Effect of RNA on Morphology and Dynamics of Membraneless Organelles

Srivastav Ranganathan, Eugene Shakhnovich
Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, United States

Abstract
Membraneless organelles (MLOs) are spatiotemporally regulated structures that concentrate multivalent proteins or RNA, often in response to stress. The proteins enriched within MLOs are often classified as high-valency “scaffolds” or low-valency “clients”, with the former being associated with a phase-separation promoting role. In this study, we employ a minimal model for P-body components, with a defined protein–protein interaction network, to study their phase separation at biologically realistic low protein concentrations. Without RNA, multivalent proteins can assemble into solid-like clusters only in the regime of high concentration and stable interactions. RNA molecules promote cluster formation in an RNA-length-dependent manner, even in the regime of weak interactions and low protein volume fraction. Our simulations reveal that long RNA chains act as superscaffolds that stabilize large RNA–protein clusters by recruiting low-valency proteins within them while also ensuring functional “liquid-like” turnover of components. Our results suggest that RNA-mediated phase separation could be a plausible mechanism for spatiotemporally regulated phase separation in the cell.

Graphical Abstract

INTRODUCTION
The cellular milieu is heterogeneous, and diverse intermolecular interactions can drive phase separation of biological macromolecules leading to formation of spatially segregated...
membraneless compartments. Membraneless organelles (MLOs) are involved in several important cellular functions including ribosome synthesis and spindle formation during cell division. A common feature of these membrane-free “organelles” is liquid–liquid phase separation (LLPS). One such class of biomacromolecular assemblies is that of RNA–protein (RNP) condensates which are present within both the nucleus and the cytoplasmic space. Examples of cytoplasmic RNP assemblies include P bodies, germ granules, and stress granules (SGs). Macromolecular assemblies formed by LLPS regulate gene expression by colocalizing RNA and translation-related proteins within them. Other functions attributed to RNP condensates include (i) sequestration of mRNAs stalled during translation in response to stress and (ii) segregation of the protein machinery involved in RNA processing and transport. Experimental studies involving SGs identified the role of the granules, elucidated their constituents, and partially characterized structural dynamics of SGs. While several experimental studies have shown the ability of RNA-binding proteins to assemble into droplets, these observations relate to overexpression of the protein of interest in vivo or at large saturation concentrations in vitro. This raises a critical question—is phase separation a relevant mechanism for assembly of MLOs in the cell at physiologically relevant, low protein concentrations? Are there other spatiotemporally regulated factors that are essential for their phase separation at low abundances of phase-separating proteins? In this context, the potential role of RNA as a LLPS promoter becomes critical. While RNA molecules are found to be an integral component of several condensates (mRNAs in SGs and P-bodies, lncRNAs such as NEAT1 in para-speckles), their potential in regulating phase behavior and the dynamics of these droplets is not systematically understood. Boeynaems et al. in their study used short homotypic RNA (<50 nucleotides) and PR-repeat peptides to show that RNA can dictate material properties of droplets in an RNA-composition-dependent manner. Maharana et al. demonstrated the ability of RNA to promote (at low RNA:protein ratio) or prevent droplet formation (high RNA:protein ratio) by prion-like RNA-binding proteins. These studies typically focus on homotypic self-assembly of RNP–protein components in the presence/absence of RNA and not on multicomponent phase separation.

Computational studies have been extensively employed to address equilibrium aspects of phase separation. Typically, the focus of these studies is to understand how factors such as valency and strength of interactions can tune phase behavior of multivalent biomolecules. Equilibrium studies by Harmon et al. analyzed the physical determinants of phase separation by using on-lattice spacer-sticker proteins. Off-lattice approaches include use of patchy particles to model multivalent biomolecules to describe how properties of scaffold (large valency proteins) and client (smaller valency) proteins can define phase behavior. In particular, the recent work by Espinosa et al. employs the patchy-particle approach to establish the importance of molecular connectivity in the phase separation of scaffold–client mixtures. A common assumption in existing computational approaches is that the distinguishing feature of scaffolds and clients is their valency alone and that all adhesive sites can participate in attractive interactions with any other free adhesive site. While a valuable first step in addressing a very complex problem, these studies fail to capture the complex reality of protein–protein interaction networks in the cell where each...
adhesive site often has a specific binding partner. Also, the primary focus of previous studies has been on studying the phase behavior of a mixture of multivalent proteins.\textsuperscript{27-29} Coarse-grained computational models have also been extensively used by Thirumalai and co-workers in their seminal works focusing on the thermodynamics of RNA folding.\textsuperscript{30-32} However, the role of physicochemical properties of RNA in modulating phase separation is less well studied computationally.

In this study, we aim to understand the mechanism by which RNA molecules could promote phase separation of a multicomponent protein mixture in a length-dependent manner. Here, we employ a two-pronged approach wherein the interacting protein components not only vary in terms of their valency but also have a defined interaction network, mimicking protein–protein interactions within a known RNP condensate, P-body. We use a hard-sphere patchy particle representation of P-body RNA-binding proteins to study the effect of RNA length on their phase behavior at low RNA/protein ratios while modeling the RNA molecules as semiflexible polymer chains. Using Langevin dynamics (LD) simulations, we probe the ability of this multicomponent protein mixture to form large assemblies at varying protein concentrations and interaction strengths. In addition, we specifically focus on the mechanism by which RNA molecules can facilitate droplet formation in the regime of low protein concentrations and weak interaction strengths by promoting interactions involving low-valency components in the RNP clusters. Our results suggest that long RNA chains can enable assembly of large clusters even at low concentration of protein scaffolds and for weak inter-protein interactions.

**MODEL**

The primary focus of this study is to elucidate how RNA molecules can influence the assembly of a multicomponent mixture of multivalent proteins into protein-rich clusters. We study how long RNA molecules (modeled as semiflexible homopolymers) at low RNA:protein ratios influence the assembly of multivalent proteins (modeled as hard spheres with interaction patches\textsuperscript{28}) into dense clusters using Langevin dynamics simulations. Below, we discuss the protein–protein interaction network modeled in this study and the simulation method in detail.

**Modeling the Protein–Protein Interaction Network.**

Membraneless RNA–protein condensates are made of a complex network of RNA-binding proteins and RNA. One such example of an RNA–protein condensate are the P-bodies or processing bodies that are associated with mRNA processing.\textsuperscript{6,34} P bodies are broadly made up of two classes of proteins: the high-valency core proteins (or scaffolds) and low-valency noncore proteins (or clients).\textsuperscript{28,33,34} Unlike *in vitro* experimental studies with engineered multivalent proteins, the adhesive sites within P-body scaffolds and clients bind uniquely to a specific interaction site on another component protein.\textsuperscript{33,34}

In Figure 1A, we represent some components of this interaction network schematically. In our minimal model the P-body contains particles of 16 types, modeling its known components such as Dcp2, Edc3, Pat1, Xrn1, Lsm1, Upf1, Dhh1 and RNA\textsuperscript{34,35} (Table 1 and Figure 1). Each of these particles has a finite valency, defined by the total number of
adhesive sites ("patches") on the particle (shown as side beads in the schematic Figure 1A). Each adhesive site has a unique identity wherein it can only interact with a binding site specific to it on another protein. Therefore, each RNA binding protein that is part of the P-body has a fixed valency with respect to every other protein type. Here, valency refers to the maximum number of adhesive interactions of any particular type that a protein can get involved in. The particle types with large valencies (Dcp2, Edc3, and Pat1 in Figure 1) are termed core proteins, while the low-valency components (Pat1, Xm1, Lsm1, and Upf1) are termed noncore proteins. Additionally, each protein component also harbors an RNA-binding domain that can engage in one interaction with an RNA molecule.

**Patchy-Particle LD Simulations.**

To probe how the RNA-binding proteins assemble into dense clusters in an RNA-dependent manner, we first model these proteins as hard spheres with varying number of adhesive interaction patches on their surface (with interaction preferences defined in Figure 1). Each of these interaction patches can engage in a maximum of one adhesive interaction with a complementary patch on another patchy-particle via a continuous square-well potential (see the Methods section). The RNA chain, on the other hand, is modeled as a semiflexible homopolymer chain, with each bead of the semiflexible polymer representing a nucleotide. In our patchy particle LD simulations, we vary the length of the RNA chain and the RNA:protein ratios to study their role in promoting protein–protein interactions. However, the higher resolution at which the RNA is modeled in the patchy particle LD studies makes simulations in the limit of large RNA:protein concentrations computationally expensive.

**METHODS**

**Langevin Dynamics Simulations.**

**Force Field.**—The polymer chains in the box are modeled by using the following interactions. Adjacent beads are connected by a simple harmonic spring with a spring constant ($k_s$) of 2 kcal/Å$^2$ and an equilibrium bond length ($r_0$) of 4 Å. The potential energy function describing the connection between adjacent coarse-grained amino acid beads is given by

$$E_{\text{stretching}} = k_s \sum_{i=1}^{M-1} \left( |r_i - r_{i+1}| - r_0 \right)^2$$

To model electrostatic interactions, we use the standard Debye–Hückel electrostatic potential with a screening length ($\kappa$) of 1 nm and a dielectric ($D$) of 80. The Debye–Hückel potential has the form

$$E_{\text{DH}} = \frac{q_i q_j}{4\pi D r} \exp\left(-\frac{r}{\kappa}\right)$$

where $q_i$ and $q_j$ are the charges of the two interacting RNA beads. Each RNA bead carries a single negative charge in our simulations. The RNA chains in our simulation are modeled
as semiflexible polymers in which any two neighboring bonds interact via a simple cosine bending potential

\[ E_{\text{bending}} = k_{\text{bend}} \sum_{i=1}^{M-2} (1 - \cos \theta_i) \]  

(3)

where \( \theta_i \) describes the angle between the \( i \)th and \((i + 1)\)th bond while \( k_{\text{bend}} \) is the energetic cost for bending (bending stiffness of linker).

**Modeling Multivalent Proteins as Patchy Particles.**—Multivalency is a key feature of phase-separating proteins, as established by numerous experimental and computational studies. Multivalent proteins are typically modeled as spacer-sticker polymer chains on- or off-lattice. However, the higher-resolution spacer-sticker polymer chains with higher degrees of freedom are computationally demanding. However, in a scenario where the microscopic features of the self-assembling chain are not central to the problem, multivalent proteins have also been modeled as patchy particles with a defined number of adhesive interaction sites. This MD patchy model for multivalent proteins has been extensively used to study phase separation.\(^{28}\) In this model, the proteins are modeled as hard spheres of diameter \( \sigma_{\text{HS}} \) with a defined number of attractive sites or patches on their surface (Figure 4A). The central body particle is larger in diameter and is only involved in repulsive hard-sphere interactions. The attractive patches are smaller particles on the surface of this hard sphere and interact with complementary patches on another particle via an attractive potential that is defined by a continuous square-well potential

\[ E_{\text{csw}}(r) = -\frac{1}{2}\epsilon_{\text{csw}} \left[ 1 - \tanh \frac{r - r_w}{\alpha} \right] \]  

(4)

Here \( r \) is the distance between the centers of two attractive patches, and \( r_w \) is the radius of the attractive well in the square-well potential. Previous computational studies have established that a cutoff \( r_w \) of 0.12\( \sigma \) ensures a valency of one per attractive patch.\(^{28}\) The strength of attraction \( \epsilon_{\text{csw}} \) is varied in our simulations in the range of 1kT–5kT. Each multivalent protein in our Langevin dynamics simulations is therefore composed of \( M + 1 \) particles, where \( M \) is the valency of the protein being modeled. It must be noted that these patchy particle proteins are defined as a multicenter rigid body with no internal dynamics in our simulations. To achieve this, we use the rigid particle definition in LAMMPS.\(^{36}\) In addition, unlike the patchy-particle simulations by Espinosa et al.,\(^{28}\) each attractive patch in our LD simulations has a specific interaction patch on another protein which is modeled on the complex interaction networks that are typical of RNA–protein condensates such as P-bodies. In other words, an attractive patch can only interact with a subset of attractive patches on other sites. The interaction rules are described in Figure 1 and Table 1.

**LD Simulation Details.**—The dynamics of patchy-particle proteins and RNA was simulated by using the LAMMPS molecular dynamics package.\(^{36}\) In these simulations, the simulator solves for Newton’s equations of motion in the presence of a viscous drag (modeling effect of the solvent) and a Langevin thermostat (modeling random collisions...
The simulations were performed under the NVT ensemble at a temperature of 310 K. The mass of the multivalent patchy particle proteins was set to 8000 Da. The size of the rigid, hard-sphere, patchy particles was set to 20 Å. In our simulations, a viscous drag is implemented via a damping coefficient, \( \gamma = m/3\pi \eta a \). Here, \( m \) is the mass of an individual bead, \( \eta \) is the dynamic viscosity of water, and \( a \) is the size of the bead. The integration time step, \( \Delta t \), was set to 15 fs, while the viscosity of the surrounding medium was set to \( 10^{-3} \) Pa·s.

Starting Configuration for Simulations.—To generate starting configurations for the LD simulations, we used PACKMOL,\(^{38}\) a tool that optimizes the packing of a defined number of particles in a defined space (with box size decided by the volume fraction of proteins in the simulation box). For our starting configurations, the patchy particle proteins are randomly placed in the simulation box, with a constraint that no two particles are located closer than twice the cutoff distance for interactions in our simulation. The initial velocities for particles were randomized to generate independent trajectories.

RESULTS

In the Absence of RNA, Large Clusters Are Disfavored at Low Volume Fractions and Weak Interactions.

Phase-separated condensates such as stress granules and P-bodies are enriched in RNA molecules.\(^{34,35,39,40}\) Also, several of the protein components within these structures are known to possess an RNA-binding domain,\(^{33}\) suggesting that RNA–protein interactions could form the molecular basis of their biogenesis. RNA-binding proteins have also been observed to phase separate at lower threshold concentrations in the presence of RNA.\(^{41-43}\) The mechanism by which RNA molecules of varying lengths could influence the assembly of multivalent proteins into condensates, for a defined interaction network, has not been systematically explored. To understand how RNA chains could influence the phase behavior of these complex mixtures of multivalent proteins in a length-dependent manner, we modeled various P-body proteins\(^{34}\) as hard spheres with attractive patches\(^{28}\) (Figure 1). The attractive sites on the surface of central core of the protein mimic adhesion sites. The number of adhesive sites on the surface defines the valency of these components. Each adhesive interaction has a specific interaction partner that is defined by the interaction matrix represented schematically in Figure 1. The RNA molecules, on the other hand, are modeled as semiflexible polymer chains with a negatively charged back-bone that prevents homotypic interaction between RNA beads. One adhesive site on each of the 16 patchy-particle protein types is involved in RNA binding.

We first performed Langevin dynamics simulations of these hard-sphere proteins in the absence of RNA molecules in the simulation box. Two key parameters dictate the phase behavior of these proteins: the total protein volume fraction (\( \phi_{prot} \)) and the strength of the protein–protein adhesive interactions (\( \varepsilon_{sp} \)). In our simulations, all multivalent proteins that make up the multicomponent mixture are present at equal concentrations. To model the phenomenon at biologically relevant interaction strengths and bulk protein densities, we vary the interaction strengths (per adhesive interaction), \( \varepsilon_{sp} \), in the range of 1kT–5kT\(^{27}\) and
protein volume fraction ($\phi_{\text{prot}}$) in the range of 0.01–0.1. The $\phi_{\text{prot}}$ values employed in our simulations are of the same order as the estimated optimal protein volume fractions in the cell.\textsuperscript{44-46} In the absence of RNA molecules, at low $\phi_{\text{prot}}$ and weak $e_{sp}$ (<4kT in Figure 2A), the multicomponent mixture of proteins exists predominantly in the monomeric state or in the form of small clusters ($\leq$10 monomers). Large clusters ($\geq$100 molecules) are only observed in the limit of strong protein–protein interactions and/or high bulk densities (Figure 2A).

To further quantify the propensity of these multivalent proteins to assemble into larger clusters, we define a second-order parameter, $\mathcal{P}(\text{bound})$, which is the probability of finding a protein component in a cluster that is 10 monomers or more in size. In Figure 2B, we plot $\mathcal{P}(\text{bound})$ for each multivalent particle type in the simulation for different values of protein–protein interaction strength. The $\mathcal{P}(\text{bound})$ values were computed over several molecules and across 1000 different simulation snapshots. Interestingly, even for a relatively high volume fraction of 0.1, for weaker protein–protein interactions ($e_{sp}$ < 4kT), all the multivalent protein components exhibit low values of $\mathcal{P}(\text{bound})$. At higher interaction strengths (4kT and 5kT in Figure 2B), the two classes of proteins—high-valency core proteins (Dcp2, Edc3, Pat1, Lsm1, Upf1, and Dhh1) and the low-valency noncore proteins (Pop2, CCR4, Not2, Upf3, Upf2, Edc1, and Hek2)—show distinct $\mathcal{P}(\text{bound})$ values. The core proteins, at high interaction strengths, are predominantly (>50%) found within large clusters while the noncore proteins are more likely to be observed in the monomeric state (rapid shuttling between the surrounding medium and within clusters). The high bound fraction for the core proteins also gets reflected in the radial distribution function, $p(r)$, for different pairs of particles. The radial distribution function values show the probability of finding two particles of a given type at any distance, normalized by the corresponding values for a pair of ideal gas particles at the same bulk density. As evident from Figure 2C, for high volume fraction, strong protein–protein interactions and in the absence of RNA molecules, the clusters are mediated by core–core interactions. For instance, the red curve in Figure 2C ($e_{sp}$ of 5kT) for core–core contacts peaks at twice the value compared to interactions involving noncore contacts.

For strong interactions and large volume fractions, high $\mathcal{P}(\text{bound})$ for self-assembled proteins is indicative of minimal exchange of monomers between the protein-rich clusters and the bulk medium in this regime. Functional biomolecular condensates, however, are known for dynamic turnover of components between the droplet phase and the cytoplasm.\textsuperscript{19,39} Concentrations of individual proteins within the cell are often in the low nanomolar range.\textsuperscript{35,47} Similarly, the protein–protein interactions that govern biomolecular condensation are transient in nature to maintain the liquid-like nature of these structures. Therefore, the predominantly core-protein rich clusters observed in the regime of large protein concentration and extremely strong protein–protein interactions might not be relevant to the cellular milieu.\textsuperscript{48}
Long RNA Molecules Promote Clustering of Multivalent Proteins Even at Low Concentrations and Weak Protein–Protein Interactions.

LD simulations in the absence of RNA molecules (Figure 2) reveal that, except for the regime of high protein volume fractions and strong interprotein interactions, the multicomponent system does not undergo phase separation. Even in this narrow regime of large $\phi_{prot}$ and $\epsilon_{sp}$, we see core protein dominant clusters with little or no recruitment of low-valency, noncore proteins. This gives rise to some critical questions. Can RNA chains alter the phase behavior of these patchy-particle proteins in the limit of biologically realistic low-to moderate bulk densities of proteins and weak protein–protein interactions? How do noncore proteins get recruited in this regime of low protein concentrations and weak protein–protein interactions? Multivalent proteins that get recruited within biomolecular condensates are often present at low concentrations\(^{21}\) in the cell and colocalize RNA chains within them.\(^{25,33,49,50}\) Therefore, to probe the effect of RNA on the phase separation of the hard-sphere protein mixture, we performed simulations in the presence of RNA chains of varying lengths—$L_{RNA}$ of 10, 50, and 120 in Figure 3A. These simulations were performed for an $\epsilon_{sp}$ of 2.5$kT$ per pairwise interaction patch and a protein volume fraction of $\phi_{prot} = 0.033$. For short RNA chains ($L_{RNA} = 10$), we do not see an increase in cluster sizes even at large RNA concentrations, suggesting a negligible phase separation promoting effect (Figure 3A, black curve). As we increased the length of the RNA chains from 10 to 50, we see an increase in cluster sizes ($L_{clus}$) for the same number of effective RNA binding sites (Figure 3A, purple curve). This effect is even more pronounced for $L_{RNA}$ of 120 (Figure 3A, green curve), suggesting that longer RNA chains are more effective in promoting the formation of large stable clusters, even at low protein concentrations and weak protein–protein adhesions. In effect, a small number of long RNA chains are more effective in promoting self-assembly than large number of short RNA chains at the same number of RNA binding sites.

To further confirm the RNA-mediated nature of self-assembly, we track protein RNA contacts in the system as a function of time for the core and noncore proteins. As evident from Figure 3B,C, during the early phases of the simulation we observe an increase in RNA–protein contacts involving noncore proteins (green curve). Interestingly, the increase in the size of the largest cluster follows this initial increase in noncore RNA contacts, suggesting that RNA enables larger clusters by stabilizing noncore proteins within them, even at this low volume fraction ($\phi_{prot} = 0.033$) and weak interactions ($\epsilon_{sp} = 2.5kT$). This effect is more pronounced for long RNA chains (Figure 3C) where we see a sharp increase in RNA–protein contacts followed by a sharp increase in cluster sizes. Overall, these results establish that long RNA chains can mediate droplet formation in a multicomponent protein mixture even at low protein concentrations.

Long RNA Chains Enable Large Clusters by Recruiting Noncore Proteins.

While the effect of long RNA chains in promoting self-assembly is evident from the large protein–RNA clusters, the mechanism by which RNA chains promote clustering is not well understood. To gain further mechanistic insights into the phenomenon, we plot the radial distribution function ($g(r)$) for different pairs of particles (core–core, core–noncore, and noncore–noncore) in our multicomponent mixture of RNA and proteins. The $g(r)$ values in Figure 4 refer to the average number of atom pairs of any type (core–core, core–noncore,
and noncore–noncore) found at any radial distance (between \( r \) and \( r + \text{dr} \)), normalized by the equivalent values for a system of ideal gas particles of the same bulk density. A peak in \( p(r) \) shows a higher propensity for two particle types to be in contact with each other. Sharper peaks are signatures of denser, solid-like organization while more diffuse peaks indicate a liquid-like arrangement. The gpu-based implementation\(^{51}\) of the radial distribution function was used to compute the \( p(r) \) in Figure 4.

For low RNA concentrations, the only stable interactions are between core proteins with high valencies (high-valency particles in Figure 1), as evident from the sharp peak in core–core \( p(r) \) (black curve for core–core interactions in Figure 4). In contrast, the \( p(r) \) for interactions involving noncore proteins exhibit flat profiles for low RNA numbers (black curve for noncore interactions in Figure 4). This shows that in the absence of RNA, the clusters can only localize core proteins and thus are smaller in size (Figure 3A). As we introduce more RNA chains in the simulation box, the structures that result show a greater likelihood of interactions involving noncore proteins, as evident from \( p(r) \) values approaching values of 5–8 (Figure 4, purple and green curves). For example, a \( p(r) \) of 5 in Figure 4 indicates that, on average, a pair of particles is 5 times more likely to be in contact as compared to a system of noninteracting, ideal gas particles for the same bulk density.

In addition, the longer chain RNA molecules are more effective in promoting noncore interactions as compared to the shorter RNA chains (Figure 4, left vs right panel). It must, however, be noted that the \( p(r) \) profiles for noncore contacts are more diffuse, that is, liquid-like as compared to the core–core interactions that exhibit a sharper profile (Figure 4, core–core vs noncore–noncore). Sharp peaks in the radial distribution function are typical of densely packed solids, while liquids exhibit diffuse peaks. Such a disparity stems from stronger net interaction strengths for core proteins due to their higher valencies. This finding is consistent with the recent \textit{in vitro} observation of dynamically “solid-like” behavior of large-valency core proteins which exhibit slower exchange times.\(^{33}\) Therefore, the \( p(r) \) profiles suggest that the presence of long RNA chains promotes larger clusters by facilitating noncore interactions.

In addition, we computed the probability, \( P(\text{bound}) \), of finding individual particle types within large clusters (>10 monomers in size). For short RNA chains (Figure 5A), we observe little or no self-assembly, with \( P(\text{bound}) \) values not exceeding 0.1 for the component proteins in the mixture. These \( P(\text{bound}) \) values, in the presence of short RNA, are akin to those observed in the absence of RNA, at weak protein–protein interaction strengths (<4kT in Figure 2B). However, as we increase the length of the RNA chains in the simulation box, we see a dramatic increase in \( P(\text{bound}) \) values across different protein components (Figure 5B,C). Also, unlike the assemblies observed at large volume fractions and strong interprotein interactions (in the absence of RNA, \( P(\text{bound}) \to 1 \)), the proteins in RNA-mediated clusters exhibit low \( P(\text{bound}) \) values, suggesting rapid exchange between the monomeric and self-assembled states. Overall, these results suggest that RNA molecules can act as superscaffolds that promote phase separation of multivalent proteins, interacting via defined networks, in a length-dependent manner.
RNA-Mediated Clustering of Proteins Causes Compaction of RNA.

One of the key functions of RNA–protein condensates is the sequestration of mRNA molecules stalled in translation (stress granules\textsuperscript{52}) or the processing of RNA.\textsuperscript{34} Single molecule experimental studies have revealed significant conformational changes in mRNA molecules depending on its translational state and cellular localization. To further test whether the coarse-grained RNA chains in our simulations exhibit behavior that is consistent with experiments, we probe the conformational dynamics of RNA chains in our simulations for increasing volume fractions of multivalent proteins. In Figure 6A, we show the distribution of the radius of gyration of RNA chains in the simulation at equilibrium, computed over 1000 different equilibrium snapshots and multiple RNA chains in the simulation box. As we increase the volume fraction of proteins within the simulation box from 0.01 to 0.1, the distribution of RNA sizes shifts to lower values, with the peak shifting from 40 to 20 Å, as protein volume fraction approaches 0.05 or higher. In Figure 6B, this gets reflected in the sharp decline in \( R_g \) at higher volume fractions, suggesting that RNP assembly results in compaction of RNA, facilitated by RNA–protein contacts within protein-rich clusters. Interestingly, this is consistent with experimental studies which show that translationally inhibited mRNA molecules get sequestered within stress granules in a compact, circularized configuration.\textsuperscript{52}

DISCUSSION

\textit{In vitro} and \textit{in silico} studies have established the ability of several RNP condensate components to form droplets even in the absence of other condensates components.\textsuperscript{42,53} An important caveat from these studies is that the phase separation of RNA-binding proteins \textit{in vitro} and \textit{in vivo} typically occurs at concentrations that are typically higher than their normal endogenous cellular concentrations.\textsuperscript{19–21} The complex nature of the intracellular space and the multicomponent nature of the condensates necessitate a systematic exploration of nonprotein components driving the phase separation of protein–RNA clusters. RNA molecules promote phase separation of several RNP proteins at low to moderate RNA:protein ratios and are also enriched within condensates like stress granules and P-bodies.\textsuperscript{34} Yet, these model studies are often far from the cellular reality where the phase-separated state is a complex mixture of multivalent proteins, each with a different valency and defined adhesive interactions.\textsuperscript{33,34}

We employ a coarse-grained computational approach that allows a systematic bottom-up exploration of RNA-mediated assembly of RNA–protein clusters for a broad range of conditions (protein volume fraction, protein–protein interaction strength, and RNA length). This approach is similar in philosophy to the \textit{in vitro} study by Banani et al.\textsuperscript{17} where they design scaffold and client peptides based on differing repeat lengths. However, the scaffolds and clients in the Banani model only differ in valencies but do not incorporate any of the complex interaction networks that are typical of protein–protein interactions within condensates. Our phenomenological simulations are therefore an attempt to extend the understanding of a client–scaffold system to a more complex reality with defined interaction networks (Figure 1) to study the phase transitions in a multicomponent mixture. Espinosa et al. have previously used a similar coarse-grained computational approach to
study the phase behavior of a multicomponent mixtures of scaffold and client proteins\textsuperscript{28} in response to critical parameters such as pH, salt concentration, temperature, and so on. However, in a variation from the Espinosa study, the scaffold and client protein particles in our phenomenological model do not just vary in valency but also in the identities of the adhesive sites. Each adhesive site on the interacting particles in our patchy-particle LD has a unique binding partner on another protein—a coarse-grained representation of the network of specific interactions within the P-body.

A systematic exploration of the behavior of this multicomponent system for a range of protein volume fractions and protein–protein interaction strength reveals that in the absence of RNA molecules clustering of proteins is observed only at large volume fractions and strong protein–protein interactions. The protein-rich clusters that assemble in this regime of concentrations and interaction strengths are predominantly enriched in core proteins (Figure 2B, $e_{sp} > 3kT$). Strikingly, the regime that favors large clusters in the absence of RNA features long dwelling times of core proteins within these clusters. Large-valency proteins (for e.g. Patl in Figure 2B) in this concentration regime are almost entirely localized within these clusters and exhibit negligible turnover (Figure 2B, $P(\text{bound}) \rightarrow 1$ for $e_{sp} = 4kT$ and $5kT$). The noncore proteins, on the other hand, do not participate in these clusters (Figure 2B, $P(\text{bound}) \ll 1$). This raises an interesting paradox. Individual phase separating proteins are found at low concentrations in the cell. Moreover, intracellular droplets are liquid-like and exhibit fast turnover dynamics, suggesting weak, promiscuous interactions\textsuperscript{54} stabilize these assemblies. How do cells achieve phase separation in the endogenous low-concentration and weak interaction regime? Earlier studies showed that mRNA molecules lower the threshold concentrations for phase separation (at low RNA:protein ratios) of prion-like RNA-binding proteins \textit{in vitro}.\textsuperscript{26} In addition, threshold concentrations for homotypic phase separation of proteins \textit{in vitro} can be severalfold higher than their cellular concentrations.\textsuperscript{21} \textit{In vitro} reconstituted droplets are also prone to exhibit maturation into solid-like assemblies compared to the more dynamic \textit{in vivo} counterparts that assemble at lower concentrations.\textsuperscript{41}

We systematically probed whether RNA can promote phase separation of multicomponent protein mixture in the regime of low protein volume fraction and weak protein–protein interactions. Consistent with model \textit{in vitro} studies, the introduction of long-RNA molecules facilitates assembly of the multicomponent protein mixture to into large clusters of 10 particles and more even at a low protein concentration and weak protein–protein interactions. A detailed exploration of the process of assembly reveals that RNA facilitates larger clusters in this regime by stabilizing interactions involving noncore proteins, as evident from Figure 3B,C where noncore RNA contacts precede the formation of large clusters! This is also consistent with the radial distribution functions which show an increased likelihood of interactions involving noncore proteins in the presence of RNA. Interestingly, the enrichment fraction $P(\text{bound})$ shows a significant increase for all protein components. However, unlike clusters observed in the absence of RNA in the regime of large protein concentrations and strong interaction, the $P(\text{bound})$ values do not approach 1 for RNA-dependent clusters. This indicates that the clusters formed in the presence of RNA chains exhibit greater turnover of individual components. \textit{In vitro} phase separation
in the absence of RNA can therefore result in droplets with altered composition as well as intradroplet dynamics compared to their *in vivo* counterparts.

The RNA-mediated demixing exhibits a strong dependence on RNA chain length, with long RNA chains (Figure 3A) enabling larger clusters even at low protein volume fractions. These results suggest that long RNA molecules behave like high-valency superscaffolds which can act as a hub that networks core and noncore proteins. This behavior of long RNA chains is analogous to the condensate-stabilizing effect of multivalent spacer sticker ligands in a recent study by Ruff et al.\(^{55}\) Similarly, the patchy particle simulation study by Espinosa et al. suggests that high valency, promiscuous binders can efficiently promote phase separation.\(^{28}\) These studies, however, model scaffolds and client molecules as collapsed structures (patchy hard spheres).\(^{28,55}\) In the protein–protein interaction network modeled in our current work, the large valency proteins (core proteins in Table 1) cannot promote large clusters in the low endogenous concentration regime because of their specificity. RNA chains, on the other hand, due to their ability to bind to several proteins, can act as physical cross-links between proteins in the cluster, allowing them to interact with their partners in this high-density phase. In addition, the weaker strength of RNA–protein interactions (1kT) compared to protein–protein interactions can result in dynamic forming and breaking of interactions within the condensate, thereby providing a liquid-like environment. The structures that assemble in the absence of RNA (strong interprotein interactions) are also different in composition to RNA-mediated assemblies for weak protein–protein interactions. In their seminal study of stress granule RNA enrichment,\(^{22,24}\) Khong et al. show that RNA enrichment within condensates is strongly correlated to the RNA length, with long RNA chains (>1 kB) more likely to be enriched within SGs. Formation of membraneless organelles is often a response to stress, and dynamic changes in RNA metabolism can trigger the reversible assembly of these structures.\(^{56}\) Spatiotemporal regulation of RNA concentration and selective enrichment of mRNA in a length-dependent fashion therefore make RNA molecules efficient enhancers and modulators of reversible phase separation in the regime of low protein concentrations and promiscuous interactions.

**ACKNOWLEDGMENTS**

The authors thank Ranjith Padinhateeri and Mark Miller for useful discussions. This work was supported by NIH GM 068670.

**REFERENCES**

(1). Banani SF; Lee HO; Hyman AA; Rosen MK Biomolecular condensates: organizers of cellular biochemistry. Nat. Rev. Mol. Cell Biol 2017, 18, 285. [PubMed: 28225081]
(2). Hyman AA; Weber CA; Jülicher F Liquid-Liquid Phase Separation in Biology. Annu. Rev. Cell Dev. Biol 2014, 30, 39–58. [PubMed: 25288112]
(3). Brangwynne CP; Tompa P; Pappu RV Polymer physics of intracellular phase transitions. Nat. Phys 2015, 11, 899–904.
(4). Han TW; et al. Cell-free Formation of RNA Granules: Bound RNAs Identify Features and Components of Cellular Assemblies. Cell 2012, 149, 768–779. [PubMed: 22579282]
(5). Rhine K; Vidaurre V; Myong S RNA Droplets. Annu. Rev. Biophys 2020, 49, 247. [PubMed: 32040349]

*J Phys Chem B*. Author manuscript; available in PMC 2022 April 19.
(6). Standart N; Weil D P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. Trends Genet. 2018, 34, 612. [PubMed: 29908710]

(7). Wurtz JD; Lee CF Stress granule formation via [ATP] depletion-triggered phase separation. New J. Phys 2018, 20, 045008.

(8). Brangwynne CP; Mitchison TJ; Hyman AA Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc. Natl. Acad. Sci. U. S. A 2011, 108, 4334–4339. [PubMed: 21368180]

(9). Proter DSW; Parker R Principles and Properties of Stress Granules. Trends Cell Biol. 2016, 26, 668–679. [PubMed: 27289443]

(10). Berchiano MT; et al. Cajal body number and nucleolar size correlate with the cell body mass in human sensory ganglia neurons. J. Struct. Biol 2017, 158, 410–420. [PubMed: 27255332]

(11). Sabari BR; et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science (Washington, DC, U. S.) 2018, 361, eaar3958.

(12). Boijia A; et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell 2018, 175, 1842. [PubMed: 30449618]

(13). Buchan JR; Yoon J-H; Parker R Stress-specific composition, assembly and kinetics of stress granules in Saccharomyces cerevisiae. J. Cell Sci 2011, 124, 228–239. [PubMed: 21172806]

(14). Xing W; Muhlråd D; Parker R; Rosen MA quantitative inventory of yeast P body proteins reveals principles of compositional specificity. bioRxiv 2018, 489658.

(15). Anderson P; Kedersha N RNA granules: post-transcriptional and epigenetic modulators of gene expression. Nat. Rev. Mol. Cell Biol 2009, 10, 430–436. [PubMed: 19461665]

(16). Jain S; et al. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell 2016, 164, 487–498. [PubMed: 26777405]

(17). Li P; et al. Phase transitions in the assembly of multivalent signalling proteins. Nature 2012, 483, 336–340. [PubMed: 22398450]

(18). Nott TJ; et al. Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. Mol. Cell 2015, 57, 936–947. [PubMed: 25747659]

(19). Feric M; et al. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell 2016, 165, 1686–1697. [PubMed: 27212236]

(20). Patel A; et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell 2015, 162, 1066–1077. [PubMed: 26317470]

(21). McSwiggen DT; Mir M; Darzacq X; Tjian R Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. Genes Dev. 2019, 33, 1619. [PubMed: 31594803]

(22). Khong A; et al. The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. Mol. Cell 2017, 68, 808. [PubMed: 29129640]

(23). Lavut A; Raveh D Sequestration of highly expressed mRNAs in cytoplasmic granules, p-bodies, and stress granules enhances cell viability. PLoS Genet. 2012, 8, e1002527. [PubMed: 22383896]

(24). Khong A; Parker R mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction. J. Cell Biol 2018, 217 (12), 4124–4140. [PubMed: 30322972]

(25). Boeynaems S; et al. Spontaneous driving forces give rise to protein—RNA condensates with coexisting phases and complex material properties. Proc. Natl. Acad. Sci. U. S. A 2019, 116, 7889–7898. [PubMed: 30926670]

(26). Maharana S; et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science (Washington, DC, U. S.) 2018, 360, 918.

(27). Harmon TS; Holehouse AS; Rosen MK; Pappu RV Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. eLife 2017, 6, e30294. [PubMed: 29091028]

(28). Espinosa JR; et al. Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components. Proc. Natl. Acad. Sci. U. S. A 2020, 117, 13238. [PubMed: 32482873]

(29). Ghosh A; Mazarakos K; Zhou HX Three archetypical classes of macromolecular regulators of protein liquid-liquid phase separation. Proc. Natl. Acad. Sci. U. S. A 2019, 116, 19474. [PubMed: 31506351]
(30). Denesyuk NA; Thirumalai D Coarse-Grained Model for Predicting RNA Folding Thermodynamics. J. Phys. Chem. B 2013, 117, 4901–4911. [PubMed: 23527587]

(31). Hyeon C; Thirumalai D Mechanical unfolding of RNA hairpins. Proc. Natl. Acad. Sci. U. S. A 2005, 102, 6789–6794. [PubMed: 15749822]

(32). Thirumalai D; Hyeon C RNA and Protein Folding: Common Themes and Variations. Biochemistry 2005, 44, 4957–4970. [PubMed: 15794634]

(33). Xing W; Muhlrad D; Parker R; Rosen MK A quantitative inventory of yeast P body proteins reveals principles of composition and specificity. eLife 2020, DOI: 10.7554/eLife.56525.

(34). Luo Y; Na Z; Slavoff S A P-Bodies: Composition, Properties, and Functions. Biochemistry 2018, 57, 2424. [PubMed: 29381060]

(35). Xing W; Muhlrad D; Parker R; Rosen MK A quantitative inventory of yeast P body proteins reveals principles of compositional specificity. eLife 2018, 489658.

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

(37). Bellesia G; Shea J-E Effect of μ-sheet propensity on peptide aggregation. J. Chem. Phys 2009, 130, 145103. [PubMed: 19368476]

(38). Martínez L; Andrade R; Birgin EG; Martínez JM PACKMOL: A package for building initial configurations for molecular dynamics simulations. J. Comput. Chem 2009, 30, 2157–2164. [PubMed: 19229944]

(39). Kedersha N; et al. Dynamic Shuttling of Tia-1 Accompanies the Recruitment of mRNA to Mammalian Stress Granules. J. Cell Biol 2000, 151, 1257–1268. [PubMed: 11121440]

(40). Ramaswami M; Taylor JP; Parker R Altered Ribostasis: RNA-Protein Granules in Degenerative Disorders. Cell 2013, 154, 727–736. [PubMed: 23953108]

(41). Lin Y; Protter DSW; Rosen MK; Parker R Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol. Cell 2015, 60, 208–219. [PubMed: 26412307]

(42). Guillén-Boixet J; et al. RNA-Induced Conformational Switching and Clustering of G3BP Drive Stress Granule Assembly by Condensation. Cell 2020, 181, 346–361. [PubMed: 32302572]

(43). Garcia-Jove Navarro M; et al. RNA is a critical element for the sizing and the composition of phase-separated RNA–protein condensates. Nat. Commun 2019, DOI: 10.1038/s41467-019-11241-6.

(44). Dill KA; Ghosh K; Schnit JD Physical limits of cells and proteomes. Proc. Natl. Acad. Sci. U. S. A 2011, 108, 17876. [PubMed: 22006304]

(45). Fulton AB How crowded is the cytoplasm? Cell 1982, 30, 345. [PubMed: 6754085]

(46). Zimmerman SB; Trach SO Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. J. Mol. Biol 1991, 222, 599. [PubMed: 1748995]

(47). Milo R; Phillips R Cell biology by the numbers, 2016.

(48). Garcia-Seisdedos H; Empereur-Mot C; Elad N; Levy ED Proteins evolve on the edge of supramolecular self-assembly. Nature 2017, 548, 244–247. [PubMed: 28783726]

(49). Wheeler JR; Matheny T; Jain S; Abrisch R; Parker R Distinct stages in stress granule assembly and disassembly. eLife 2016, DOI: 10.7554/eLife.18413.

(50). Protter DSW; Parker R Principles and Properties of Stress Granules. Trends Cell Biol. 2016, 26, 668–679. [PubMed: 27289443]

(51). Levine BG; Stone JE; Kohlmeyer A Fast analysis of molecular dynamics trajectories with graphics processing units-Radial distribution function histogramming. J. Comput. Phys 2011, 230, 3556. [PubMed: 21547007]

(52). Adivarahan S; et al. Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway. Mol. Cell 2018, 72, 727. [PubMed: 30415950]

(53). Wang J; et al. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell 2018, 174, 688–699. [PubMed: 29961577]

(54). Kroschwald S; et al. Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. eLife 2015, 4, e06807. [PubMed: 26238190]

(55). Ruff KM; Dar F; Pappu RV Ligand effects on phase separation of multivalent macromolecules. Proc. Natl. Acad. Sci. U. S. A 2021, 118, e201784118. [PubMed: 33653957]
(56). Bond U Stressed out! Effects of environmental stress on mRNA metabolism. FEMS Yeast Res. 2006, 6, 160. [PubMed: 16487339]
Figure 1.
Patchy particle model and the interaction network. Schematic representation of some P-body proteins with different valencies. The colors on the side beads represent an adhesive site that is specific to its complementary adhesive site on another component of the same color. Therefore, each particle has a valency that is specific to every other particle in the simulation. The RNA is modeled as a semiflexible polymer chain. The length of the RNA chain, $L_{RNA}$, is a parameter that we vary in our simulations.
Figure 2.
Phase separation in the absence of RNA (A) Assembly sizes in the absence of RNA, for varying protein volume fractions and interaction strengths. The size of the largest cluster ($L_{\text{clus}}$) is plotted as a function of interaction strength (per interaction patch). $\phi_{\text{prot}}$ refers to the volume fraction of proteins in the simulation, $\phi_{\text{prot}} = N_{\text{tot}}(4/3 \pi R^3)/L^3$. $N_{\text{tot}} = 1000$, and $L$ is the size of the simulation box. $L_{\text{clus}}$ values were computed across 500 different equilibrium snapshots and five independent simulation trajectories. (B) Probability of finding a multivalent protein in a cluster that is larger than 10 molecules in size. Dcp2, Edc3, Patl, and Lsml are core proteins. Different color bars represent different values of interaction strength, $\epsilon_{\text{sp}}$ (per patch). (C) Radial distribution function, $p(r)$, for core—core, core—noncore, and noncore—noncore interactions showing the probability of finding two particles of a given type within a certain distance of each other (normalized by the corresponding value for a pair of ideal gas particles at the same bulk density).
Figure 3.
RNA-mediated phase separation. (A) The values along the x-axis indicate increasing RNA concentrations represented as effective number of RNA beads. Here, $L_{RNA}$ and $N_{RNA}$ refer to the length and the total number of RNA chains, respectively. $L_{clus}$ is the size of the single largest cluster in the simulation, at equilibrium, expressed as a fraction of the total number of monomers ($N_{tot}$). The total number of protein particles ($N_{tot}$) in the simulation box was 1000. A $L_{clus}$ of 0.1, therefore, refers to a cluster that is composed of 100 protein monomers. $L_{clus}$ values were computed across 500 different equilibrium snapshots and five independent simulation trajectories. Panels B and C show the RNA-mediated nature of assembly (single trajectory), with an increase noncore RNA contacts (shaded blue in the figure) preceding the formation of larger sized clusters. $\varepsilon_{sp}$ for protein—protein interactions was set to 2.5kT. The interaction strength between any RNA-binding patch and RNA was set to 1kT (per RNA bead). The curves for the largest duster size were smoothed by using cubic spline fitting as a guide to the eye.
Figure 4.
RNA promotes noncore interactions. Radial distribution function, $p(r)$, for core—core (top), core—noncore (center), and noncore—noncore (lowest panel) for $L_{RNA}$ of 10 (right panel) and 120 (left panel). The purple, cyan, and black curves represent different numbers of RNA binding sites in the simulation box. The standard errors (SD/$\sqrt{n}$) in $p(r)$ were computed over 1000 different equilibrium snapshots.
**Figure 5.**

Probability of finding each protein component within a multimeric cluster (≥10 monomers) in the presence of increasing RNA concentrations. Panels A, B, and C show results in the presence of RNA chains of length 10, 50, and 120, respectively. Higher values of $P_{\text{bound}}$ indicate a greater likelihood of a component being found within RNA—protein clusters.
RNA compaction within cluster. (A) Distribution of radius of gyration of RNA chain at increasing bulk density of proteins. At higher volume fractions, where RNA—protein clusters form, the RNA chains assume a more compact configuration. (B) Comparison of conformational statistics (radius of gyration) of a sequestered RNA chain within an RNP cluster and a free RNA chain. Statistics in this figure were computed over 1000 different equilibrium snapshots and multiple RNA chains in the simulation box.
Figure 7.
Summary. Addition of long chain RNA molecules can promote large protein-rich clusters even at low concentrations \([C]\) and weak protein—protein interactions. These clusters enrich noncore proteins (yellow balls) more than the clusters that are observed at large protein concentrations and strong protein—protein interactions. Blue balls represent core proteins.
Table 1.
Protein—Protein Interactions in P-Body Proteins (Based on the Study by Xing et al.33)

| Protein Interaction                      | Protein—RNA Interactions | Total Interactions | Core/Noncore |
|------------------------------------------|--------------------------|--------------------|--------------|
| Dcp2, Edc3 + Patl (10 in total), Dhhl, Edcl | 1                        | 13                 | core         |
| Edc3, Dcp2, Edc3, Dhhl, Upfl, Patl       | 1                        | 6                  | core         |
| Patl, Dcp2/Xml (same site), Dhhl, Edc3, Lsml—7, Upfl | 1                      | 8                  | core         |
| Xml, Patl                                | 1                        | 2                  | noncore      |
| Lsml, Patl, Lsml (2)                     | 1                        | 4                  | core         |
| Upfl, Patl, Edc3, Upfl2                  | 1                        | 4                  | core         |
| Dhhl, Patl/Edc3 (same site), Dcp2, (Pop2) | 1                        | 4                  | core         |
| Pop2, CCR4, (Dhhl)                       | 1                        | 3                  | noncore      |
| CCR4, Pop2                               | 1                        | 2                  | noncore      |
| Not2                                     | 1                        | 1                  | noncore      |
| Upf3, Upf2                               | 1                        | 2                  | noncore      |
| Upf3, Upf1, Upf3                         | 1                        | 3                  | noncore      |
| Hek2                                     | 1                        | 1                  | noncore      |
| Shp1                                     | 1                        | 1                  | noncore      |
| Edc3, Dcp2                               | 1                        | 2                  | noncore      |
Table 2.

| notation (or terminology) | physical interpretation | definition |
|---------------------------|-------------------------|------------|
| $\phi_{\text{prot}}$     | volume fraction of proteins in the simulation box | $\phi_{\text{prot}} = \frac{N_{\text{tot}} \left( \frac{4}{3} \pi R^3 \right)}{L^3}$, where $R$ is the size of the protein monomer and $L$ is the length of the cubic simulation box; this is a dimensionless quantity used to represent protein concentrations |
| $L_{\text{clus}}$       | size of the single largest cluster | |
| $N_{\text{tot}}$        | total number of protein monomers in the system. | $N_{\text{tot}} = 1000$ in our patchy-particle simulations |
| $N_{\text{RNA}}$        | number of RNA chains in simulation | parameter that is varied in our simulations to vary the protein/RNA ratio. |
| $L_{\text{RNA}}$        | length of semiflexible RNA chains in LD simulations | represented in terms of number of nucleotides |
| $R_g$                   | radius of gyration of RNA | |