

# **Liquid-Liquid Phase Separation of Patchy Particles Illuminates Diverse Effects of Regulatory Components on Protein Droplet Formation**

Valery Nguemaha<sup>1</sup> and Huan-Xiang Zhou<sup>1,2,\*</sup>

<sup>1</sup>Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, United State

<sup>2</sup>Department of Chemistry and Department of Physics, University of Illinois at Chicago, Chicago, IL 60607, United States

\*Correspondence and request for materials should be addressed to H.-X.Z. (e-mail: [hzhou4@fsu.edu](mailto:hzhou4@fsu.edu))

## Abstract

Recently many cellular functions have been associated with membraneless organelles, or protein droplets, formed by liquid-liquid phase separation (LLPS). Proteins in these droplets often contain RNA-binding domains, but the effects of RNA on LLPS have been controversial. To gain better understanding on the roles of RNA, here we used Gibbs-ensemble simulations to determine phase diagrams of two-component patchy particles, as models for mixtures of proteins with RNA or other regulatory components. Protein-like particles have four patches, with attraction strength  $\epsilon_{PP}$ ; regulatory particles experience mutual steric repulsion but have two attractive patches toward proteins, with the strength  $\epsilon_{PR}$  tunable. At low  $\epsilon_{PR}$ , the regulator, due to steric repulsion, preferentially partitions in the dispersed phase, thereby displacing the protein into the droplet phase and promoting LLPS. At moderate  $\epsilon_{PR}$ , the regulator starts to partition and displace the protein in the droplet phase, but only to weaken bonding networks and thereby suppress LLPS. At  $\epsilon_{PR} > \epsilon_{PP}$ , the enhanced bonding ability of the regulator initially promotes LLPS, but at higher amounts, the resulting displacement of the protein suppresses LLPS. These results illustrate how RNA can have disparate effects on LLPS, thus able to perform diverse functions in different organelles.

## Introduction

There is intense current interest in membraneless organelles, thought to be formed by liquid-liquid phase separation (LLPS) of protein-RNA mixtures.<sup>1-5</sup> Upon phase separation, these organelles emerge as droplets containing condensed component proteins, from the cytoplasm or nucleoplasm where the component proteins are dilute. Membraneless organelles have been associated with many cellular functions. In particular, the condensation of component proteins may facilitate the assembly of complexes for biochemical processes, as exemplified by the nucleolus, an organelle with a primary function in ribosome pre-assembly. Component proteins usually contain intrinsically disordered regions (IDRs) and RNA-binding domains. The effects of RNA on LLPS have been investigated in many experimental studies, but conflicting results have been reported.<sup>2,6-15</sup> The aim of this study was to use computation to elucidate potential roles of RNA and other regulatory components in LLPS and the underlying physical mechanisms.

A hallmark of LLPS (in contrast to, e.g., aggregation) is thermodynamic reversibility. That is, droplets can easily dissolve upon raising temperature or salt concentration and can reform when conditions are reverted. The reversibility comes because the two phases, droplet and dispersed, are in thermodynamic equilibrium. The two phases separate when the molecules can achieve the same low free energy by adopting two distinct types of configurations, with disparate concentrations and extents of intermolecular bonding.<sup>16</sup> The low free energy is achieved by distinct means in the two phases: e.g., high entropy (due to low concentration) in the dispersed phase but low enthalpy (from intermolecular bonding) in the droplet phase. Raising temperature or salt concentration leads to reduced stabilization of the droplet phase by intermolecular bonding. Beyond a critical point, the distinction between the two phases vanishes, and hence droplets dissolve. The thermodynamic equilibrium between the two phases is characterized by a phase diagram, or binodal, which is specified by the concentrations in the two coexisting phases for conditions below the critical point. In general, the stronger the intermolecular bonding, the higher the

critical point.

Among the experimental studies on LLPS of protein-RNA mixtures, most reported that the addition of RNA promoted LLPS, indicated by either reduced threshold protein concentrations for droplet formation or increased critical salt concentrations.<sup>2,7,8,10-13</sup> Reinforcing the promotional effects, both Saha et al.<sup>12</sup> and Smith et al.<sup>13</sup> found that molecules that competed for RNA binding with the droplet-forming proteins also suppressed phase separation. Intriguingly, Burke et al.,<sup>6</sup> Zhang et al.,<sup>9</sup> and Banerjee et al.<sup>14</sup> observed that RNA promoted LLPS up to a point in the RNA to protein molar ratio; further increase in RNA led to LLPS suppression. The mechanism of such a dual effect remains unresolved; Zhang et al. suggested screening of interprotein electrostatic interactions by RNA (akin to salt) at higher concentrations, whereas Banerjee et al. speculated that increased RNA led to a charge inversion of the protein component. In yet another variation, Wei et al.<sup>15</sup> recently found that RNA had little effects on the threshold protein concentration and critical salt concentration, but significantly reduced the protein concentration in the droplet phase. The binodals in the absence and presence of RNA were fitted to a polymer theory, with RNA treated only implicitly. Based on the fitting parameter values, the authors attributed the leftward shift of the high-concentration arm of the binodal in the presence of RNA to the latter's weakening of interprotein attraction.

It should be recognized that the effects of RNA on protein LLPS are a reflection of intermolecular interactions rather than anything uniqueness about RNA. Indeed, similar promotional effects were observed in one case when RNA was replaced by heparin (which, like RNA, is an acidic polymer and thus presumably binds to the same sites on the droplet-forming protein)<sup>10</sup> and in another case, instead of RNA, the regulatory component was arginine-rich peptides (expected to bind to acidic tracks, rather than RNA-binding sites, of the droplet-forming protein).<sup>11</sup> LLPS also occurs for, and in fact, was first observed on, globular proteins.<sup>17,18</sup> The phase separation of droplet-forming proteins was suppressed, with decreased critical temperatures, when another protein was added.<sup>19,20</sup> For both disordered and globular proteins, LLPS was

promoted by crowding agents such as polyethylene glycol (PEG) or Ficoll.<sup>7,8,21,22</sup> Interestingly, in the latter study, bovine serum albumin (BSA) was found to suppress the phase separation of hnRNPA1 or its IDR fragment, but promoted, as if as a crowding agent, the phase separation of a protein-RNA mixture. Moreover, lysozyme exhibited a concentration-dependent dual effect on the LLPS of the IDR fragment of FUS (FUS<sub>IDR</sub>). In short, RNA and other regulatory components can exert similarly diverse effects on the LLPS of disordered and globular proteins. This is not surprising, since the LLPS of disordered and globular proteins have a common physical basis, though their phase behaviors show characteristic differences.<sup>23</sup>

Experimental studies have presented many observations on LLPS that beg for mechanistic answers. In principle, theoretical calculations and molecular simulations can provide these answers, but still face significant technical challenges. Polymer theories,<sup>24,25</sup> coarse-grained simulations,<sup>26</sup> and hybrids thereof<sup>27</sup> have been used to determine phase diagrams of chain molecules, in particular as models of intrinsically disordered proteins (IDPs), and to address several important questions. These include how phase boundaries are affected by charge patterns along the sequence, chain length, charge mutations, and salt. The study that relates most closely to the present one is by Lin et al.,<sup>25</sup> who treated the phase separation of a mixture of two IDP sequences within a polymer theory. All these approaches have limitations and hence their predictions are necessarily qualitative. In particular, polymer theories often provide only a very crude account of excluded-volume effects.<sup>23</sup>

Here we used Gibbs-ensemble simulations to investigate how regulatory components affect protein droplet formation, a pressing question that has received only scant attention in previous theoretical and simulation studies. While the promotion of LLPS by crowding agents can be understood from the perspective of excluded-volume effects,<sup>28,29</sup> when RNA or another regulatory component is attractive toward the droplet-forming component and hence is recruited into the droplet phase, the physical rules governing the foregoing disparate effects on LLPS are yet to be established. Toward that aim, we chose a relatively simple model for

protein-regulator mixtures, to calculate phase diagrams over a wide range of parameter values. To our knowledge, this is the first computational study devoted to uncovering the physical determinants underlying the roles of RNA and other regulatory components in LLPS. Our results qualitatively recapitulate the behaviors observed in the experimental studies, and provide a physical basis, in terms of the relative strengths of protein-protein and protein-regulator interactions, to reconcile the apparently conflicting effects of RNA on droplet formation.

## Results

**A simple model for protein-regulator mixtures.** The droplet phase of protein-RNA mixtures, as true for LLPS in general, is stabilized by bonding networks of the molecules.<sup>23</sup> As alluded to above, we reason that the disparate effects on LLPS are generic instead of unique to RNA, and hence can be captured by models that possess common features of mixtures of proteins with RNA or other regulatory components, in particular, polyvalent interactions. Here we used patchy particles to model protein and regulatory molecules (Figure 1a). These particles, with attractive patches decorated on a hard core, which in our case is spherical, have been used to model proteins in LLPS.<sup>30-33</sup>

The details of the protein and regulator models were chosen based on a number of observations regarding LLPS of protein-regulator mixtures. First of all, in many experimental studies where LLPS has been demonstrated, there exists a single “driver” protein that can form droplets on its own, whereas RNA or other regulatory components can modulate the phase boundaries. Correspondingly, interprotein interactions have to be strongly attractive (for droplet formation), whereas inter-regulator interactions are weakly attractive or even purely repulsive. The latter scenario is certainly true for RNA, due to the uniform anionic charges of the phosphates. Meanwhile protein-regulator interactions are expected to be highly dependent on the particular molecules involved. Lastly, all particles occupy volume so they experience steric repulsion, which, much like the attractive interactions

between patches, should play a role in determining phase diagrams, as demonstrated by effects of crowding agents.

In our models, protein-like (“P”) and regulatory (“R”) particles have 4 and 2 patches, respectively, to be denoted as  $m_S$ , where  $S = P$  or  $R$ . The patches together cover the same fraction,  $\chi$ , of the surface areas of the hard cores, which have the same diameter  $\sigma$  for both types of particles. Each patch is a spherical disk, with the spanning polar angle  $\theta_S$  given by

$$\cos \theta_S = 1 - 2\chi/m_S \quad (1)$$

Each patch is identified by the unit vector from the particle center to the patch center, denoted by  $\mathbf{n}$  with appropriate indices. The centers of the patches are positioned as a tetrahedron on a P particle but at the poles on an R particle. The greater spreading of the patches on a P particle gives it an advantage in forming bonds with multiple partner particles. A bond is formed when a patch on one particle is in contact with a patch on another particle.

The interaction energy between any two particles,  $i$  and  $j$ , has a directional square-well form,<sup>34</sup>

$$U_{ij} = U_{sw}(r_{ij}) U_{pp}(\hat{\mathbf{r}}_{ij}, \mathbf{\Omega}_i, \mathbf{\Omega}_j) \quad (2)$$

where  $r_{ij}$  is the distance between the particle centers,  $\hat{\mathbf{r}}_{ij}$  is the unit vector between them,  $\mathbf{\Omega}_i$  and  $\mathbf{\Omega}_j$  denote the orientations of the particles. The square-well part has the usual form,

$$U_{sw}(r_{ij}) = \begin{cases} \infty, & r_{ij} \leq \sigma \\ -\varepsilon_{ij}, & \sigma < r_{ij} \leq \sigma + \lambda \\ 0, & r_{ij} \geq \sigma + \lambda \end{cases} \quad (3)$$

where  $\varepsilon_{ij}$  and  $\lambda$  are the strength and width, respectively, of the interparticle attraction. The interparticle attraction operates only when the patches of the two particles form a bond. The latter condition is specified by  $\hat{\mathbf{r}}_{ij} \cdot \mathbf{n}_{i\alpha} > \cos \theta_{i\alpha}$  and  $-\hat{\mathbf{r}}_{ij} \cdot \mathbf{n}_{j\beta} > \cos \theta_{j\beta}$ , where  $\alpha$  or  $\beta$  refers to any of the patches on particle  $i$  or  $j$ , and  $\theta_{i\alpha}$  or  $\theta_{j\beta}$  is the corresponding spanning polar angle of the patch. Hence

$$U_{pp}(\hat{\mathbf{r}}_{ij}, \mathbf{\Omega}_i, \mathbf{\Omega}_j) = \begin{cases} 1, & \text{if an } i-j \text{ bond is formed} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

Note that each patch can simultaneously form bonds with multiple partner particles. The total energy of the two-component patchy particle system is the sum of all pairwise interactions between particles.

In the present work,  $\chi$  was fixed at 0.7 and  $\lambda$  was fixed at  $0.5\sigma$ . We set  $\varepsilon_{RR} = 0$  (so the R particles experienced only steric repulsion between each other), and varied the strength of attraction between P and R particles,  $\varepsilon_{PR}$ , over a wide range, from 0 to  $1.5\varepsilon_{PP}$ .

**Phase diagrams.** For each  $\varepsilon_{PR}$ , we used Gibbs-ensemble Monte Carlo simulations<sup>35</sup> to determine phase diagrams for the protein-regulator model system over a wide range of molar ratios between P and R particles. In these simulations, fixed numbers of P and R particles were placed in a box with a fixed volume and maintained at a constant temperature. The box was divided into two regions, and the volume and the particle numbers in each region were changed to achieve equality in pressure and chemical potentials. After reaching equilibrium, the two regions represent different phases, with distinct concentrations (denoted by  $\rho$  with appropriate subscript and superscript) for the two types of particles (Figure 1b). We refer to the low- and high-concentration phases by I and II. A binodal consists of coexistence concentrations,  $\rho_P^I$  and  $\rho_P^{II}$ , at a series of temperatures. For notational simplicity, hereafter all lengths and energies are, respectively, in units of the particle diameter  $\sigma$  and the strength of P-P attraction,  $\varepsilon_{PP}$  (equivalent to defining  $\sigma = 1$  and  $\varepsilon_{PP} = 1$ ). Likewise the temperature,  $T$ , is in units of  $\varepsilon_{PP}/k_B$ , where  $k_B$  is the Boltzmann constant. We denote the ratio of total R and P particle numbers as  $X$ .

The effects of the R particles on the phase diagrams differ qualitatively depending on  $\varepsilon_{PR}$  (Figure 2). They fall into three regimes, at low, moderate, and high  $\varepsilon_{PR}$  values. The low- $\varepsilon_{PR}$  regime is represented by the results at  $\varepsilon_{PR} = 0$  (Figure 2a). In this case, R particles promote LLPS. The critical temperature ( $T_c$ ) for the pure P system (i.e.,  $X = 0$ ) is 0.75 (in agreement with a previous study<sup>34</sup>). With increasing  $X$ ,  $T_c$  increases (indicated by an upward gray arrow in Figure 2a) and the



binodal broadens in both the low- and high-concentration arms, in a near symmetric fashion. Hence the threshold P concentration ( $\rho_P^I$ ) for phase separation, at a fixed temperature below  $T_c$  (e.g.,  $T = 0.7$ , indicated by a horizontal line in Figure 2a), is a decreasing function of  $X$  (Figure 2a inset).

As  $\varepsilon_{PR}$  increases from 0, the ability of the R particles deteriorates in promoting phase separation. At around  $\varepsilon_{PR} = 0.25$ , they begin to suppress phase separation. The suppressing effects are illustrated by the results at  $\varepsilon_{PR} = 0.5$  (Figure 2b), which are essentially opposite to those at  $\varepsilon_{PR} = 0$ . Specifically, now  $T_c$  decreases monotonically with  $X$ , and the binodal narrows in both the low- and high-concentration arms, though more pronounced in the former than in the latter. With further increase in  $\varepsilon_{PR}$ , the  $T_c$  decrease by the R particles is even somewhat greater at  $\varepsilon_{PR} = 0.75$  (Supplementary Fig. S1a) but then becomes tempered at  $\varepsilon_{PR} = 1.0$  (Figure 2c). Meanwhile, the effect on  $\rho_P^I$  nearly vanishes, with a hint of a small decrease around  $X = 0.1$  for  $T = 0.68$  (Figure 2c inset). Instead the most prominent effect is on  $\rho_P^{II}$ , which decreases by 35% at  $X = 0.43$  relative to the counterpart in the pure P system.

At even higher  $\varepsilon_{PR}$ , a third regime emerges, as illustrated by the results at  $\varepsilon_{PR} = 1.35$  (Figure 2d; see also Supplementary Fig. S1b for results at  $\varepsilon_{PR} = 1.25$ ). Here the effects of the R particles depend on the R to P molar ratio. With increasing  $X$ ,  $T_c$  initially increases, reaching a maximum at  $X = 0.35$  (indicated by a parabolic gray arrow in Figure 2d). Then  $T_c$  gradually decreases, going below the counterpart for the pure P system at approximately  $X = 1.3$ . The threshold P concentration for phase separation has a similar turnover behavior, reaching a minimum at  $X = 0.7$  for  $T = 0.7$  (Figure 2d inset). In contrast,  $\rho_P^{II}$  is a monotonically decreasing function of  $X$ . When  $\varepsilon_{PR}$  is further increased to 1.5, the maximal  $T_c$  shifts to a higher  $X$  value of approximately 0.65 and promotion of LLPS (i.e.,  $T_c >$  the counterpart for the pure P system) persists over a wider range of  $X$  (Supplementary Fig. S1c).

**Partitioning of R particles in two phases.** While Figure 2 displays the

concentrations of the P particles in the two coexisting phases, the phase equilibrium is fully characterized by the concentrations of both types of particles. In Figure 3 we present the compositions of the two phases for the system with the four representative  $\varepsilon_{PR}$  values at a given temperature. At each  $X$ , the compositions are shown as a pair of points,  $(\rho_P^I, \rho_R^I)$  and  $(\rho_P^{II}, \rho_R^{II})$ , connected by a black line (known as a tie line).

At  $\varepsilon_{PR} = 0$ , the R concentrations in the droplet phase are miniscule for all  $X$  values, while those in the dispersed phase increase with increasing  $X$  (Figure 3a). These R particles preferentially partition in the dispersed phase, as expected for volume-exclusion crowders. In doing so, they displace the P particles into the droplet phase, as indicated by higher and higher  $\rho_P^{II}$  and lower and lower  $\rho_P^I$  at increasing  $X$ . The latter is already seen in Figure 2a as near symmetric broadening of the binodal. Preferential partitioning in the dispersed phase occurs at  $\varepsilon_{PR}$  up to 0.2, where steric repulsion of R particles still dominates.

With increasing  $\varepsilon_{PR}$ , more and more R particles are recruited into the droplet phase. Whereas the R particles still prefer the dispersed phase at  $\varepsilon_{PR} = 0.5$ , as indicated by tie lines with a negative slope (i.e.,  $\rho_R^I > \rho_R^{II}$ ; Figure 3b), the preference switches to the droplet phase at  $\varepsilon_{PR} = 1.0$ , where the tie lines now have a positive slope (Figure 3c). The preference of the R particles for the droplet phase further strengthens at  $\varepsilon_{PR} = 1.35$  (Figure 3d). As more and more R particles are recruited into the droplet phase, more and more P particles are displaced from it. The latter trend is already seen in Figure 2c,d as significant leftward shifts by the high-concentration arm of the binodals. The simple reason for the displacement is that both P and R particles occupy volume so that the total concentration of particles that can be accommodated in the droplet phase is limited.

**Correlation between  $T_c$  and average number of bonds.** The critical temperature for LLPS is determined by the strength of the bonding networks in the droplet phase. We measured the latter by the average number of bonds,  $n_b$ , formed per particle at a given temperature. In Figure 4 we present  $T_c$  and  $n_b$  side by side, as functions of  $X$ . For all

the four representative  $\varepsilon_{PR}$  values, the dependences of  $T_c$  on  $X$  qualitatively follow those of  $n_b$ . Specifically, at  $\varepsilon_{PR} = 0$ , both  $T_c$  and  $n_b$  increase with increasing  $X$  (though a turnover may be expected at very high  $X$ ; Figure 4a), whereas at  $\varepsilon_{PR} = 0.5$  and 1.0 both  $T_c$  and  $n_b$  decrease monotonically with  $X$  (Figure 4b,c). At  $\varepsilon_{PR} = 1.35$ , both  $T_c$  and  $n_b$  exhibit a turnover, with the maxima occurring at nearly the same  $X$  value (around 0.3).

**Physical mechanisms for diverse effects of R particles.** We are now in a position to elucidate the physical mechanisms behind the various effects of the R particles on phase separation, by examining how they affect  $n_b$ . At low  $\varepsilon_{PR}$ , as typified by the situation at  $\varepsilon_{PR} = 0$ , the R particles preferentially partition into the dispersed phase, thereby displacing more and more of the P particles into the droplet phase. As  $\rho_R^{\text{II}}$  increases, each P particle is surrounded by more and more other P particles, and hence can form more bonds. This explains why  $n_b$  at  $T = 0.7$  increases from 6.6 at  $X = 0$  to 7.5 at  $X = 1.5$  (Figure 4a). Therefore volume-exclusion crowders promote LLPS, as observed for both protein-RNA mixtures and globular proteins,<sup>7,8,21,22</sup> by preferentially partitioning into the dispersed phase and thereby passively displacing droplet-forming proteins into the droplet phase to form stronger bonding networks.

To better explain how the R particles affect  $n_b$  at moderate and high  $\varepsilon_{PR}$ , we further decomposed it.  $n_b$  is the average of  $n_{b,P}$ , the number of bonds formed per P particle, and  $n_{b,R}$ , the counterpart per R particle, weighted by their populations ( $\rho_P^{\text{II}}$  and  $\rho_R^{\text{II}}$ ):

$$n_b = (\rho_P^{\text{II}} n_{b,P} + \rho_R^{\text{II}} n_{b,R}) / \rho^{\text{II}} \quad (5)$$

where  $\rho^{\text{II}} = \rho_P^{\text{II}} + \rho_R^{\text{II}}$  is the total concentration of particles in the droplet phase. Since a P particle can form bonds with both P and R partners (whereas an R particle can only have P partners), we further separated  $n_{b,P}$  by the types of partner particles,

$$n_{b,P} = n_{b,P-P} + n_{b,P-R} \quad (6)$$

In Figure 5 we present the decomposition of  $n_b$  as well as  $\rho_R^{\text{II}} / \rho^{\text{II}}$  at  $\varepsilon_{PR} = 0.5$ , 1.0, and 1.35 for a given temperature. At  $\varepsilon_{PR} = 0.5$ , the R particles have lower

bonding ability than the P particles (recall that  $\varepsilon_{PP} = 1$  by definition). Indeed, when surrounded by pure P particles, a P on average forms 7 bonds [ $n_{b,P}$  at  $X = 0$  in Figure 5a, at  $T = 0.68$ ], while an R forms only 4.5 bonds (extrapolation of  $n_{b,R}$  to  $X = 0$ ). As  $X$  increases, more and more R particles are recruited into the droplet phase, and hence they weight down the overall  $n_b$ . In addition, at the highest  $X$  value (i.e., 0.43) studied,  $n_{b,P}$  is significantly lower (to 6.5 bonds) than the counterpart in the pure P system. The latter comes about because, due to the displacement of P particles by R particles, the loss in  $n_{b,P-P}$ , the number of bonds with P partners, is not fully compensated by the gain in  $n_{b,P-R}$ , the number of bonds with R partners.

The behavior of  $n_b$  at  $\varepsilon_{PR} = 1.0$  is similar to that at  $\varepsilon_{PR} = 0.5$ , but shows subtle differences (Figure 5b). The bonding ability of the R particles is now higher. Still, when surrounded by pure P particles, an R still lags behind a P in the number of bonds formed (6.5 for R versus 7 for P), due to less spreading of the patches on the R particles. Moreover,  $n_{b,R}$  now decreases significantly with increasing  $X$ , due to decreasing  $\rho_P^{\text{II}}$  and hence a less number of bonding partners. The resulting widening gap between  $n_{b,P}$  and  $n_{b,R}$ , coupled with the increasing R population, accounts for most of the decrease in  $n_b$  with increasing  $X$ . Lastly,  $n_{b,P}$  itself decreases with increasing  $X$ , again because the gain in  $n_{b,P-R}$  does not keep pace with the loss in  $n_{b,P-P}$ . In short, at moderate  $\varepsilon_{RP}$ , the R particles, by partitioning into the droplet phase and displacing P particles, actively weaken bonding networks there and thereby suppress LLPS.

At  $\varepsilon_{PR} = 1.35$ , the stronger R-P attraction (relative to the P-P attraction) enables an R to form more bonds than a P when each is surrounded by pure P particles (8 for R versus 6.5 for P at  $T = 0.7$ ; Figure 5c). With the enhanced bonding ability for R, as  $X$  increases from 0 and more R particles displace P particles, the gain in bonds for a P formed with R particles is greater than the loss in bonds with other P particles (Figure 5d, left). Consequently,  $n_{b,P}$ , the average number of bonds formed by a P, initially increases with increasing  $X$ . After  $X = 0.25$ ,  $n_{b,P}$  levels off. Meanwhile  $n_{b,R}$  decreases monotonically with  $X$  due to the displacement of partners (i.e., P particles)

by R particles (Figure 5d, right). At around  $X = 0.3$ , the weighted average of  $n_{b,P}$  and  $n_{b,R}$ , i.e.,  $n_b$ , thus reaches maximum, and hence here the R particles have optimal ability in promoting LLPS. As  $X$  increases further and further and more and more P particles are displaced from the droplet phase, the R particles cannot form enough bonds so eventually they suppress phase separation.

## Discussion

We have calculated the phase diagrams of a model protein-regulator system over wide ranges of parameter values, in order to elucidate potential roles of RNA and other regulatory components in LLPS and the underlying physical mechanisms. The calculation results demonstrate that regulatory components can exhibit diverse effects on phase separation, depending on the relative strengths of protein-protein and protein-regulator interactions and the regulator to protein molar ratio (Figure 6). If the regulator is at most only weakly attractive toward the protein, it behaves like a volume-exclusion crowder in promoting LLPS, with increasing effects at higher regulator to protein molar ratios. (This trend may stop and even reverse eventually when the regulator is forced into the droplet phase and disrupts bonding networks there). In contrast, with moderate attraction toward the protein, the regulator suppresses LLPS (“active suppressor”), with effects monotonically increasing with the regulator to protein molar ratio. With strong protein-regulator interactions, the regulator initially promotes LLPS (“active promoter”), reaching a maximal effect at a certain regulator to protein molar ratio. But at higher and higher amounts of the regulator, the promotional effect subsides and eventually turns into suppression of LLPS.

One obvious limitation of our model is that the protein and regulator are treated as rigid, whereas proteins forming membraneless organelles usually contain IDRs, and regulatory RNA can be flexible molecules as well. We contend, as further discussed below, our findings and the resulting physical insights transcend the divide between rigid and flexible molecules. Other recent theoretical and simulation studies

have modeled flexible chain molecules.<sup>24-27</sup> These studies have not addressed the central question posed here, i.e., the effects of regulatory components on protein LLPS. It will be interesting to extend these studies to test the applicability of our conclusions on flexible molecules.

To elucidate the underlying physical mechanisms of the effects of regulatory components, we have found it useful to characterize bonding networks in the droplet phase, specifically, by the average number of bonds per particle and its decomposition by species. In general, molecular interaction networks, or transient clusters, are crucial to the physical understanding of dense macromolecular systems, such as those in cellular environments.<sup>16,36</sup>

Our results qualitatively recapitulate the behaviors observed in experimental studies of protein-crowder, protein-protein, and protein-RNA mixtures, and provide a physical basis to reconcile conflicting reported results for the effects of RNA on protein droplet formation. Crowding agents such as PEG and Ficoll are found to promote LLPS for protein-RNA mixtures and globular proteins.<sup>7,8,21,22</sup> Our calculation results with low  $\varepsilon_{PR}$  show that the promotional effect of such crowders comes because they preferentially partition in the dispersed phase, thereby passively displacing droplet-forming proteins into the droplet phase to strengthen the bonding networks there.

The critical temperature for LLPS of  $\gamma$ D crystallin decreased upon adding  $\beta$ B1 crystallin.<sup>19</sup> Similarly, droplet formation of a monoclonal antibody and hnRNPA1, respectively, was suppressed by human serum albumin<sup>20</sup> and BSA.<sup>22</sup> The changes in  $T_c$  and the tie lines of the protein mixtures in the first two studies are similar to our results obtained at moderate  $\varepsilon_{PR}$ . In this case, the regulatory protein has weaker attraction toward the droplet-forming protein than the latter toward itself. Still, the regulatory protein is recruited to the droplet phase, and thereby displaces the droplet-forming protein and weakens the bonding networks there. Both the recruitment and displacement features identified by our study were directly visualized by fluorescence microscopy for the hnRNPA1-BSA systems, whereas PEG and Ficoll,

also as expected, were excluded from the droplet phase.<sup>22</sup> Displacement occurs simple because there is a limit to the total concentration of macromolecules in the droplet phase.

Promotional effects of RNA on LLPS have been reported in a number of studies.<sup>2,7,8,10-13</sup> We explain these results by suggesting that the RNA molecules in these studies have stronger attraction toward the droplet-forming proteins than the latter toward themselves, and therefore actively strengthen bonding networks in the droplet phase. This suggestion is quite reasonable, since the droplet-forming proteins all contain specific RNA-binding domains to achieve strong bonding with RNA. We further hypothesize that the promotional effects in these studies were obtained when the amounts of RNA added were sufficiently low, and predict that adding higher amounts of RNA would eventually lead to LLPS suppression. Indeed, such dual effects have been observed in three studies.<sup>6,9,14</sup> The present study reveals that the suppression at high RNA-to-protein molar ratios arises from the displacement of proteins by RNA from the droplet phase, leading to an insufficient level of proteins to maintain the strength of the bonding networks of the protein and RNA molecules. This mechanism is distinct from those by Zhang et al.<sup>9</sup> and Banerjee et al.<sup>14</sup>, who speculated that RNA would either screen interprotein electrostatic interactions like diffuse salt ions or invert the protein net charge by strong binding to the protein surface. While these specific effects of RNA may exist, our mechanism is generally applicable to RNA and any other macromolecular regulators.

As support for our contention that the findings in the present study are applicable to both rigid and flexible molecules, a concentration-dependent dual effect was also found for lysozyme, a globular protein, on the LLPS of FUS<sub>IDR</sub>.<sup>22</sup> In comparison, BSA did not show any sign of promoting the FUS<sub>IDR</sub> LLPS. We can now explain the different effects of lysozyme and BSA in terms of their relative strengths of attraction toward FUS<sub>IDR</sub>. That is, we predict that FUS<sub>IDR</sub> is more strongly attracted to lysozyme than to BSA. Interestingly, a recent small-angle neutron scattering study indicated exactly such differential strengths of attraction of the latter two proteins toward FlgM,

an intrinsically disordered protein.<sup>37</sup> While two regulatory proteins (lysozyme and BSA) can have divergent effects (active promoter vs. active suppressor) on the phase separation of a given protein ( $FUS_{IDR}$ ), a given regulatory protein can also have distinct effects on the phase separation of two different proteins. The latter was observed with BSA acting as an active suppressor for the LLPS of hnRNPA1 but as a volume-exclusion crowder for the LLPS of a protein-RNA mixture.<sup>22</sup> Again, the distinct effects can be attributed to the differential strengths of attraction of the regulatory protein toward the droplet-forming proteins. We note that the strengths of intermolecular interactions can be directly measured by the second virial coefficient or the binding affinity, determined at concentrations below the threshold for LLPS.<sup>13,15</sup>

In another intriguing report,<sup>15</sup> polyadenylate RNA had little effects on the threshold LAF-1 concentration for LLPS and critical salt concentration, but significantly reduced the protein concentration in the droplet phase. These effects are similar to our calculation results at  $\varepsilon_{PR} = 1.0$ , where the dominant effect of the R particles is to reduce  $\rho_P^{II}$  (Figure 2c). This comes about because most of these R particles partition into the droplet phase and displace P particles, thereby reducing  $\rho_P^{II}$  while largely preserving the strength of the bonding networks (Figure 5b). In contrast, Wei et al.<sup>15</sup> related the reduction of  $\rho_P^{II}$  to weakening of interprotein attraction by RNA. That RNA would displace proteins in the droplet phase was not considered in the model used for their data fitting.

Our study has identified the relative strengths of protein-protein and protein-regulator interactions as an important physical determinant for the effects of regulatory components on protein droplet formation. In particular, active promotion of LLPS occurs only when RNA or other regulators have stronger attraction toward proteins than the latter toward themselves. Conversely, RNA is expected to behave like volume-exclusion crowders when droplet-forming proteins do not have much affinity for RNA. This can happen when droplet-forming proteins are deleted of their RNA-binding regions. Burke et al.<sup>6</sup> actually tested such a construct,  $FUS_{IDR}$ , for RNA



effects and the results were null. It is possible that the  $FUS_{IDR}$  concentration in that test was too low, and that a higher  $FUS_{IDR}$  concentration may exist where LLPS can be induced by adding RNA. We hope that these and other predicted trends in Figure 6 will motivate experimental tests.

Different membraneless organelles may use RNA for different roles, some as clients whereas others for active promotion of LLPS. Our study demonstrates that these different roles can be fulfilled by modulating the relative strengths of protein-protein and protein-RNA interactions. Furthermore, if RNA does promote LLPS, it can do so only within a certain concentration range. This turnover behavior affords cells the ability to use RNA expression level for regulating the formation of membraneless organelles.

## Methods

The phase diagrams of the protein-regulator model system for  $\epsilon_{PR}$  (strength of P-R attraction) between 0 and 1.5 and  $X$  (R to P molar ratio) between 0 and 1.5 were determined from Gibbs-ensemble Monte Carlo simulations.<sup>35</sup> In most simulations, the total number of particles was 512 and the total concentration,  $\rho_0$ , was 0.3, a setup where the binodal for the pure P system has been obtained previously;<sup>34</sup> our simulations at  $X = 0$  reproduced this result. Some simulations were also carried out with a total of 1000 particles (while preserving  $\rho_0$ ); the resulting coexistence concentrations were unchanged. In other simulations  $\rho_0$  was changed to 0.1 and the results are qualitatively consistent with those at  $\rho_0 = 0.3$ . Each binodal was obtained from simulations at a series of temperature ranging from 0.65 to 0.80. At each temperature, the first 1 to 1.5 million cycles were discarded to ensure equilibration, and the subsequent 2 to 3 million cycles were used for data collection. Each cycle consisted of 500 attempts of particle displacement, 500 attempts of particle exchange, and 5 attempts of volume exchange.

The critical point for each binodal was determined by fitting the temperature dependence of the coexisting P concentrations,  $\rho_P^I$  and  $\rho_P^{II}$ , to the law of rectilinear

diameter<sup>38</sup>

$$\frac{1}{2}(\rho_P^I + \rho_P^{II}) = \rho_P^c + A(T - T_c) \quad (7a)$$

$$\rho_P^{II} - \rho_P^I = B(T_c - T)^{0.32} \quad (7b)$$

where  $\rho_P^c$  is the critical protein concentration, and  $A$  and  $B$  are other fitting parameters. Each fit was performed first by including data at all available temperatures and then gradually removing data that were closer to  $T_c$ , until the fitted values of  $\rho_P^c$  and  $T_c$  were stable.

**Data availability.** The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

## References

- 1 Brangwynne, C. P. *et al.* Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-1732, doi:10.1126/science.1172046 (2009).
- 2 Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340, doi:10.1038/nature10879 (2012).
- 3 Hyman, A. A., Weber, C. A. & Julicher, F. Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* **30**, 39-58, doi:10.1146/annurev-cellbio-100913-013325 (2014).
- 4 Brangwynne, C. P., Tompa, P. & Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat Phys* **11**, 899-904, doi:10.1038/nphys3532 (2015).
- 5 Qin, S. & Zhou, H. X. Protein folding, binding, and droplet formation in cell-like conditions. *Curr Opin Struct Biol* **43**, 28-37, doi:10.1016/j.sbi.2016.10.006 (2017).
- 6 Burke, K. A., Janke, A. M., Rhine, C. L. & Fawzi, N. L. Residue-by-residue view of In vitro FUS granules that bind the C-terminal domain of RNA polymerase II. *Molecular cell* **60**, 231-241, doi:10.1016/j.molcel.2015.09.006

- (2015).
- 7 Lin, Y., Protter, D. S., Rosen, M. K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular cell* **60**, 208-219, doi:10.1016/j.molcel.2015.08.018 (2015).
  - 8 Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123-133, doi:10.1016/j.cell.2015.09.015 (2015).
  - 9 Zhang, H. *et al.* RNA controls polyQ protein phase transitions. *Molecular cell* **60**, 220-230, doi:10.1016/j.molcel.2015.09.017 (2015).
  - 10 Feric, M. *et al.* Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686-1697, doi:10.1016/j.cell.2016.04.047 (2016).
  - 11 Mitrea, D. M. *et al.* Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *eLife* **5**, doi:10.7554/eLife.13571 (2016).
  - 12 Saha, S. *et al.* Polar positioning of phase-separated liquid compartments in cells regulated by an mRNA competition mechanism. *Cell* **166**, 1572-1584.e1516, doi:10.1016/j.cell.2016.08.006 (2016).
  - 13 Smith, J. *et al.* Spatial patterning of P granules by RNA-induced phase separation of the intrinsically-disordered protein MEG-3. *eLife* **5**, doi:10.7554/eLife.21337 (2016).
  - 14 Banerjee, P. R., Milin, A. N., Moosa, M. M., Onuchic, P. L. & Deniz, A. A. Reentrant phase transition drives dynamic substructure formation in ribonucleoprotein droplets. *Angew Chem Int Ed Engl* **56**, 11354-11359, doi:10.1002/anie.201703191 (2017).
  - 15 Wei, M. T. *et al.* Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat Chem* **9**, 1118-1125, doi:10.1038/nchem.2803 (2017).
  - 16 Zhou, H. X. & Pang, X. Electrostatic interactions in protein structure, folding, binding, and condensation. *Chem Rev* **118**, 1691-1741,

- doi:10.1021/acs.chemrev.7b00305 (2018).
- 17 Tombs, M. P., Newsom, B. G. & Wilding, P. Protein solubility: Phase separation in arachin-salt-water systems. *Chem Biol Drug Des* **6**, 253-277, doi:10.1111/j.1399-3011.1974.tb02384.x (1974).
- 18 Tanaka, T., Ishimoto, C. & Chylack, L. T., Jr. Phase separation of a protein-water mixture in cold cataract in the young rat lens. *Science* **197**, 1010-1012 (1977).
- 19 Wang, Y., Lomakin, A., McManus, J. J., Ogun, O. & Benedek, G. B. Phase behavior of mixtures of human lens proteins Gamma D and Beta B1. *Proc Natl Acad Sci U S A* **107**, 13282-13287, doi:10.1073/pnas.1008353107 (2010).
- 20 Wang, Y., Lomakin, A., Latypov, R. F. & Benedek, G. B. Phase separation in solutions of monoclonal antibodies and the effect of human serum albumin. *Proc Natl Acad Sci U S A* **108**, 16606-16611, doi:10.1073/pnas.1112241108 (2011).
- 21 Annunziata, O. *et al.* Effect of polyethylene glycol on the liquid-liquid phase transition in aqueous protein solutions. *Proc Natl Acad Sci U S A* **99**, 14165-14170, doi:10.1073/pnas.212507199 (2002).
- 22 Protter, D. S. W. *et al.* Intrinsically disordered regions can contribute promiscuous interactions to RNP granule assembly. *Cell Rep* **22**, 1401-1412, doi:10.1016/j.celrep.2018.01.036 (2018).
- 23 Zhou, H. X., Nguemaha, V., Mazarakos, K. & Qin, S. Why do disordered and structured proteins behave differently in phase separation? *Trends Biochem Sci*, in press (2018).
- 24 Lin, Y. H., Forman-Kay, J. D. & Chan, H. S. Sequence-specific polyampholyte phase separation in membraneless organelles. *Phys Rev Lett* **117**, 178101, doi:10.1103/PhysRevLett.117.178101 (2016).
- 25 Lin, Y.-H., Brady, J. P., Forman-Kay, J. D. & Chan, H. S. Charge pattern matching as a ‘fuzzy’ mode of molecular recognition for the functional phase separations of intrinsically disordered proteins. *New J Phys* **19**, 115003

- (2017).
- 26 Dignon, G. L., Zheng, W., Kim, Y. C., Best, R. B. & Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput Biol* **14**, e1005941, doi:10.1371/journal.pcbi.1005941 (2018).
  - 27 Lytle, T. K. & Sing, C. E. Transfer matrix theory of polymer complex coacervation. *Soft Matter* **13**, 7001-7012, doi:10.1039/c7sm01080j (2017).
  - 28 Zhou, H. X., Rivas, G. & Minton, A. P. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* **37**, 375-397, doi:10.1146/annurev.biophys.37.032807.125817 (2008).
  - 29 Cho, E. J. & Kim, J. S. Crowding-induced phase separation of Lennard-Jones particles: implications to nuclear structures in a biological cell. *J Phys Chem B* **116**, 3874-3879, doi:10.1021/jp3006525 (2012).
  - 30 Sear, R. P. Phase behavior of a simple model of globular proteins. *J Chem Phys* **111** (1999).
  - 31 Liu, H., Kumar, S. K. & Sciortino, F. Vapor-liquid coexistence of patchy models: relevance to protein phase behavior. *J Chem Phys* **127**, 084902, doi:10.1063/1.2768056 (2007).
  - 32 Quinn, M. K. *et al.* How fluorescent labelling alters the solution behaviour of proteins. *Phys Chem Chem Phys* **17**, 31177-31187, doi:10.1039/c5cp04463d (2015).
  - 33 Kastelic, M., Kalyuzhnyi, Y. V. & Vlachy, V. Modeling phase transitions in mixtures of  $\beta$ - $\gamma$  lens crystallins. *Soft Matter* **12**, 7289-7298, doi:10.1039/c6sm01513a (2016).
  - 34 Kern, N. & Frenkel, D. Fluid-fluid coexistence in colloidal systems with short-ranged strongly directional attraction. *J Chem Phys* **118**, 9882-9889 (2003).
  - 35 Panagiotopoulos, A. Z. Direct determination of phase coexistence properties of fluids by Monte Carlo simulations in a new ensemble. *Mol Phys* **61**,

- 813-826 (1987).
- 36 Yu, I. *et al.* Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *eLife* **5**, doi:10.7554/eLife.19274 (2016).
  - 37 Banks, A., Qin, S., Weiss, K. L., Stanley, C. B. & Zhou, H. X. Intrinsically disordered protein exhibits both compaction and expansion under macromolecular crowding. *Biophys J* **114**, 1067-1079, doi:10.1016/j.bpj.2018.01.011 (2018).
  - 38 Guggenheim, E. A. The principle of corresponding states. *J Chem Phys* **13**, 253-261 (1945).

## Acknowledgements

This work was supported by National Institutes of Health Grant GM118091 and US-Israel Binational Science Foundation Grant 2015376.

## Author Contributions

V.N. and H.-X.Z. designed the research. V.N. performed the research. V.N. and H.-X.Z. wrote the manuscript.

## Additional Information

**Supplementary information** accompanies this paper

**Competing Interests:** The authors declare that they have no competing interests.

## Figure Captions

**Figure 1.** Illustration of the patchy particle model for protein-regulator mixtures and its two phases. **(a)** P and R particles, with cores in blue and red, respectively, and patches in white. **(b)** The low- and high-concentrations phases, after reaching equilibrium in a Gibbs ensemble Monte Carlo simulation. Patches are not displayed.

**Figure 2.** Phase diagrams. Results at  $\varepsilon_{PR} = 0.0, 0.5, 1.0$ , and  $1.35$  are shown in **(a) – (d)**, respectively. Open symbols are calculation results, and solid curves are fits to equation (7). The critical points from the fits are shown as filled symbols, with an arrowed gray line or parabola indicating the change with increasing  $X$ . A horizontal line through the left arm of the binodal indicates the temperature chosen for displaying results in the inset here and also in subsequent figures. Inset: ratio of threshold P concentrations between the pure P system and P-R mixture; a smooth curve is drawn to guide the eye.

**Figure 3.** Partitioning of P and R particles in the two phases. Results at  $\varepsilon_{PR} = 0.0, 0.5, 1.0$ , and  $1.35$  are shown in **(a) – (d)**, respectively, at the temperatures indicated. A pair of points connected by a black line represent  $(\rho_P^I, \rho_R^I)$  and  $(\rho_P^{II}, \rho_R^{II})$  for the system at a given  $X$ .

**Figure 4.** Correlation between critical temperature and average number of bonds per particle in the droplet phase. Results at  $\varepsilon_{PR} = 0.0, 0.5, 1.0$ , and  $1.35$  are shown in **(a) – (d)**, respectively. A smooth curve through the  $T_c$  data points is drawn to guide the eye; neighboring data points for  $n_b$  are connected by line segments. The latter are for the temperatures indicated.

**Figure 5.** The average number of bonds per particle in the droplet phase and its decomposition by species. **(a) – (c)** displays bond information (left ordinate axis) and molar fraction of R particles in the droplet phase,  $\rho_R^{II}/\rho^{II}$ , (right ordinate axis) at  $\varepsilon_{PR} = 0.5, 1.0$ , and  $1.35$ , respectively, for the temperatures indicated. **(d)** Illustration of bonding networks in the droplet phase at  $\varepsilon_{PR} = 1.35$ , where R particles strengthen bonding at small  $X$  (left,  $X = 0.25$ ) but disrupt bonding at large  $X$  (left,  $X = 1.0$ ). In the central cluster, an R particle and its bonded P neighbors (9 in the left and 4 in the right)

are shown as full-size spheres in red and blue, respectively, with bonds in gray; nonbonded P and R neighbors (2 Ps in the left and 4 Rs in the right) are shown as full-size spheres in cyan and magenta, respectively. The latter coloring scheme is also used for the more distant P and R particles, shown with radius reduced to 40%.

**Figure 6.** Expected effects of regulatory (“R”) particles on the droplet formation of protein (“P”) particles. Suppression of LLPS occurs at moderate P-R attraction (blue color), whereas promotion of LLPS occurs at either very weak or strong P-R attraction (red color). The promotion can turn into suppression of LLPS at high R to P ratios.













