Separation of sticker-spacer energetics governs the coalescence of biomolecular condensates

Aniruddha Chattaraj, Eugene I Shakhnovich

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA

Correspondence: shakhnovich@chemistry.harvard.edu

Abstract

Biopolymer condensates often emerge as a multi-droplet state and never coalesce into one large droplet within the experimental timespan. This contradicts the prediction of classical polymer physics which suggests the existence of one large droplet beyond the phase transition. Previous work revealed that the sticker-spacer architecture of biopolymers may dynamically stabilize the multi-droplet state. Here, we simulate the condensate coalescence using metadynamics approach and reveal two distinct physical mechanisms underlying the fusion of droplets. Condensates made of sticker-spacer polymers readily undergo a kinetic arrest when stickers exhibit slow exchange while fast exchanging stickers at similar levels of saturation allow merger to equilibrium states. On the other hand, condensates composed of homopolymers fuse readily until they reach a threshold density. We also show that the inter-condensate exchange of chains offers a general mechanism that drives the fusion. We map the range of mechanisms of kinetic arrest from slow sticker exchange dynamics to density mediated in terms of energetic separation of stickers and spacers. In that vein, our simulations predict that the biological condensates may show a re-entrant dynamic merger behavior at a particular range of sticker-spacer energetics. Mechanistic underpinnings revealed in this work have broader implications in polymer physics and condensate biology.

Introduction

Biomolecular condensates are membrane-less intracellular compartments that emerge via phase transition [1, 2]. Condensates perform a range of spatiotemporal biochemical tasks across multiple scales [3]. Dysregulation of condensate biology has been implicated in many pathological conditions, including neurodegenerative diseases [4, 5].

Condensate formation is a coupling between two distinct phase transitions – density transition and percolation transition [6-8]. Biopolymers involved in such processes often have an architecture of multivalent associative heteropolymers. Such polymers are commonly modelled with a “sticker-spacer” framework [9, 10]. A “sticker” is a cohesive region of the polymer sequence that may engage in inter or intra-chain interactions. Two successive stickers are interspersed by less “sticky” linker regions known as “spacers”. When stickers of same type interact with each other, it is called homotypic interactions. On the contrary, heterotypic interaction refers to the interactions between different sticker types. Linear multi-domain proteins connected by flexible linkers (poly-SH3, for example) serves as a prototypical example of sticker-spacer polymer where the structured domains act as stickers and the linkers behave as spacers [11, 12]. Intrinsically disordered proteins (IDPs) form another important class where amino acids with aromatic side chains may serve as stickers while rest of the amino acids function as spacers [13]. Stickers form transient physical crosslinks (“bonds”) to generate multi-chain network (“cluster”) and beyond a concentration threshold, the system undergoes a network transition known as percolation or gelation when the system shows tendency to create large (infinite in thermodynamic limit) networks. With the presence of flexible spacers, such percolation is accompanied by a density transition where the percolated clusters separate out from the solution to create a “polymer-rich” dense phase.

Physics of biomolecular phase separation immensely benefits from classical theories of homopolymers [14, 15] which describes the interplay of entropy and energy in determining the mixed or demixed configuration of such systems. Classical theory predicts two outcomes: a dispersed state at lower concentrations and a single large droplet coexisting with dilute phase at higher concentrations. However, till now , the majority of in-vitro and cellular experiments revealed a multi-droplet state where the droplets dynamically exchange components with each other but

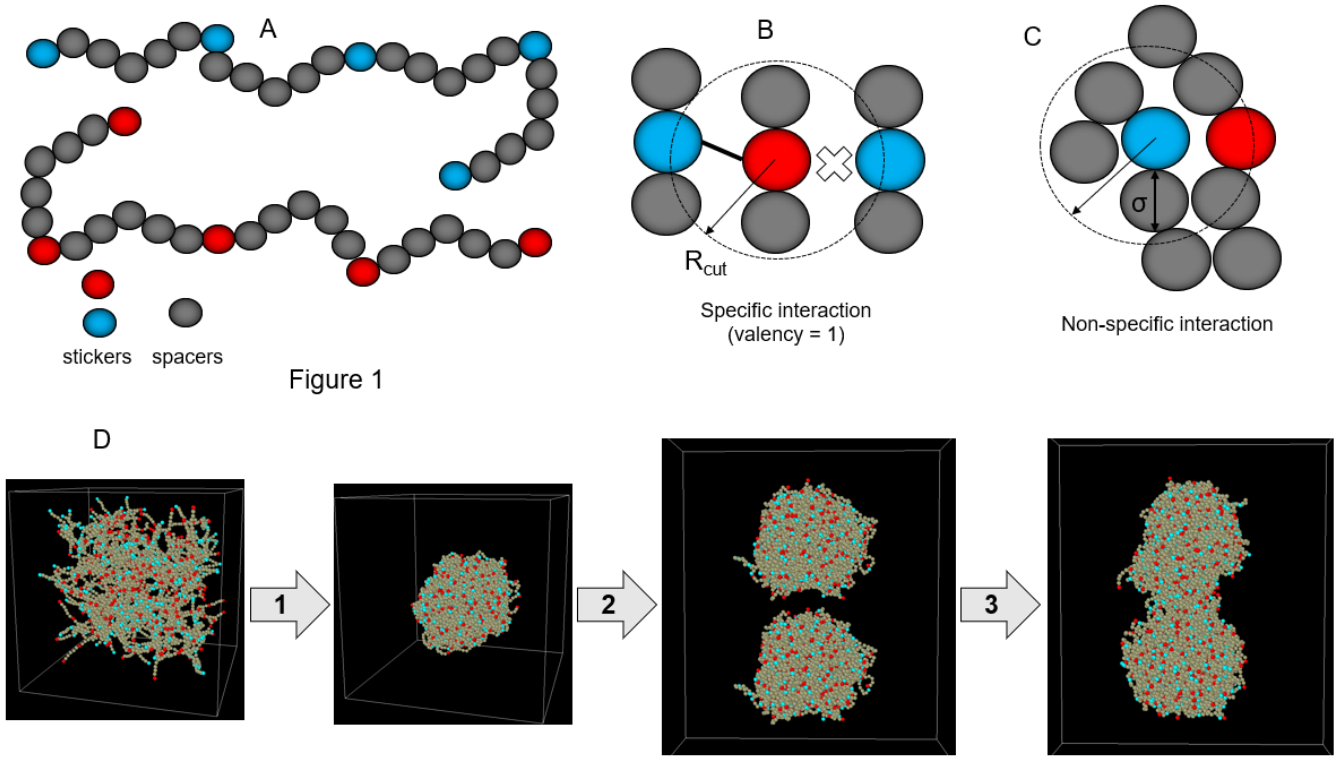


Figure 1: Workflow of the two-cluster fusion simulation by metadynamics approach. (A) Illustration of the two-component system. Each component is an associative heteropolymer consisting of 5 stickers (red and cyan beads) and 20 spacers (grey beads). We only allow heterotypic interactions, that is, red stickers interact with cyan stickers, but red-red or cyan-cyan are not allowed. (B) Stickers engage in specific interactions. Once a pair of stickers come closer to a cut-off distance (Rcut), they may form a "bond" with probability, p_{on}. The bond may dissociate with a probability, p_{off}. The bond is modelled as a harmonic distance constraint with stretching energy, Es. Both probabilities (p_{on}, p_{off}) are set to 1. The lifetime of the bond depends on Es. Specific interactions mimic cognate biomolecular interactions. Once a pair of stickers is bonded, they cannot engage with another sticker that may be present within Rcut. In other words, each sticker has a valency of 1. (C) Spacers interact via non-specific forces, modelled by Lennard-Jones potential. Each bead (both stickers and spacers) can exert a long-range attractive force (within a cut-off radius, dotted circle) and short-range repulsive force which determines the bead diameter (σ). One bead can interact with multiple beads, permitted by volume exclusions. The magnitude of the non-specific energy (Ens) dictates the dwelling time, that is, how much time two beads will spend on each other's vicinity. (D) Setup of the metadynamics simulations. In step 1, we start with 200 randomly distributed chains (100 red-types + 100 cyan-types) and bias the system, along the order parameter - Rg^{system} , to condense into one large cluster. Rg^{system} represents the radius of gyration of the entire system. In step 2, using standard Langevin dynamics (without any bias), we relax the cluster with a pair of Es and Ens. The relaxed cluster is then copied with an initial separation of $3*Rg^{cluster}$, where $Rg^{cluster}$ is the radius of gyration of the relaxed configuration. Finally, in step 3, we make the two clusters fuse with biasing potential along the distance between cluster centers. Same energy pair (Es, Ens) is used as in step 2.

rarely coalesce to become one. Multiple mechanisms have been proposed to explain this conundrum. Living cells operate far from the thermodynamic equilibrium by consuming energy. Such active processes may play important roles in determining the condensate size distribution [16]. Also, certain proteins may act as surfactant [17-19] by selectively getting adsorbed to the condensate surface. Such spatial architecture stabilizes multiple condensates.

However, in-vitro systems devoid of such complexities also demonstrate long-living droplets. We have recently proposed that such droplets exist in a dynamically arrested metastable state [8]. The interplay of diffusion and reaction timescales determines the size distribution of condensates. The concept of dynamic arrest has been invoked recently in explaining observations regarding multi-layered condensates [20] or assembly of membrane-bound condensates [21]. Competition between nucleation and coalescence kinetics [22] is also proposed as a mechanism to regulate the condensate size distribution.

Maturation or ageing [23] is another interesting kinetic phenomenon where material properties of condensates gradually change over time. Specifically, the freshly formed droplets show dynamic molecular rearrangements and fuse readily once they come into contact. Surprisingly, these behaviors are reversed after waiting for a long time, that is, these “aged” condensates partially fuse and show a reduced degree of dynamicity in terms of molecular rearrangements. A recent modeling study [24], pertinent to our current investigation, explored the role of kinetic arrest in condensate ageing and the resultant shape. They have shown that the strength of inter-protein interactions serves as a control parameter that alters the material properties of condensates.

Despite all the existing efforts, a generic mechanistic understanding underlying coalescence of condensates is lacking. In the current study, we seek to reveal such mechanisms. What is the physical force that drives the fusion of two droplets? What is the role of sticker-spacer architecture in steering such processes? How similar are they compared to homopolymers?

Using a heterotypic system, composed of two types of sticker-spacer model polymers, we have identified two distinct mechanisms underlying the fusion of condensates. Here we report that the relative energetic contributions of stickers and spacers is instrumental in determining “fusibility” of condensates. We show that fusion is governed by an intricate interplay of energy and entropy, and the sticker saturation effect triggers a kinetically arrested metastable state that prevents two condensates from fusing. In contrast, homopolymers undergo a density-mediated arrest where two tightly packed condensates cannot inter-penetrate. We believe that the mechanistic

underpinnings revealed in this study will inspire the development of new theory in polymer physics. It should also help in designing biopolymers with desired material properties.

Methods

1. Model Construction

We used bead-spring polymers for coarse-grained representation of biomacromolecules (proteins and nucleic acids), with similar force-fields as our previous studies [8, 25]. We considered a pair of sticker-spacer polymers (Fig. 1A) where stickers engage in heterotypic interactions (red and cyan), but homotypic interactions are not allowed. Since the focus of the current study is to derive generic principles underlying condensate coalescence, we did not attempt to model any specific system. However, the engineered protein systems like SH3-PRM or SUMO-SIM are best examples of such heterotypic systems [26] where stickers of complementary types form cross-links to create intra-condensate networks.

Each polymeric chain contains 25 beads (5 stickers + 20 spacers), connected by harmonic bonds. Stretching energy of each harmonic bond, $E_{bond} = K_b * (R - R_0)^2$ where K_b is the spring constant and R_0 is the equilibrium bond distance. R measures the distance between the bonded beads at any given time. Each bond is a harmonic distance constraint that enables permanent connectivity within a chain. In our model, R_0 = 10 Å and K_b = 3 $\frac{kcal}{mol*\text{Å}^2}$.

To allow chain flexibility, angle (θ) between three successive beads is modelled with a cosine function: $E_{bending} = \kappa * (1 - cos\theta)$ where κ decides the bending stiffness. In our model, $\kappa = 2\ kcal.mol^{-1}$.

Each pair of beads interacts via a non-bonded isotropic interaction, modelled by Lennard-Jones (LJ) potential: $E_{LJ} = 4 * \epsilon * [\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}]$ where σ represents the bead diameter and r is the separation between the beads. ϵ determines the strength of attractive potential which we will describe in the next section. The LJ potential enforces excluded volume effect when diameters of two beads overlap. It also models weak inter-particle attractive force at intermediate distances. To achieve computational efficiency, the LJ potential is neglected beyond a cut-off distance (R_{max}). In our model, $\sigma = 10$ Å, $R_{max} = 2.5 * \sigma$.

To encode molecular site-specific biochemistry, we introduced inter-sticker bonds (Fig. 1B). When two stickers approach within a cut-off radius (R_{cut}), they can form a “bond” with a probability, p_{on}.

The bond can be broken with a probability of p_{off}, at or greater the distance of R_{cut}. These bonds are harmonic distance constraint (equilibrium distance = R_0^s) governed by a spring constant parameter. Energy of these inter-sticker bonds, $E_{Sb} = Ks * (R - R_0^s)^2$. Since the spring constant (Ks) decides the stiffness of the bond, it also sets the energy cost of breaking the bond. We will refer to this tunable spring constant as $E_{specific}$ or Es. These are "specific" interactions because once a sticker pair is bonded, they can't form another bond (Fig. 1B) with complementary stickers that are still within R_{cut}. In other words, each sticker has a valency of 1. In our model, $p_{on} = p_{off} = 1,\ R_0^s = 1.122 * \sigma,\ R_{cut} = R_0^s + 1.5$ Å.

To distinguish from specific inter-sticker interactions, the energy parameter (ϵ) of attractive component of LJ potential will be called Ens ($E_{nonspecific}$). Unlike specific interactions, beads governed by Ens can form multiple "contacts" with neighboring beads permitted by volume exclusions (Fig. 1C). The magnitude of Ens decides the dwell time, that is, how long a group of interacting beads spend time on each other's vicinity. We will use "bonds" and "contacts" to refer to the specific and non-specific interactions. In this work, we systematically alter Es and Ens to probe their effects on structure and fusion property of multi-chain clusters.

It is important to emphasize the unit convention of the energy parameters. As stated earlier, Es is basically the spring constant of the reversible bonds (inter-stickers) with a unit of kcal / (mol *Å^2). Since we don't change the equilibrium bond distances, for convenience, we write Es in the unit of energy (kcal /mol). Strictly speaking, there is always a proportionality constant with the unit of 1/Å^2. It follows from the LJ potential expression that Ens has a unit of energy (kcal /mol). We express the energies (both Es and Ens) in the unit of thermal energy or k_BT where k_B is the Boltzmann constant, and T is the absolute temperature of the system. We will, hereafter, simply use "kT" for notational simplicity.

2. Simulation details

We have used the LAMMPS package [27, 28] to perform Langevin dynamics simulations of our two-component polymer system. Langevin dynamics captures the Brownian motion of particles by introducing a stochastic force at each timestep, on top of standard Newtonian dynamics. We performed our simulations in a cubic box (fixed volume) with periodic boundary conditions. Simulation temperature is 310 K. Mass of each bead is set to 1000 Dalton which roughly corresponds to 10 amino acids. With this spatial resolution, a chain composed of 25 beads refers to 250 amino acids. The viscosity of the simulation medium is described with a "damp" parameter

which is set to 500 femtoseconds (fs). The “damp” parameter is inversely proportional to the viscosity of the solvent. Simulation timestep = 25 fs.

We have used metadynamics simulations [29] to facilitate the processes of polymer coalescence and cluster fusion (Fig. 1D). Metadynamics is an enhanced sampling scheme where the auxiliary gaussian potentials are imposed along a user defined order parameter (also known as collective variables, reaction coordinates etc.) to reconstruct the free energy profile of the process of interest. We have previously employed metadynamics [25, 30] to study how proteins self-assemble into liquid or solid-like condensates. In this work, we performed two types of metadynamics simulations (Fig. 1D). First, we used R_g^{system} ,as the order parameter to bias in 200 million timesteps the coalescence of 200 randomly distributed chains into one large cluster (step 1). R_g^{system} is the radius of gyration of the center of masses of 200 chains. To be precise, we considered the 13th (middle) bead of each chain as its center and used their locations to compute the R_g^{system}. For this step, Es = 3kT and Ens = 0.5kT. The fully clustered state represents the free energy minimum (Fig. S1). We then used standard Brownian dynamics (without any bias, 100 million timesteps) to relax the cluster with a selected pair of energetic parameters Es and Ens (step 2). We then copied the relaxed cluster into two as an initial condition with an initial separation between the two clusters of $3 * R_g^{cluster}$ where $R_g^{cluster}$ is the radius of gyration of the relaxed configuration. We then performed a two-cluster fusion metadynamics simulation (step 3) with a biasing potential along the distance between the cluster centers. The center of a cluster is again defined with the location of the middle bead of each chain.

3. Data analysis

For metadynamics simulations, we analyzed the time evolution of order parameters and related potential mean force (PMF) profile. To analyze physical properties of the clusters (density, sticker saturation etc.), we used the configuration files (“restart” files in LAMMPS) containing information of coordinates and topology of the system. The topology information is converted into a network and average properties like cluster (network) density, shortest path length, degree distribution of nodes (sticker saturation) etc. are computed. To compute the shortest path length distribution (discussed in Fig. 3C), we select one node from the sticker-spacer network and calculate the shortest paths to all other nodes. A topological path connects two nodes in a network and the number of intermediate nodes defines the path length.

To calculate the bond entropy (discussed in Fig. 5), we used the distribution of bonds inside (intra) and between (inter) the clusters. We divide the total bonds into three states or “labels” – cluster11,

cluster22 and cluster12, where 1 and 2 are the cluster indices. So, each bond may belong to one of the three states. The bond distribution of the system is then characterized by an information entropy, $H = -[p_{11}\, logp_{11} + p_{22}\, logp_{22} + p_{12}\, logp_{12}]$, where p_{11}, p_{22}, p_{12} compute the probability of a bond being intracluster1, intracluster2 and intercluster respectively. For this three-state representation, p = 0.33 for a fully mixed (fused) configuration; hence maximum entropy, H_{max} = 1.1.

To compute the surface-to-volume ratio (also discussed in Fig. 5), we considered the coordinates of all 10000 beads (400 chains) and fitted it to a convex hull. We then computed the surface and volume of the topological object at multiple timepoints to get the time course.

4. Software

Moltemplate: To create the model polymers, we made use of the Moltemplate [31] package which enables the user to create multiple types of chains in a template based manner.

Packmol: The polymers are packed inside the simulation volume using the PACKMOL [32] package.

LAMMPS: The Langevin dynamics simulations are performed using the LAMMPS software package [28]. We used the "COLVARS" functionality [33] to perform metadynamics within LAMMPS.

OVITO: We used the OVITO (basic version) software to visualize the particle motions [34].

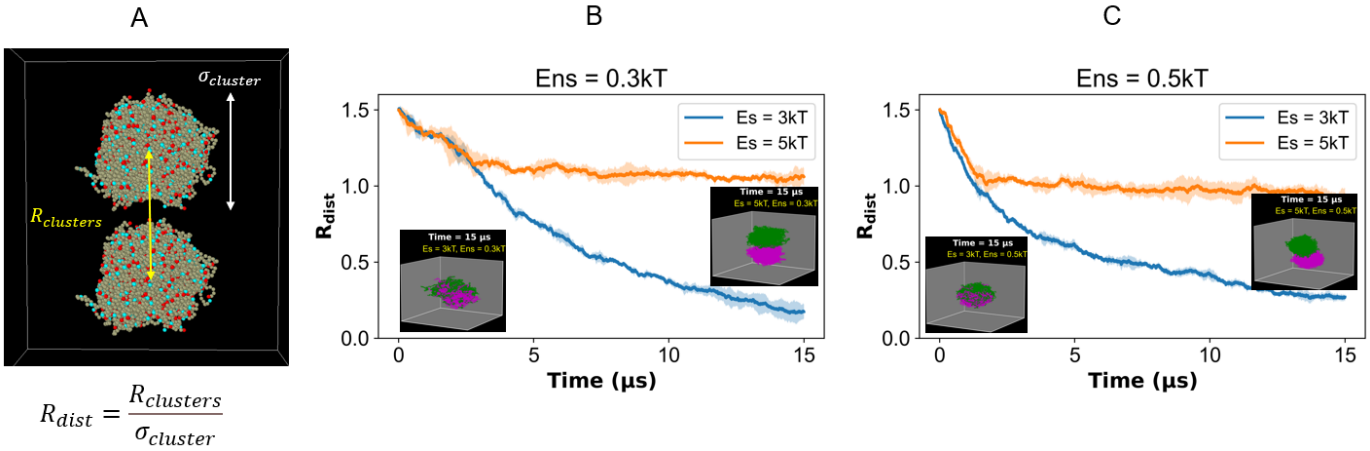
Python: We used custom python scripts to setup simulations and analyze the data.

We have organized and released example simulations and python code. GitHub Link

Results

1. Strong sticker interactions yield metastable clusters

Using a series of two-cluster fusion simulations, we first seek to understand how the fusion tendency of clusters depends on specific interaction strengths (Es). The dynamics of fusion is observed with a reduced parameter – "relative distance" or R_{dist} (Fig. 2A). When $R_{dist} \sim 1$, two clusters are in contact, and it gradually goes down as the fusion proceeds. For a completely fused state, $R_{dist} = 0$. We note that the analysis depends on the sphericity of the clusters. For perfectly



$$R_{dist} = \frac{R_{clusters}}{\sigma_{cluster}}$$

Figure 2: Quantification of cluster fusion. (A) We express the inter-cluster distance with a dimensionless "relative distance" parameter: R_{dist} which is the inter-cluster distance ($R_{clusters}$) normalized by the cluster diameter ($\sigma_{cluster}$). (B) Fusion behavior of clusters under stronger (Es = 5kT, orange line) and weaker (Es = 3kT, blue line) specific interaction strengths. Non-specific interaction strength, Ens = 0.3kT. Each trajectory is an average over 5 stochastic runs (Solid line: mean, fluctuation envelop: standard deviation). Insets display the representative snapshots of last simulation timeframe (parameters mentioned in the labels). Two clusters are colored differently for visual clarity. (C) Same configurations as B, except Ens = 0.5kT.

spherical clusters, R_{dist} = 1 reflects the surface contact. Our simulated clusters are far from perfect spheres, so R_{dist} provides an approximate measure of the extent of cluster penetration.

Firstly, we delineate the phase transition boundary from the cluster relaxation dynamics (Fig. 1D, step 2). When we relax the cluster under different pairs of Es and Ens, without any bias, the minimum combination (Es, Ens) which stabilizes the cluster is 3kT, 0.3kT (Fig.S2). Using this pair as our threshold, we gradually increase both Es and Ens and observe their fusion properties. Fig. 2B shows that, with stronger Es (5kT), two clusters do not fuse (orange line, $R_{dist} \geq 1$) even though they touch each other. At a relatively lower Es (3kT, blue line), clusters readily fuse. These behaviors are replicated at a different non-specific interaction strength, Ens = 0.5kT (Fig. 2C). For both Ens, two clusters (colored in green and magenta) cannot penetrate ("arrested state") for the higher Es, but fuse readily for the lower Es. Although Figs. 2B and 2C show qualitative similar behaviors, it is worthy to highlight some subtle differences between the fusion profiles. For Ens = 0.5kT (Fig. 2C), the orange line (Es = 5kT) runs closer to the Rdist = 1 level, compared to the Es = 5kT line in Ens = 0.3kT case (Fig. 2C). The higher non-specific contact energy provides incentive to the clusters to initiate fusion, but stronger Es (5kT) prevents it. On the other hand, once the fusion proceeds, blue line with Ens = 0.5kT reaches to a slightly higher level (last point in Fig. 2C) than the Ens = 0.3kT case (last blue point in Fig. 2B). This could be a consequence of higher Ens mediated bead density (shown in Fig. 2A) that makes the clusters more viscous. The movie S1 helps us to better appreciate the differences between the fusion behaviors under these four parameter combinations. Now, if we start our simulations with 400 randomly distributed chains, they coalesce into one large cluster, with a free energy minimum around the clustered state (Fig. S3). So, the sequence of clustering events seems to play an important role. Under same parameter combinations (Es = 5kT, Ens = 0.5kT), two clusters, each carrying 200 chains, encounter an arrested state (Fig. 2C) while randomly distributed 400 proteins can bypass such a state to form one large cluster (Fig. S3). Hence, we conclude that the stronger Es mediated arrested configuration signifies a non-equilibrium metastable state.

2. Kinetic arrest stems from long-living inter-sticker bonds

Next, we elucidate the physical properties of the individual clusters that can be mapped into their fusion tendencies (Fig. 3). As shown in Figs. 2B and 2C, we consider four parameter combinations – two of them undergo fusion (C1 and C2, labelled in green) and the rest two display an arrest (C3 and C4, labelled in red). Density of clusters correlates with the strength of non-specific interactions between spacers , Ens (Fig. 3A) and does not predict the fusion outcome. In fact, clusters with similar density (C2 and C4) show very different fusion behavior. When we look at the

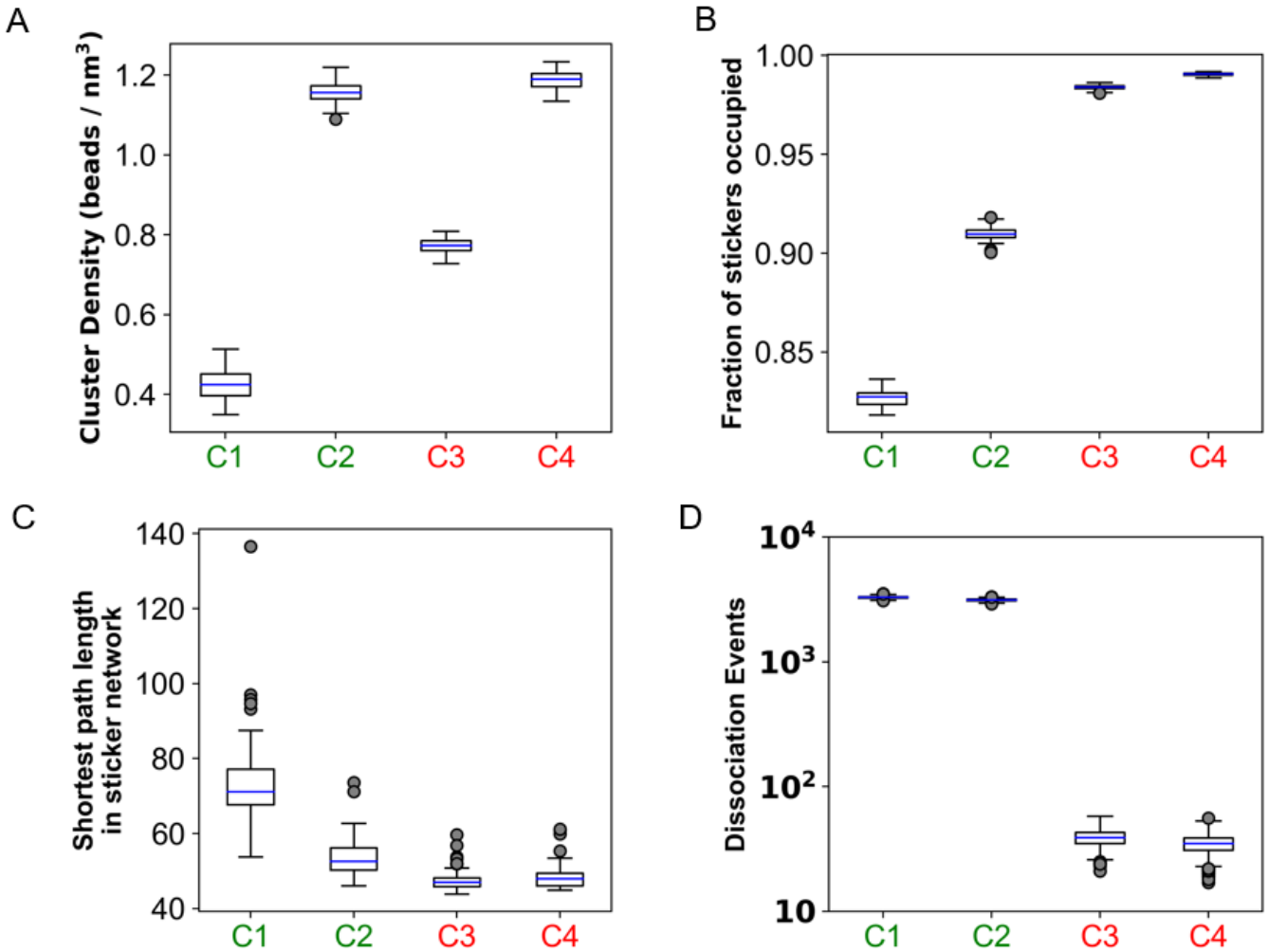


Figure 3: Mapping fusion behavior to physical properties of the clusters. Parameter combinations (Es, Ens) are divided into two categories: C1 (3kT, 0.3kT) and C2 (3kT, 0.5kT) are labelled in green as they undergo fusion; C3 (5kT, 0.3kT) and C4 (5kT, 0.5kT) are labelled in red which do not fuse. (A) Density of clusters under four parameter combinations. $\mathrm{Density} = \frac{\mathrm{N_{beads}}}{\frac{4}{3}*\pi*\mathrm{R_g^3}}$ where $\mathrm{R_g}$ is radius of gyration of the cluster and $\mathrm{N_{beads}}$ is the total number of beads (stickers + spacers) present in the cluster. (B) Fractional occupancy of stickers that indicates, on average, what fraction of the total sticker population are bonded. (C) Shortest path length (also called Geodesic length) distribution of the sticker-spacer network. Each network has 5000 nodes (beads). Starting from one node, the shortest topological path to all other nodes are computed. Hence, the path length is in the unit of bead count. (D) Dissociation events refer to how many times an inter-sticker bond is broken in between two successive observation points. Events are plotted in log scale. To derive the distributions (A – D), for each condition, we run 10 stochastic trials and sample 10 timepoints from each trial; hence each distribution is collected over 100 snapshots. To display the distribution, we used standard boxplot representation or “five-number summary” consisting of the minimum, the maximum, the sample median, and the first and third quartiles.

degree of sticker saturation within the clusters (Fig. 3B), C3 and C4 exhaust almost 100% of stickers, that is, almost all the stickers are bonded with their complementary sites. The less dense cluster (C1) has 16% stickers open while the denser cluster (C2) saturates around 92% stickers. This data also shows that, for identical Es, Ens can alter the sticker saturation by changing the density of the system. Importantly, 92% sticker saturation does not prevent cluster fusion.

When we analyze the connectivity within the polymer network (Fig. 3C), except C1, all other clusters have similar topological distance (defined in Method) distributions. More open stickers in C1 manifests in a wider distribution of network paths. However, we notice that the well-connected network in C2 does not prevent the fusion event. The radial distributions (Fig. S4) of fully saturated chains within clusters do not yield any significant difference, although there are few unsaturated chains towards the cluster periphery for C1 and C2. This suggests that, for C1 and C2, there are a few peripheral chains that are not fully saturated, in contrast to C3 and C4 where peripheral chains are fully saturated.

While mapping cluster property (static and dynamic) to their fusibility, the most striking correlation we observe upon examining the rate of bond dissociation (Fig. 3D). The rate of bond dissociations implies the extent of dynamic re-organization within the cluster. Stronger Es (5kT) makes the bonds long-lived (C3 and C4) whereas weaker Es (3kT) keeps a high turnover of bonds (C1 and C2) that allow the cluster to retain its dynamic nature. Overall, strong specific interactions trigger a kinetic arrest where the fully saturated stickers form a compact network with *long-lived bonds*.

3. Separation of sticker-spacer energetics yields re-entrant phase behavior

Next, we explore the interplay of Es and Ens in determining the cluster fusion behavior. Specifically, we seek to understand whether there is any difference in fusion of clusters composed of heteropolymers (sticker-spacer) vs homopolymers (spacer only). By tuning the difference between Es and Ens, we can systematically create a spectrum of energy separation between stickers and spacers.

Fig. 4A displays the phase diagram where extent of fusion is plotted against a wider combination of Es and Ens. First, we notice that the lower left corner of the diagram has no cluster signifying a zone below the phase transition limit. Next dramatic difference we observe is the trend of fusion along the Ens dimension for a lower Es (left regime) vs higher Es (right regime). If we consider the clusters consisting of pure homopolymers (Es = 0, leftmost column), fusion happens readily for a range of intermediate Ens strengths. When we ramp up Ens to very high values, fusibility of

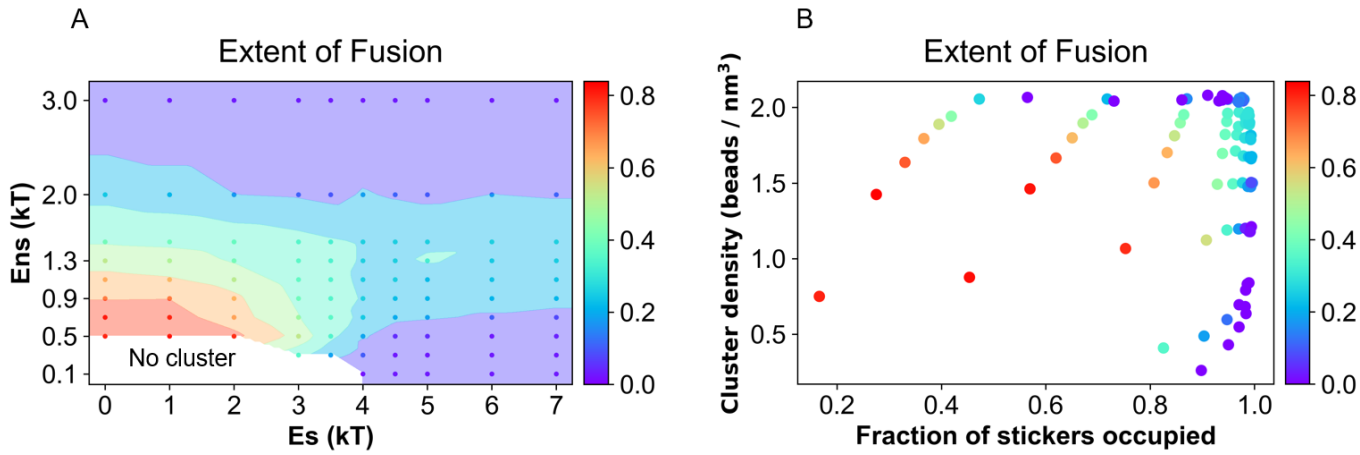


Figure 4: Phase diagram of cluster fusion behaviors. (A) Fusion of clusters as a function of a wide range of specific (Es) and non-specific (Ens) interactions. Dotted lines indicate the points where the simulations are carried out. These discrete points are then extrapolated to generate the continuous phase space encoded by the color scheme. (B) The fusion tendencies are mapped to cluster density and extent of sticker saturation. The color bar displays the extent of fusion which is defined as $1 - R_{dist}^{tf}$, where R_{dist}^{tf} is the relative inter-cluster distance at the final timeframe of the simulation. Each point is an average over 5 stochastic trials.

clusters gradually goes down, leading to a complete arrest at Ens = 3kT. Inspection of the density reveals that these clusters increasingly become denser (Fig. S5A) and beyond some threshold value, they cannot penetrate. This density mediated kinetic arrest for homopolymers is a fundamentally different mechanism than the sticker saturation mediated kinetic arrest observed for sticker-spacer polymers (Fig. 2). We next turn our attention to the high Es regime where intra-cluster stickers are more saturated (Fig. S5B). For weaker Ens values (lower right section), these systems readily undergo kinetic arrest. But we see an interesting trend upon increasing the Ens values. The kinetic arrest is partially rescued for intermediate Ens values (middle right). But *clusters re-enter again the arrested state for large Ens values*. When we map this fusion behavior into the sticker saturation vs. cluster density space, the trend becomes clearer (Fig. 4B). Clusters with very high densities and nearly 100% saturated stickers show minimal to zero fusion. Lower to intermediate values of these two parameters permit fusion.

4. Interplay of energy and entropy drives the fusion of clusters

We have shown that loss of dynamic bond organization prevents fusion of two clusters composed of sticker-spacer polymers. For homopolymers, it is the cluster density that dictates the fate of arrest. But what is the physical force that drives the fusion of clusters? To answer that question, we consider the fusion trajectory for a heteropolymer (Es = 3kT, Ens = 0.3kT) and a homopolymer system (Es = 0kT, Ens = 0.9kT) (Fig. 5). For heteropolymers, firstly we note that the potential energy does not change much as the fusion proceeds (Fig.5A). There is a slight downward trend in potential energy which is steered by the non-specific energy (Fig. S6A) which stems from increased pairwise contacts upon fusion. Most importantly, the total number of bonds in the system does not change during the course of fusion (Fig. 5B); the number of inter-cluster bonds gradually rises at the expense of intra-cluster bonds. The quantitative effect of this bond exchange can be captured by an information entropy, which is called bond entropy hereafter (Fig. 5C). This parameter describes how the total number of bonds are distributed across the two-cluster system. As the clusters exchange bonds, bond entropy keeps rising. For a fully fused state, entropy has a maximum value of 1.1 (described in Methods). Since the change in potential energy is minimal and total number of bonds remains constant, we can think of the fusion process as an entropy driven one. Although the bond entropy formulated here does not capture all components of the thermodynamic entropy (translation, conformations etc. of polymeric chains), it persistently goes up as the fusion continues to happen. Now, when there are not enough inter-cluster bonding opportunities due to intra-cluster sticker saturation, fusion comes to a halt due to the lack of bond exchange (Fig. S7). Analysis of neighbor exchange (Fig. S8) reveals that fusion happens when

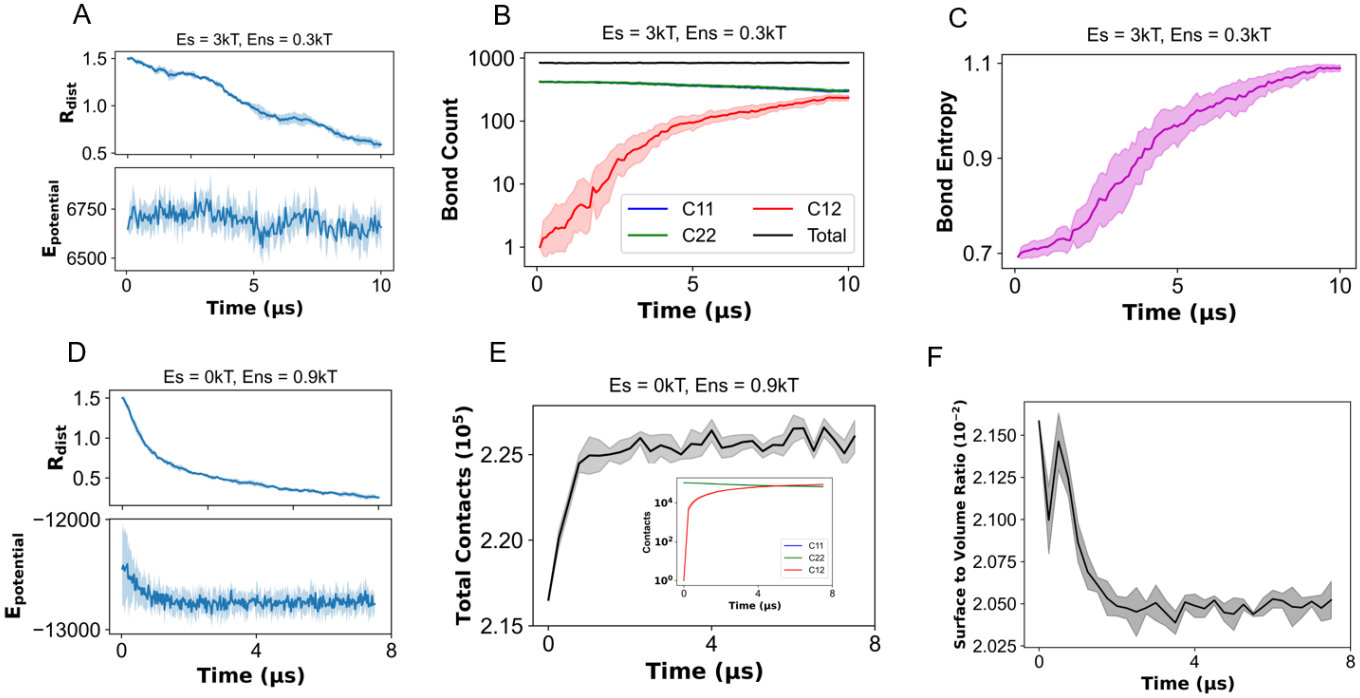


Figure 5: Driving forces of cluster fusion. First three panels (A,B,C) represent a heteropolymer system (Es = 3kT, Ens = 0.3kT), while last three panels (D,E,F) represent a homopolymer system (Es = 0kT, Ens = 0.9kT). (A, D) Potential energy profile (unit: kcal/mol) along with the cluster fusion profile. R_{dist} is defined in Fig. 2A. (B, E) Number of bonds and contacts respectively. “Bond” is a link (spring) between two heterotypic stickers while any two beads within the Lennard-Jones potential cut-off distance (25 Å) are counted to have one “contact”. C11, C22 and C12 stand for intra-cluster-1, intra-cluster-2 and inter-cluster, respectively. “Total” indicates the entire system. (C) Measure of bond entropy for the entire system. (F) Surface to volume ratio of the two-cluster system as the fusion proceeds. Each analysis is an average of 5 stochastic trials.

inter-cluster neighbors, on average, exceed one. Chains from cluster-1 need to “grab” more than one chain from cluster-2 to initiate fusing into one another. For a case with kinetic arrest, two clusters cannot engage more than one chain at a time (Fig. S8B).

Next, we turn our attention to homopolymers (Fig. 5D). First, we see a significant drop in potential energy as the clusters begin to fuse; however, after the initial drop, potential energy plateaus even though the fusion continues to happen (upper panel, Fig. 5D). The potential energy profile is again shaped by the pairwise contact energy (Fig. S6B). This drop in potential energy is accompanied by a rise in contact counts (Fig. 5E). When two clusters begin to merge, a fraction of the surface beads establish new contacts. Since Ens is relatively high (0.9kT), this increase in contacts favors a energy driven fusion. Movement of beads from surface to interior manifests as a reduction of surface tension. Fig. 5F shows the gradual drop in surface to volume ratio as fusion starts to happen. We also notice that the potential energy, total contact counts and surface-to-volume ratio all stabilize around 2 μs, but the fusion parameter (R_{dist}) continues to go down. This suggests an entropic contribution towards the later stage of fusion. We can, again, invoke the concept of exchange entropy to rationalize this phenomenon. The steady rise of inter-cluster contacts (Fig. 5E, inset) indicates that the exchange of beads drives the process towards completion. Hence, we observe a combined effect of energy and entropy. The initial energy drop clearly shows an energy inceptive in initiating the fusion and the contact exchange between the clusters later serves as additional incentive to drive two clusters into a fully fused state. When Ens is very high, clusters become too dense (solid-like). In this scenario, beads from one cluster can no longer flow into the other, causing a kinetic arrest.

Discussion

In this work, we identified two distinct physical mechanisms that may potentially trigger kinetic arrest of biomolecular condensates. Heteropolymers may undergo a sticker saturation mediated arrest while homopolymers need to cross a threshold density to encounter arrest.

Condensate forming proteins typically conform to a sticker-spacer architecture. There is a separation of energy scale (Fig.6A) in these heteropolymers where stickers drive the energetics by creating transient bonds and spacers govern the solubility of the inter-sticker network by tuning the flexibility of the linker regions. For such systems, we have shown that the ability of clusters to exchange stickers decides the fusion propensity. As more and more intra-cluster stickers are saturated, purely due to kinetic considerations, the ability of the two clusters to coalesce gradually

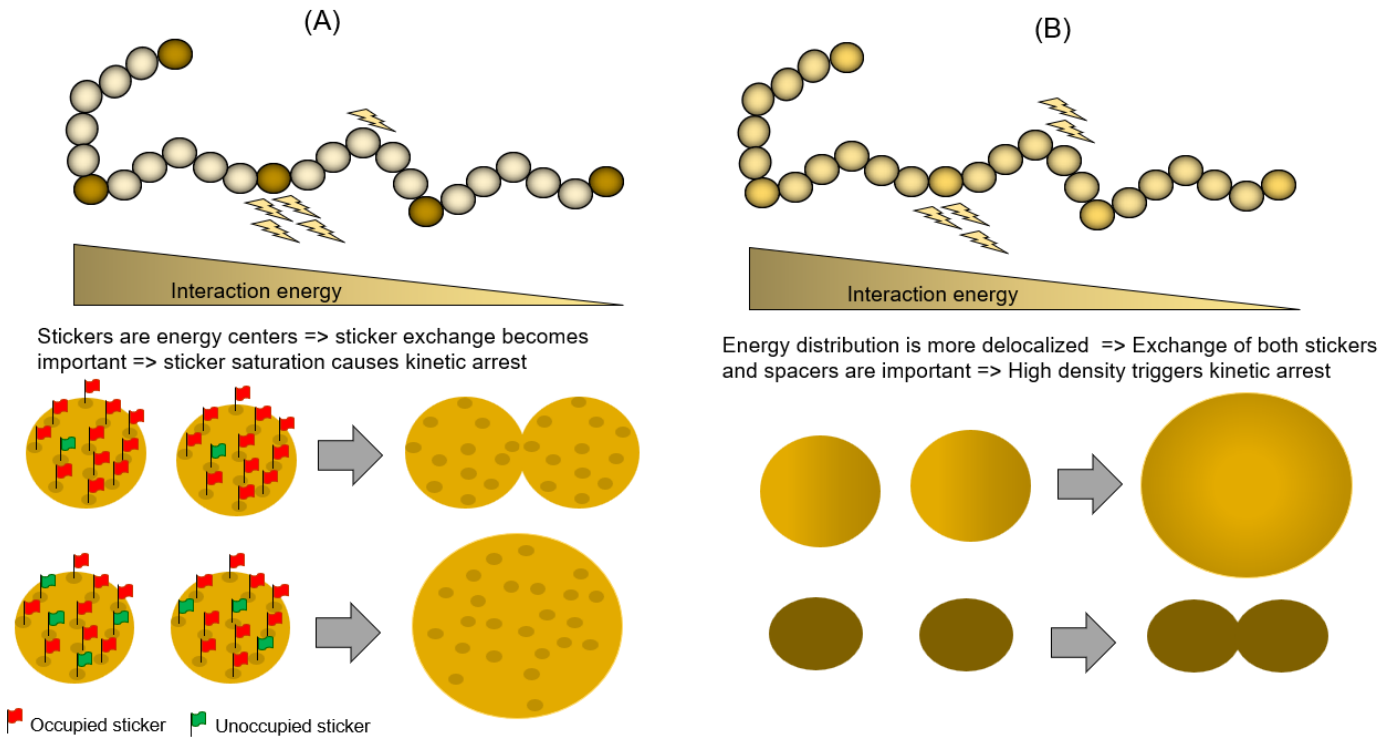


Figure 6: Graphical summary of two distinct mechanisms of kinetic arrest. (A) Energetic contribution from stickers and spacers are widely different. Since stickers are energy centers, exchange of stickers drives the fusion of two clusters. If stickers are mostly saturated within a cluster (middle panel), exchange becomes kinetically unfavorable, and the clusters don't fuse. With more stickers available (lower panel), clusters will readily merge. (B) Energy is evenly distributed between stickers and spacers. System benefits from both the spacer-exchange mediated entropic gain and surface energy minimization. Sticker saturation mediated kinetic arrest is alleviated. When density of such clusters becomes too high, they again cannot fuse due to viscosity, leading to a kinetic arrest.

decreases. When 100% stickers are saturated or the effective cluster valency is fully exhausted, they can't fuse when come into contact. In terms of driving force, since stickers are energy centers, exchangeability of stickers are sole determinant of fusion. Hence, for a fully saturated cluster, fusion is not possible as the long-lived inter-sticker bonds do not open up readily to promote exchange entropy. It is worthy to highlight that the single valency nature of individual stickers is the causal factor underlying the kinetic arrest at low density of clusters

Biological systems may exploit such architecture to reversibly tune the "fusibility" of condensates. Multiple receptor tyrosine kinase systems that are crucial in cell signaling, undergo phase separation . Tyrosine in phosphorylated (pY) form can recruit SH2 domain containing proteins. The phosphorylation level of tyrosine residues is balanced by a competing effect of kinases and phosphatases. Saturation of all pY stickers by intra-cluster SH2 domains could be a potential strategy to trigger an arrested state and stabilize a multi-cluster configuration. This state can be reversed by ramping up dephosphorylation and creating more unoccupied stickers that are again chemically modifiable.

On the other hand, for homopolymers, energy is equally distributed amongst all the beads (Fig. 6B). So, exchange of each bead is equally "desirable". These systems fuse readily at a wider range of Ens (Fig. 4). However, with higher density, fusibility goes down gradually which eventually leads to an arrest at a threshold density. Both energy and entropy play important roles in driving such fusions. Since non-specific interactions are not restricted by any "single valency" effect, multiple contacts allow two clusters to exchange beads easily. The beginning of fusion transfers multiple beads from surface to interior; such surface energy minimization is a generic fusion mechanism of two liquid droplets. For homopolymeric condensates, the surface energy initiates the merger and the entropic boost from contact exchange drives the process towards the finish line.

For biological systems, we can expect a full spectrum where the energy separation between stickers and spacers can be controlled gradually. When the gap in energetic contributions between stickers and spacers reduces, exchange of both stickers and spacers become important and sticker saturation alone can't trigger the arrest. Conversely, stickers alone would control fusibility if they were the major energetic contributors. The interplay between sticker-saturation and density mediated kinetic arrest might give rise to a re-entrant phase behavior. Again, by reversibly tuning the energy gap, potentially through post-translational modifications, cellular systems can switch between arrest vs no-arrest state.

Work presented here highlights a clear connection between the structural details of biopolymers and the material properties of mesoscopic structures that they assemble into. Studying viscoelastic properties of condensates has been an active field of investigation due to its implications in biological functions. There is a growing interest in synthetic condensates with desirable material properties for therapeutic purposes. Our work reveals that material properties of condensates can be tuned by creating sticker-spacer architecture with different degrees of energy separation. Finally, the rich physics of associative heteropolymers may inspire to develop new theories relevant to soft and condensed matter physics.

Acknowledgement

We are grateful to Junlang Liu and Srivastav (Aditya) Ranganathan for fruitful discussions. A.C. greatly appreciates all the help with LAMMPS software and metadynamics simulations that Aditya provided. This work was supported by NIH (Grant 5R35GM139571).

Author Contributions

A.C. and E.I.S. designed the research, A.C. performed simulations, A.C. and E.I.S. analyzed the data and wrote the paper, E.I.S. secured the funding.

References

1. Banani, S.F., et al., *Biomolecular condensates: organizers of cellular biochemistry.* Nat Rev Mol Cell Biol, 2017. **18**(5): p. 285-298.
2. Shin, Y. and C.P. Brangwynne, *Liquid phase condensation in cell physiology and disease.* Science, 2017. **357**(6357).
3. Lyon, A.S., W.B. Peeples, and M.K. Rosen, *A framework for understanding the functions of biomolecular condensates across scales.* Nat Rev Mol Cell Biol, 2021. **22**(3): p. 215-235.
4. Mathieu, C., R.V. Pappu, and J.P. Taylor, *Beyond aggregation: Pathological phase transitions in neurodegenerative disease.* Science, 2020. **370**(6512): p. 56.
5. Wang, B., et al., *Liquid-liquid phase separation in human health and diseases.* Signal Transduct Target Ther, 2021. **6**(1): p. 290.
6. Choi, J.-M., A.S. Holehouse, and R.V. Pappu, *Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions.* Annual Review of Biophysics, 2020. **49**(1): p. 107-133.
7. Mittag, T. and R.V. Pappu, *A conceptual framework for understanding phase separation and addressing open questions and challenges.* Mol Cell, 2022. **82**(12): p. 2201-2214.
8. Ranganathan, S. and E.I. Shakhnovich, *Dynamic metastable long-living droplets formed by sticker-spacer proteins.* Elife, 2020. **9**.
9. Semenov, A.N. and M. Rubinstein, *Thermoreversible Gelation in Solutions of Associative Polymers. 1. Statics.* Macromolecules, 1998. **31**(4): p. 1373-1385.
10. Choi, J.-M., F. Dar, and R.V. Pappu, *LASSI: A lattice model for simulating phase transitions of multivalent proteins.* PLOS Computational Biology, 2019. **15**(10): p. e1007028.
11. Li, P., et al., *Phase transitions in the assembly of multivalent signalling proteins.* Nature, 2012. **483**(7389): p. 336-340.
12. Harmon, T.S., et al., *Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins.* Elife, 2017. **6**.
13. Martin, E.W., et al., *Valence and patterning of aromatic residues determine the phase behavior of prion-like domains.* Science, 2020. **367**(6478): p. 694-699.
14. Flory, P.J., *Thermodynamics of high polymer solutions.* The Journal of chemical physics, 1942. **10**(1): p. 51-61.
15. Huggins, M.L., *Some properties of solutions of long-chain compounds.* The Journal of Physical Chemistry, 1942. **46**(1): p. 151-158.
16. Zwicker, D., A.A. Hyman, and F. Jülicher, *Suppression of Ostwald ripening in active emulsions.* Phys Rev E Stat Nonlin Soft Matter Phys, 2015. **92**(1): p. 012317.
17. Cuylen, S., et al., *Ki-67 acts as a biological surfactant to disperse mitotic chromosomes.* Nature, 2016. **535**(7611): p. 308-312.

18. Wang, Z., et al., *Cellular proteins act as surfactants to control the interfacial behavior and function of biological condensates.* Developmental Cell, 2023. **58**(11): p. 919-932.e5.
19. Folkmann, A.W., et al., *Regulation of biomolecular condensates by interfacial protein clusters.* Science, 2021. **373**(6560): p. 1218-1224.
20. Erkamp, N.A., et al., *Spatially non-uniform condensates emerge from dynamically arrested phase separation.* Nature Communications, 2023. **14**(1): p. 684.
21. Snead, W.T., et al., *Membrane surfaces regulate assembly of ribonucleoprotein condensates.* Nat Cell Biol, 2022. **24**(4): p. 461-470.
22. Lee, D.S.W., et al., *Size distributions of intracellular condensates reflect competition between coalescence and nucleation.* Nature Physics, 2023. **19**(4): p. 586-596.
23. Jawerth, L., et al., *Protein condensates as aging Maxwell fluids.* Science, 2020. **370**(6522): p. 1317-1323.
24. Garaizar, A., et al., *Kinetic interplay between droplet maturation and coalescence modulates shape of aged protein condensates.* Scientific Reports, 2022. **12**(1): p. 4390.
25. Ranganathan, S. and E. Shakhnovich, *The physics of liquid-to-solid transitions in multi-domain protein condensates.* Biophysical Journal, 2022. **121**(14): p. 2751-2766.
26. Banani, S.F., et al., *Compositional Control of Phase-Separated Cellular Bodies.* Cell, 2016. **166**(3): p. 651-663.
27. Plimpton, S., *Fast Parallel Algorithms for Short-Range Molecular Dynamics.* Journal of Computational Physics, 1995. **117**(1): p. 1-19.
28. Thompson, A.P., et al., *LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales.* Computer Physics Communications, 2022. **271**: p. 108171.
29. Laio, A. and F.L. Gervasio, *Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science.* Reports on Progress in Physics, 2008. **71**(12): p. 126601.
30. Ranganathan, S., et al., *Phosphorylation sites are evolutionary checkpoints against liquid–solid transition in protein condensates.* Proceedings of the National Academy of Sciences, 2023. **120**(20): p. e2215828120.
31. Jewett, A.I., et al., *Moltemplate: A Tool for Coarse-Grained Modeling of Complex Biological Matter and Soft Condensed Matter Physics.* Journal of Molecular Biology, 2021. **433**(11): p. 166841.
32. Martínez, L., et al., *PACKMOL: a package for building initial configurations for molecular dynamics simulations.* J Comput Chem, 2009. **30**(13): p. 2157-64.
33. Fiorin, G., M.L. Klein, and J. Henin, *Using collective variables to drive molecular dynamics simulations.* Molecular Physics, 2013. **111**(22-23): p. 3345-3362.

34. Stukowski, A., *Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool.* Modelling and Simulation in Materials Science and Engineering, 2010. **18**(1): p. 015012.

Supplementary Figures and Movies

Movie S1 Legend

Fusion dynamics of two clusters under four combinations of energy parameters.

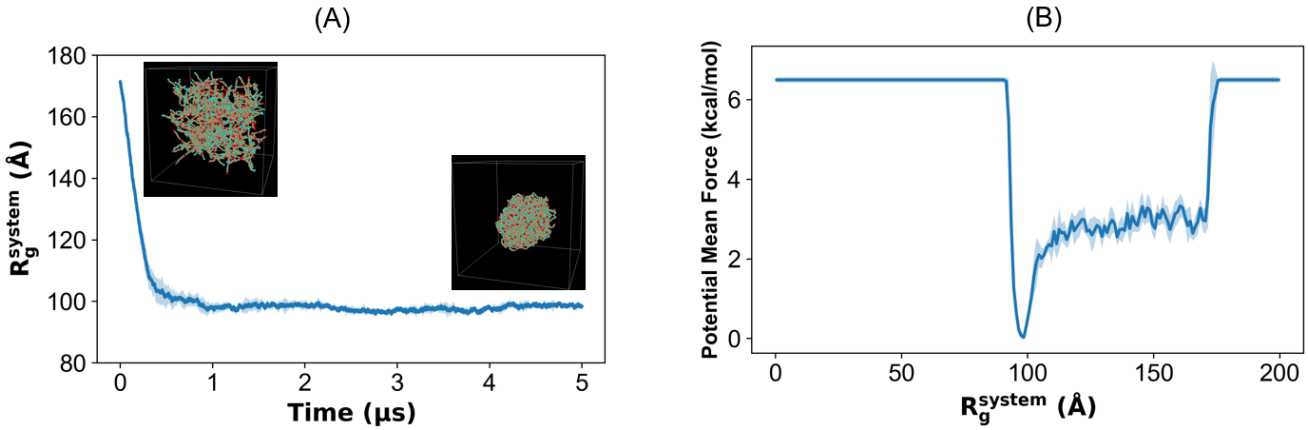


Figure S1: Coalescence dynamics of 200 chains into one large cluster. Energy parameters, Es = 3kT, Ens = 0.5kT. (A) Timecourse of the metadynamics order parameter R_g^{system}. Insets show two snapshots depicting dispersed and clustered states. (B) Free energy profile where the minimum corresponds to the fully coalesced state. Each line is an average of 5 stochastic trials. The solid line is the mean and fluctuation envelop represents the standard deviation.

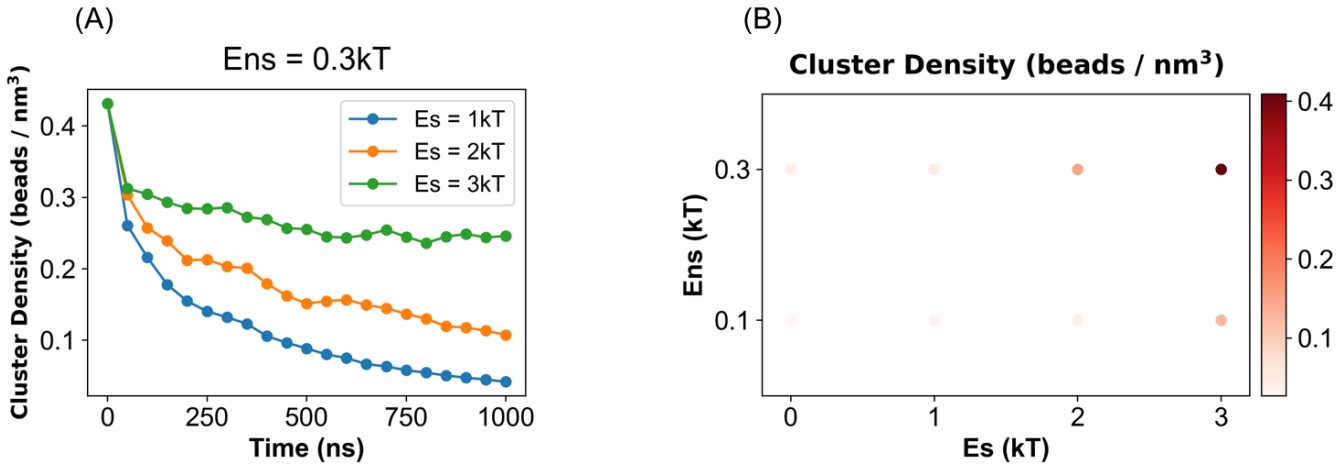


Figure S2: Quantification of phase transition boundary from the relaxation dynamics. (A) Timecourse of the cluster density, for a fixed Ens = 0.3kT, as a function of specific interaction strength (Es). (B) Phase diagram of the cluster density which is computed at the last timepoint of relaxation trajectory. $Density = \frac{N_{beads}}{\frac{4}{3}*\pi*R_g^3}$ where R_g is radius of gyration of the cluster and N_{beads} is the total number of beads (stickers + spacers) present in the cluster. Es = 3kT and Ens = 0.3kT is the lowest energy parameters combination where cluster retains its stability.

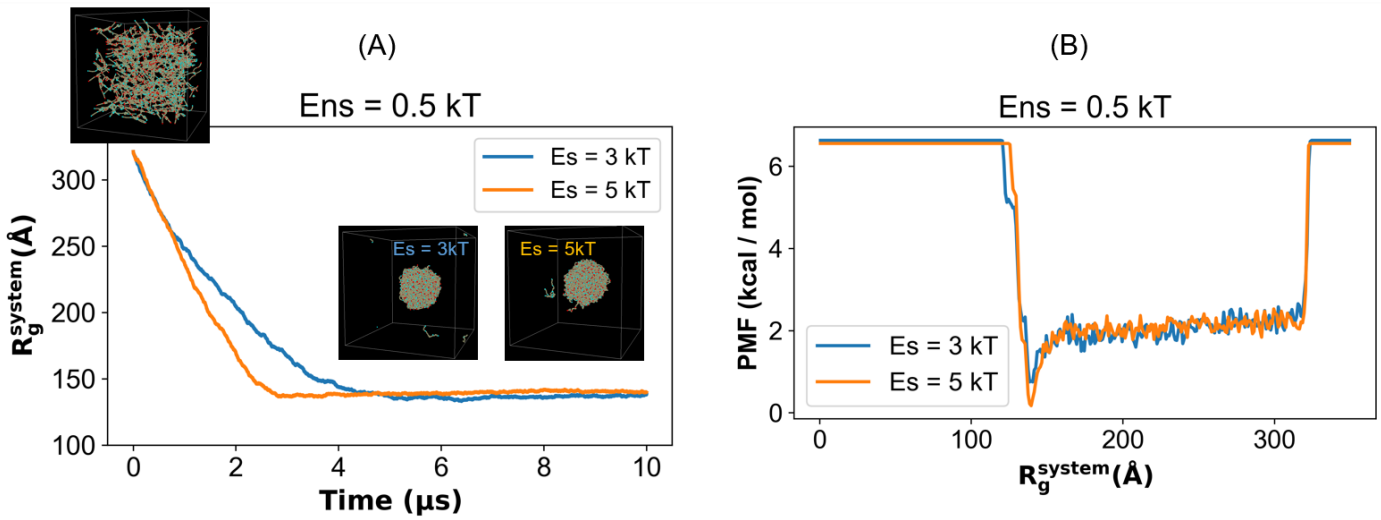


Figure S3: Coalescence dynamics of 400 chains into one large cluster starting from a uniform distribution. (A) Timecourse of the metadynamics order parameter R_g^{system}. Insets show snapshots depicting dispersed (Initial state, upper left) and clustered states (lower right). (B) Free energy profile where the minimum corresponds to the fully coalesced state. Each line is an average of 5 stochastic trials.

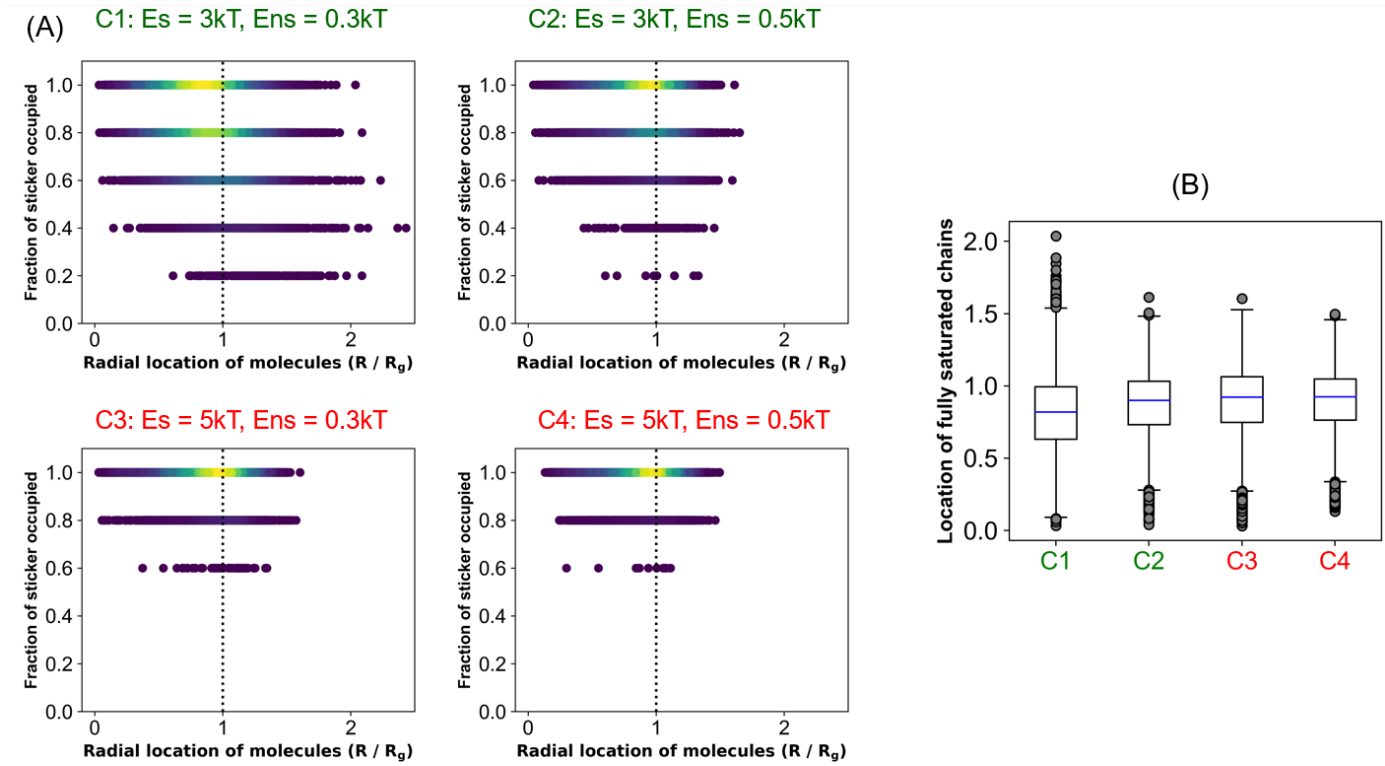


Figure S4: Spatial distribution of saturated stickers within clusters. (A) Extent of sticker saturation as a function of radial locations of chains, for four combinations of Es and Ens. The color bar represents frequency where brighter colors are more frequent compared to dark ones. We compute distance (R) of each chain center from the cluster center and normalize by cluster radius (Rg). For a perfectly spherical cluster, R ≤ Rg. Since clusters are far form a perfect sphere, $\frac{R}{R_g} \geq 1$ roughly indicates chains situated towards the surface. (B) Distribution of radial location of fully saturated chains (5 out of 5 stickers are occupied). To derive the distributions, for each condition, we run 10 stochastic trials and sample 10 timepoints from each trial; hence each distribution is collected over 100 snapshots. To display the distribution, we used standard boxplot representation or "five-number summary" consisting of the minimum, the maximum, the sample median, and the first and third quartiles.

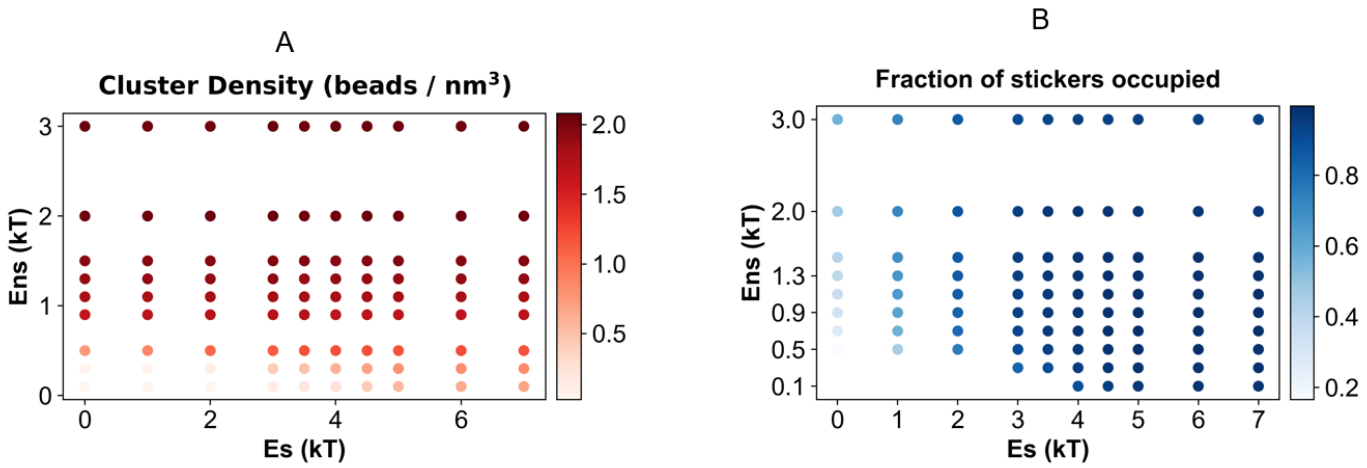


Figure S5: Quantification of the physical properties of clusters across the phase diagram. (A) Density of clusters (B) Extent of sticker saturation as a function of Es and Ens.

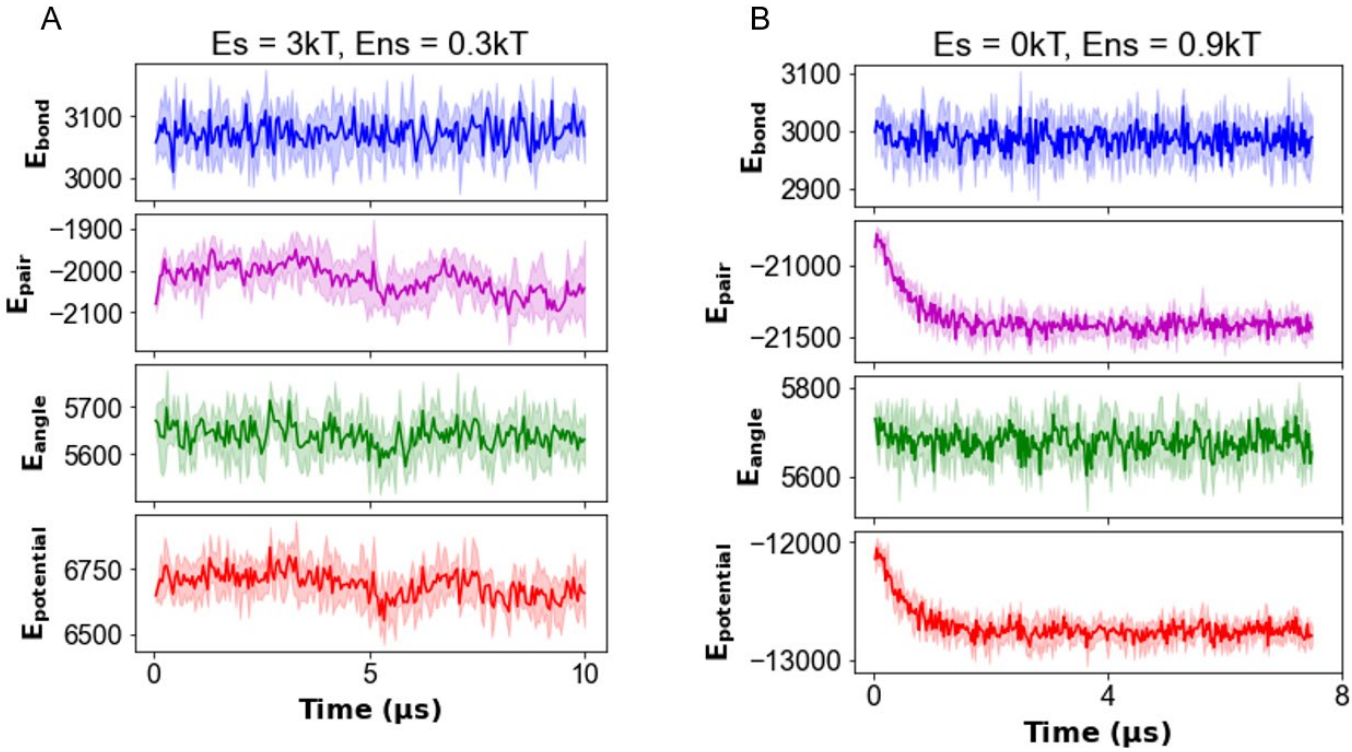


Figure S6: Energy profiles during fusion of two clusters made of (A) Heteropolymers (B) Homopolymers. E_{bond} includes all the bonds (permanent and breakable) present in the system. E_{pair} refers to the pairwise energy. E_{angle} is angular energy. $E_{potential} = E_{bond} + E_{pair} + E_{angle}$. Energy unit is kcal/mol. Each line is an average of 5 stochastic trials. The solid line is the mean and fluctuation envelop represents the standard deviation.

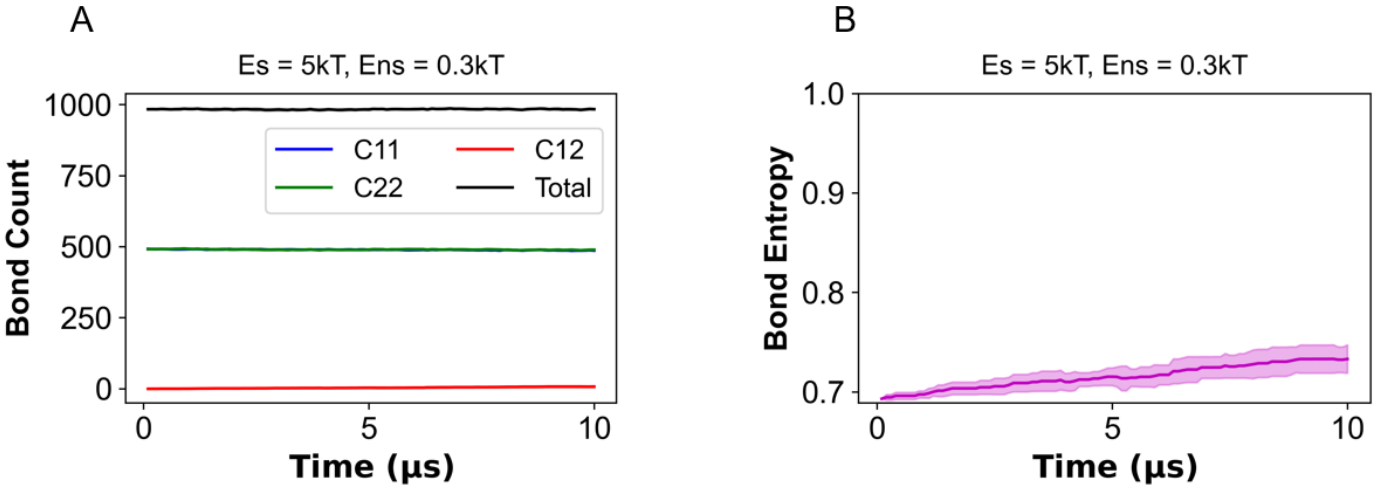


Figure S7: Lack of bond exchange triggers kinetic arrest of sticker-saturated clusters. (A) Number of bonds. C11, C22 and C12 stand for intra-cluster-1, intra-cluster-2 and inter-cluster, respectively. “Total” indicates the entire system. (B) Measure of bond entropy for the entire system.

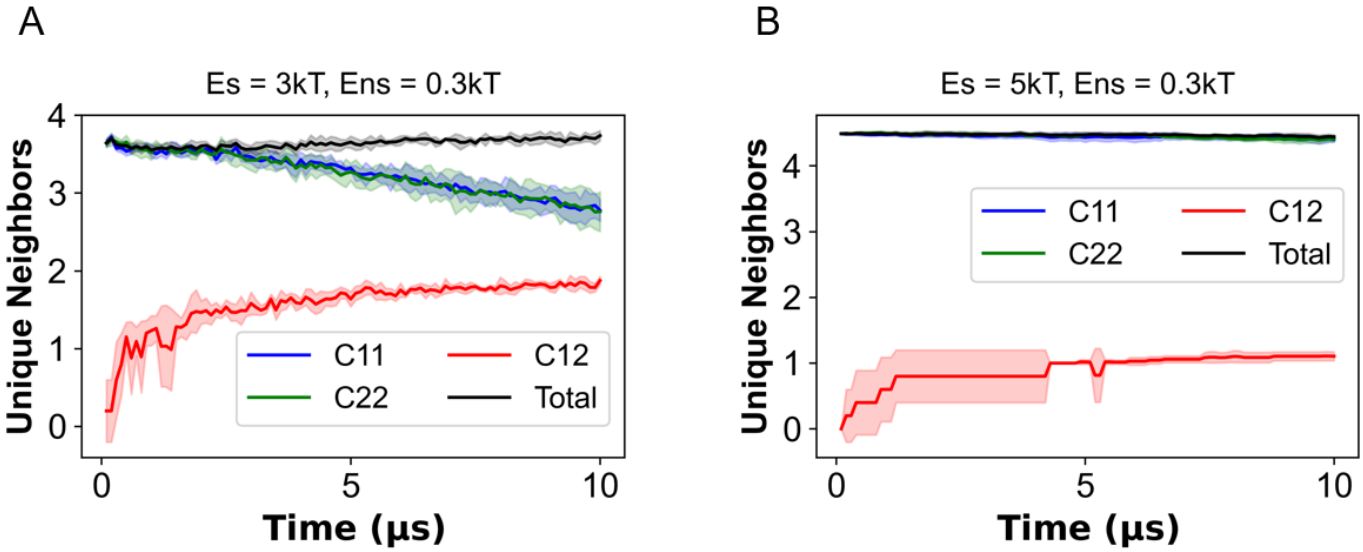


Figure S8: Neighbor exchange during cluster fusion. Two energy combinations are shown when clusters (A) fuse and (B) do not fuse. C11, C22 and C12 stand for intra-cluster-1, intra-cluster-2 and inter-cluster, respectively. “Total” indicates the entire system.