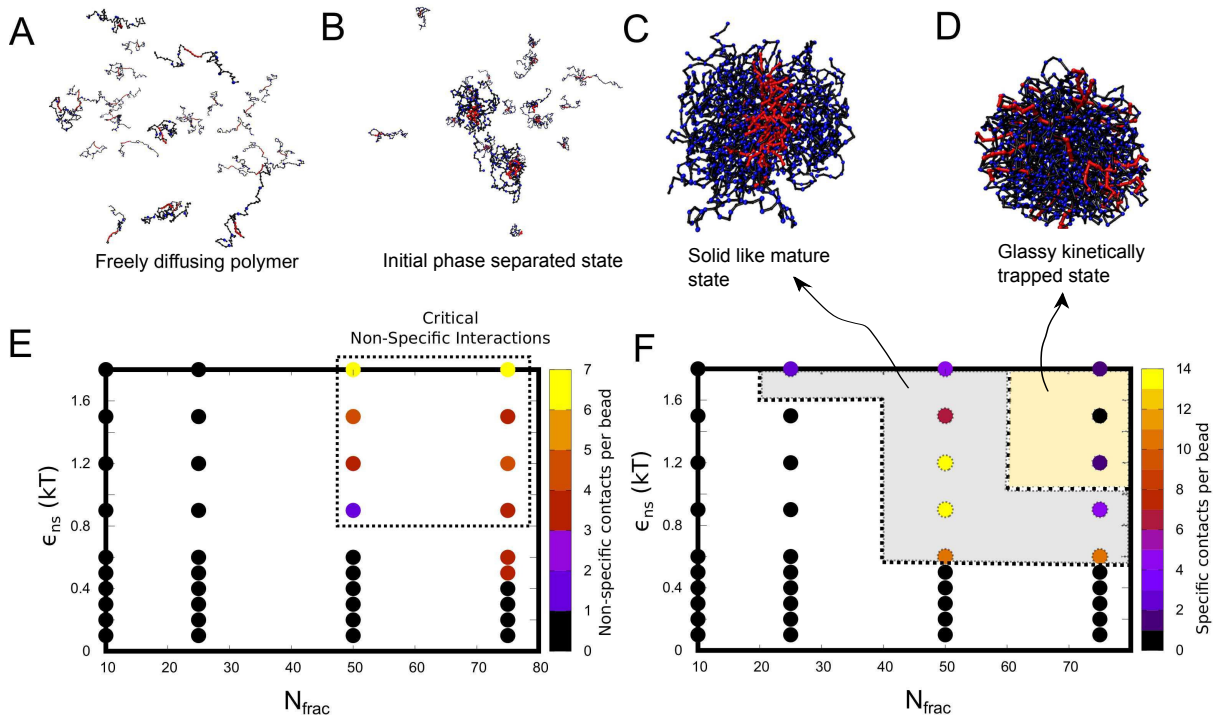

Figure 7: A) Snapshots of A) freely diffusing polymer in the simulation. B) early non-specifically bound initial multimer state, C) mature multimer and D) glass-like, kinetically trapped multimeric state for a 3-bead polymeric system. E) Non-specific contacts ,N_{ns}, per bead is plotted for varying non-specific interaction ϵ_{ns} and fraction of non-specifically interacting beads (N_{frac}, given in percentage). The dotted box indicate the critical values of non-specific interaction and N_{frac}–fraction of non-specific beads required for formation mature multimer complex. F) Specific contacts, N_{sp}, as a function of varying non-specific interaction, ϵ_{ns}, and fraction of non-specifically interacting beads. The grey area covered inside dotted line is the region where the formation of multimer leading to mature complex is plausible. The simulations were performed for a fixed specific interaction strength, ϵ_{sp}, of 1.2 k_BT.

tion ϵ_{ns} and fraction of non-specifically interacting residues (N_{frac}) to probe the parameter regime, wherein functional complex formation (red-red beads interaction or specific contacts formation) could occur.

At very low fraction of non-specifically interacting residues N_{frac}, the polymers are not involved in non-specific contacts (Fig. 7E, black dots). In this parameter regime of ϵ_{ns} and N_{frac}, the core regions of these chains do not form mature complexes suggesting that polymers would rather be freely diffusing state (see a snapshot of simulation in Fig. 7A) and the crucial role played by the non-functional residues (with ϵ_{ns}) information of functional assemblies. In

the regime of high non-specific interaction ($2k_BT$) with a low fraction of non-specific residues (25%) and low non-specific interaction ($0.6k_BT$) with a high fraction of non-specific residues (75%), there will be few clusters of multimers in the system as shown in the Fig. 7B, due to random interaction of polymers with cumulatively high interactions present. Further, in the parameter regime demarcated by the dotted boxed region in Fig. 7E wherein non-specific contacts start appearing, the polymer assemblies also form functional contacts (Fig. 7F, grey shaded region), suggesting polymers would have formed a mature multimer (see a snapshot of simulation in Fig. 7C). However, for extreme values of ϵ_{ns} and N_{frac}, there is a reduction in functional contacts (Fig. 7F, yellow shaded region) and the system would rather be in a trapped state as shown in Fig. 7D. This is because high non-specific interaction in many beads (75%) makes them interact more than the specific beads. These results corroborate the two-bead model's findings wherein a critical non-specific interaction strength is vital for self-assembly en route to maturation. Also, very few interactions between the residues of the non-core region can drive the formation of critical interactions between the core regions. In such a regime with a few interaction sites and a narrow parameter range for self-assembly, subtle modifications such as phosphorylation of these non-core residues could significantly alter the phase behavior, critical concentrations, and dynamicity of the condensates.[18,28–30] Conversely, few mutations resulting in extremely 'sticky' non-specific regions might result in loss of the ability of the polymers to form functional complexes.

Discussion and Conclusion

Biological reactions rely heavily on specificity and the ability of biomolecules to not only bind to the correct binding partners but also assume mature interaction configurations. However, the concentration of biomolecules like proteins is typically very low ($\approx$ nM) within cells, making the reactions malleable to tuning by external factors. The cellular environment is crowded with different types of biomolecules diffusing and promiscuously interacting with each other.[9]

Large stretches within protein sequences are not involved in the protein-protein interactions. This leads to an interesting question: can regions within proteins that are not involved in mature functional interfaces influence the process of formation of functional protein-protein complexes? At biologically relevant concentrations, could these non-interfacial residues have co-evolved to tune the binding affinities and dissociation constants of these proteins? How does the interplay between specific and promiscuous interactions shape the fate of these complexes?

Given the sheer size of the proteome and the vast number of protein-protein interactions that underlie cellular processes, phenomenological models could help us shed light on the physical mechanisms underlying the process of the formation of mature complexes.[25,31,32] Previous theoretical studies have attempted to understand the effect of promiscuous interactions on the ability of proteins to bind specifically to their partners.[19] However, there is a void in bottom-up studies elucidating the potential role of promiscuous interactions in promoting the formation of mature protein-protein complexes. To study this, we employed coarse-grained semi-flexible polymer chains and simulated them using Langevin dynamics and replica-exchange Langevin dynamics. The summary of our findings are as follows: (1) simulating dimer dynamics, we find that a narrow window of non-specific interaction strengths leads to a quick meeting of functional patches and formation of stable dimer complex (Fig. 2). (2) We show that the narrow optimal range of non-specific interaction strength also affects the formation of functional complexes with multiple polymers (a mature state with solid-like core) (Fig. 3). (3) Using RELD, we show that the optimal parameter regime leads to a free energy landscape having minima corresponding to the matured functional state of polymeric assemblies with specific interactions stabilising the assemblies (Fig. 5). (4) We then examined a heterogeneous mix of different types of polymers and showed that, depending on the nature of interactions, this leads to different kinds of complexes (Fig. 6). (5) We also simulated polymers varying the fraction of non-specifically interacting beads and their strengths and showed that there is a rich phase diagram emerging; the phase diagram

suggests how assembly states of polymers might be altered (by going from one phase to the other) in realistic situations (Fig. 7).

Various interactome studies suggest that non-interacting or noisy regions make up a large fraction of the proteome.[13] Further, several RNA bindings proteins are known to harbor intrinsically disordered regions, which facilitate RNP-granule assembly.[33,34] The spacer-sticker model for proteins has been previously employed to study the role of physical parameters such as specific interaction valencies and strengths in shaping equilibrium aspects of phase separation.[15,35] In these models, the stickers are considered to be the primary drivers of phase separation. This raises an interesting question – could the spacer regions play a role in shaping the kinetics of phase separation? In other words, can the non-specific interactions involving spacer regions influence inter-sticker interactions? This question gains importance in the context of membrane-less organelles, which facilitate the formation of inter-protein contacts within a high-density condensate phase.

Our simulations point towards a potential role of the non-interacting surfaces in facilitating the formation of mature protein-protein complexes. The current study provides a mechanistic basis for the narrow range of non-specific interaction strengths that promote protein-protein interactions. For the sake of simplicity, we modeled proteins as semi-flexible polymers with no defined native configuration. This makes it computationally tractable and allows us to sample a wider parameter space. The role of non-specific interactions in promoting the maturation of protein complexes with defined structural folds and lower conformational flexibility is to be further probed. Also, our simulations were either performed with one type or two types of polymers. In future, the simulations can be extended to multiple types of polymers in a system.

Methods

Despite the growing body of experimental literature on biomolecular phase separation, the multi-component nature of the problem makes it hard to unravel experimentally. Hence, Molecular dynamics simulations and computational models are very valuable and have been traditionally used to understand various aspects of biomolecular structural dynamics.[36] However, the scale of this problem makes it hard to probe systematically using conventional atomistic simulations. The widespread instances of the phenomenon mean that it can be treated as a problem of self-assembling polymer chains. To understand the contribution of noisy interactions in driving the formation of mature polymer complexes, we developed a semi-flexible polymer model.[37–41] In order to study the process of macromolecular chains finding each other, we performed Langevin dynamics simulations. In our study, we consider N flexible polymer chains, each made up of 'X' red beads and 'Y' blue beads, in a cubic box with periodic boundary conditions. The red beads represent residues involved in native functional interactions, while blue beads represent residues with the non-functional interaction Fig. 1.

Polymer chains of length 'M' beads are bonded via the harmonic potential with energy,

$$\mathrm{E_{stretching}} = \mathrm{k_s} \sum_{\mathrm{i=1}}^{\mathrm{M-1}} (|\vec{\mathrm{r}}_\mathrm{i} - \vec{\mathrm{r}}_{\mathrm{i+1}}| - \mathrm{r}_0)^2, \tag{1}$$

where $\vec{\mathrm{r}}_\mathrm{i}$ and $\vec{r}_{i+1}$ refer to i^{th} and $(i+1)^{th}$ bead positions, respectively; r_0 refers to the equilibrium bond length and k_s represents the spring constant. This interaction ensures the connectivity between the beads of a polymer chain. To model bending rigidity, any two neighboring bonds in a polymer interact via the following potential

$$\mathrm{E_{bending}} = \kappa \sum_{\mathrm{i=1}}^{\mathrm{M-2}} (1 - \cos\, \theta_\mathrm{i}), \tag{2}$$

where θ_i refers to the angle between i^{th} and $(i+1)^{th}$ bond, and κ is the bending stiffness.

All other non-bonded, inter-bead interactions were modeled using the Lennard-Jones (LJ) potential,

$$\mathrm{E_{nb}} = 4\epsilon \sum_{\mathrm{i<j}} \left[\left(\frac{\sigma}{|\vec{\mathrm{r_i}} - \vec{\mathrm{r_j}}|} \right)^{12} - \left(\frac{\sigma}{|\vec{\mathrm{r_i}} - \vec{\mathrm{r_j}}|} \right)^{6} \right], \tag{3}$$

for all $|\vec{\mathrm{r_i}} - \vec{\mathrm{r_j}}| < \mathrm{r_c}$, where $\mathrm{r_c}$ refers to the cutoff distance beyond which the non-bonded potentials are neglected. The cutoff for the LJ calculations was kept at 2.5 times of σ. Note that this function has two parts: a part that models steric repulsion when the beads overlap and an attractive part otherwise. For ease of implementation, in this model, this attractive part is used as an *effective interaction* that accounts for all attractive short-range forces. ϵ signifies the strength of the attractive interaction and has the units of energy. In our study, we choose different values of ϵ for interactions between red-red beads, as compared to those involving the blue beads. Since the interactions involving the blue beads model weak, non-specific contacts, they are lower in magnitude than the red-red interactions, which describe mature, native contacts. We systematically vary these two interactions to understand their role in driving the formation of mature complexes, keeping $\epsilon_{\mathrm{sp}} > \epsilon_{\mathrm{ns}}$. Note that the bending energy parameter (κ) controls the stiffness of the individual polymer chains, while the LJ potential parameter (ϵ) controls the interaction between any two beads (inter-chain and intra-chain).

Simulation parameters: In our study, we used the LAMMPS molecular dynamics package to perform the dynamic simulations,[42] where the simulator solves Newton's equations with viscous force and a Langevin thermostat, ensuring an NVT ensemble with temperature T=310 K. An integration timestep (dt) of 30 fs was used for the simulations. Polymer chains of length M=80 beads and mass of each bead was considered to be 110 Da (average mass of amino acids). The size parameter for the beads is taken as $\sigma = 4.5$Å. The parameters for the bonded springs was fixed as $\mathrm{r_0} = 4.5$Å and $\mathrm{k_s} = 10$ kcal/mol. The damping time for the Langevin thermostat was 1.2 ps. Similar values for these parameters have been previously used to perform coarse-grained protein simulations by Bellesia et al. and Bieler et al.[43,44]

Replica Exchange Langevin Dynamics: In order to calculate the free energy profiles for

the dimeric and multi-polymer systems (Fig. 2C and Fig. 5), we employed the parallel tempering Langevin dynamics technique. In this approach, 24 independent, non-interacting copies (replicas), each at different temperatures between 309 K - 401 K each 4 K apart, were simulated under the NVT ensemble. In order to ensure proper sampling of equilibrium states, an exchange of configurations between different replicas is attempted every 0.5 ns. In this algorithm, a successful swap of configurations between replicas (at different temperatures) occurs based on the Boltzmann probability.[45,46] Each of the individual replicas was simulated for a timescale of 9 μs. Using the combination of conventional Molecular dynamics and the Monte Carlo algorithm, this technique has been commonly used to enable efficient sampling of the conformational space for problems such as protein folding and self-assembly.

Acknowledgements

The authors would like to thank Prof. B.J. Rao, Prof. Dibyendu Das and Prof. M.Muthukumar for their suggestions and discussions related to the work.

References

(1) Boeynaems, S.; Alberti, S.; Fawzi, N. L.; Mittag, T.; Polymenidou, M.; Rousseau, F.; Schymkowitz, J.; Shorter, J.; Wolozin, B.; Van Den Bosch, L.; Tompa, P.; Fuxreiter, M. Protein Phase Separation: A New Phase in Cell Biology. *Trends in Cell Biology* **2018**, *28*, 420–435.

(2) Caragine, C. M.; Haley, S. C.; Zidovska, A. Nucleolar dynamics and interactions with nucleoplasm in living cells. **2019**, 1–21.

(3) Saha, S.; Weber, C. A.; Nousch, M.; Adame-Arana, O.; Hoege, C.; Hein, M. Y.; Osborne-Nishimura, E.; Mahamid, J.; Jahnel, M.; Jawerth, L.; Pozniakovski, A.; Eckmann, C. R.; Jülicher, F.; Hyman, A. A. Polar Positioning of Phase-Separated Liquid

Compartments in Cells Regulated by an mRNA Competition Mechanism. *Cell* **2016**, *166*, 1572–1584.e16.

(4) Bergeron-Sandoval, L.-P.; Safaee, N.; Michnick, S. Mechanisms and Consequences of Macromolecular Phase Separation. *Cell* **2016**, *165*, 1067 – 1079.

(5) Landry, C.; Levy, E.; Rabbo, D. A.; Tarassov, K.; Michnick, S. Extracting Insight from Noisy Cellular Networks. *Cell* **2013**, *155*, 983 – 989.

(6) Landry, C. R.; Levy, E. D.; Michnick, S. W. Weak functional constraints on phosphoproteomes. *Trends in Genetics* **2009**, *25*, 193 – 197.

(7) Levy, E. D.; Michnick, S. W.; Landry, C. R. Protein abundance is key to distinguish promiscuous from functional phosphorylation based on evolutionary information. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2012**, *367*, 2594–2606.

(8) Tarassov, K.; Messier, V.; Landry, C. R.; Radinovic, S.; Molina, M. M. S.; Shames, I.; Malitskaya, Y.; Vogel, J.; Bussey, H.; Michnick, S. W. An in Vivo Map of the Yeast Protein Interactome. *Science* **2008**, *320*, 1465–1470.

(9) Deeds, E. J.; Ashenberg, O.; Gerardin, J.; Shakhnovich, E. I. Robust protein–protein interactions in crowded cellular environments. *Proceedings of the National Academy of Sciences* **2007**, *104*, 14952–14957.

(10) Das, R. K.; Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 13392–13397.

(11) Statt, A.; Casademunt, H.; Brangwynne, C. P.; Panagiotopoulos, A. Z. Model for disordered proteins with strongly sequence-dependent liquid phase behavior. **2019**, 23–27.

(12) Vacha, R.; Frenkel, D. Relation between Molecular Shape and the Morphology of Self-Assembling Aggregates: A Simulation Study. *Biophys. J.* **2011**, *101*, 1432 – 1439.

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

(14) Osmanović, D.; Rabin, Y. Effect of non-specific interactions on formation and stability of specific complexes. *Journal of Chemical Physics* **2016**,

(15) Choi, J.-M.; Holehouse, A. S.; Pappu, R. V. Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annual Review of Biophysics* **2020**, *49*, 107–133, PMID: 32004090.

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

(17) Zhang, J.; Muthukumar, M. Simulations of nucleation and elongation of amyloid fibrils. *The Journal of Chemical Physics* **2009**, *130*, 035102.

(18) Hofweber, M.; Dormann, D. Friend or foe — post-translational modifications as regulators of phase separation and RNP granule dynamics. *Journal of Biological Chemistry* **2018**,

(19) Lukatsky, D. B.; Afek, A.; Shakhnovich, E. I. Sequence correlations shape protein promiscuity. *Journal of Chemical Physics* **2011**, *135*, 1–15.

(20) Sharma, R.; Raduly, Z.; Miskei, M.; Fuxreiter, M. Fuzzy complexes: Specific binding without complete folding. *FEBS Letters* **2015**, *589*, 2533–2542.

(21) Protter, D. S.; Parker, R. Principles and Properties of Stress Granules. *Trends in Cell Biology* **2016**, *26*, 668 – 679.

(22) Das, R. K.; Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 13392–13397.

(23) Feric, M.; Vaidya, N.; Harmon, T. S.; Mitrea, D. M.; Zhu, L.; Richardson, T. M.; Kriwacki, R. W.; Pappu, R. V.; Brangwynne, C. P. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **2016**, *165*, 1686–1697.

(24) Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology* **2017**, *18*, 285.

(25) Tiwary, A. K.; Zheng, Y. Protein phase separation in mitosis. *Current Opinion in Cell Biology* **2019**, *60*, 92–98.

(26) Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019**, *176*, 419–434.

(27) Ray, S. et al. α-Synuclein aggregation nucleates through liquid–liquid phase separation. *Nature Chemistry* **2020**,

(28) Aumiller, W. M.; Keating, C. D. Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. *Nature Chemistry* **2016**,

(29) Ferreon, J. C.; Jain, A.; Choi, K.-J.; Tsoi, P. S.; MacKenzie, K. R.; Jung, S. Y.; Ferreon, A. C. Acetylation Disfavors Tau Phase Separation. *International journal of molecular sciences* **2018**, *19*, 1360.

(30) Alexander, C.; Toptygin, D.; Austin Elam, W.; TP, S.; Hilser, V. Phosphorylation Increases Persistence Length and End-to-End Distance of a Segment of Tau Protein. *Biophys. J.* **2016**, *110*, 362 – 371.

(31) Majumdar, A.; Dogra, P.; Maity, S.; Mukhopadhyay, S. Liquid-Liquid Phase Separation Is Driven by Large-Scale Conformational Unwinding and Fluctuations of Intrinsically Disordered Protein Molecules. *Journal of Physical Chemistry Letters* **2019**, *10*, 3929–3936.

(32) Shin, Y.; Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **2017**, *357*.

(33) Murthy, A. C.; Dignon, G. L.; Kan, Y.; Zerze, G. H.; Parekh, S. H.; Mittal, J.; Fawzi, N. L. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nature Structural and Molecular Biology* **2019**,

(34) Patel, A. et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **2015**,

(35) Ranganathan, S.; Shakhnovich, E. I. Dynamic metastable long-living droplets formed by sticker-spacer proteins. *eLife* **2020**,

(36) Ma, B.; Nussinov, R. Molecular dynamics simulations of alanine rich β-sheet oligomers: Insight into amyloid formation. *Protein Sci.* **2002**, *11*, 2335–2350.

(37) Gillam, J. E.; MacPhee, C. E. Modelling amyloid fibril formation kinetics: mechanisms of nucleation and growth. *J. Phys.: Condens. Matter* **2013**, *25*, 373101.

(38) Hnisz, D.; Shrinivas, K.; Young, R. A.; Chakraborty, A. K.; Sharp, P. A. A Phase Separation Model for Transcriptional Control. *Cell* **2017**, *169*, 13 – 23.

(39) Ranganathan, S.; Maji, S. K.; Padinhateeri, R. Defining a Physical Basis for Diversity in Protein Self-Assemblies Using a Minimal Model. *Journal of the American Chemical Society* **2016**, *138*, 13911–13922.

(40) Ghosh, S.; Ghosh, D.; Ranganathan, S.; Anoop, A.; P, S. K.; Jha, N. N.; Padinhateeri, R.; Maji, S. K. Investigating the Intrinsic Aggregation Potential of Evolutionarily Conserved Segments in p53. *Biochemistry* **2014**, *53*, 5995–6010, PMID: 25181279.

(41) Jha, N. N.; Anoop, A.; Ranganathan, S.; Mohite, G. M.; Padinhateeri, R.; Maji, S. K. Characterization of amyloid formation by glucagon-like peptides: role of basic residues in heparin-mediated aggregation. *Biochemistry* **2013**, *52*, 8800–10.

(42) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1 – 19.

(43) Bellesia, G.; Shea, J.-E. Effect of β-sheet propensity on peptide aggregation. *J. Chem. Phys.* **2009**, *130*.

(44) Bieler, N. S.; Knowles, T. P. J.; Frenkel, D.; Vácha, R. Connecting Macroscopic Observables and Microscopic Assembly Events in Amyloid Formation Using Coarse Grained Simulations. *PLoS Comput. Biol.* **2012**, *8*, e1002692.

(45) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters* **1999**,

(46) Qi, R.; Wei, G.; Ma, B.; Nussinov, R. *Methods in Molecular Biology*; 2018.