scientific reports

Check for updates

OPEN

Size conservation emerges spontaneously in biomolecular condensates formed by scaffolds and surfactant clients

Ignacio Sanchez-Burgos[1,2], Jerelle A. Joseph[1,2,3], Rosana Collepardo-Guevara[1,2,3] & Jorge R. Espinosa[1,2]✉

Biomolecular condensates are liquid-like membraneless compartments that contribute to the spatiotemporal organization of proteins, RNA, and other biomolecules inside cells. Some membraneless compartments, such as nucleoli, are dispersed as different condensates that do not grow beyond a certain size, or do not present coalescence over time. In this work, using a minimal protein model, we show that phase separation of binary mixtures of scaffolds and low-valency clients that can act as surfactants—i.e., that significantly reduce the droplet surface tension—can yield either a single drop or multiple droplets that conserve their sizes on long timescales (herein 'multidroplet size-conserved' scenario'), depending on the scaffold to client ratio. Our simulations demonstrate that protein connectivity and condensate surface tension regulate the balance between these two scenarios. The multidroplet size-conserved scenario spontaneously arises at increasing surfactant-to-scaffold concentrations, when the interfacial penalty for creating small liquid droplets is sufficiently reduced by the surfactant proteins that are preferentially located at the interface. In contrast, low surfactant-to-scaffold concentrations enable continuous growth and fusion of droplets without restrictions. Overall, our work proposes one thermodynamic mechanism to help rationalize how size-conserved coexisting condensates can persist inside cells—shedding light on the roles of protein connectivity, binding affinity, and droplet composition in this process.

To fulfill their biological functions, cells must organize their contents into different compartments. One way of achieving spatiotemporal organization is via the formation of membrane-bound organelles. Another, that does not rely on membranes, is the self-assembly of proteins and nucleic acids into biomolecular condensates through the process of liquid–liquid phase separation (LLPS)[1–3]. In LLPS, a protein solution demixes into a protein-rich liquid phase—the condensate—in coexistence with a protein-poor liquid matrix. Condensates are ubiquitous within cells, with some of the most famous examples including stress granules[4], P granules[5], nuclear speckles[6], and the nucleolus[7].

The interest in understanding the molecular grammar[8,9] and biophysical determinants of cellular LLPS has significantly increased in recent years due, in part, to the realization that condensates take part in a wide-range of cellular functions—including genome silencing[10,11], signaling[12], buffering cellular noise[13], and formation of super-enhancers[14], among many others[15,16]. Furthermore, neurodegenerative diseases such as Alzheimer's, Parkinson's or ALS have been associated with aberrant or misregulated phase transitions[17–19].

The chief drivers of biomolecular LLPS are multivalent molecules (i.e., those possessing multiple interaction sites) known as 'scaffolds'[20]. Scaffolds are defined as the molecules—e.g., a single type of protein, a combination of various cognate proteins, or mixtures of proteins and nucleic acids[21]—that are able to sustain LLPS on their own. Above a critical concentration, scaffolds establish a sufficiently high number of transient attractive scaffold–scaffold interactions per unit of volume (the 'liquid-network connectivity'[22]) to compensate for the entropic loss upon demixing. In addition, scaffolds can establish interactions with biomolecules that are unable to phase separate on their own, known as 'clients', recruiting them to the condensates via scaffold–client interactions. The

[1]Maxwell Centre, Cavendish Laboratory, Department of Physics, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, UK. [2]Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK. [3]Department of Genetics, University of Cambridge, Downing Site, Cambridge CB2 3EH, UK. ✉email: jr752@cam.ac.uk

recruitment of clients can alter significantly the stability and structural properties of condensates[22–24]. Therefore, together, scaffolds and clients shape the biophysical properties of the biomolecular condensates that play fundamental roles within the cell[25–28].

A striking observation is the presence of condensates inside cells that do not grow beyond a certain size[7,29–31]. Basic thermodynamics suggest that over time, LLPS should result in the formation of a single large condensate rather than multiple coexisting small droplets[32]. The latter case is disfavored because, cummulatively, it yields a high surface area to volume ratio (i.e., high interfacial free energy penalty), while the former ensures that the interfacial free energy penalty is minimized. Despite of the thermodynamic preference for single condensates over multidroplet systems, both types of architectures are present within cells. Indeed, single large condensates have been observed[26,33], but also diverse coexisting size-restricted droplets have been reported in different in vivo systems, such as in the amphibian oocyte nucleolus[7], in the *Saccharomyces cerevisiae* cytoplasmic processing bodies[34] and elsewhere[35–37]. Additionally, size-conserved multidroplet architectures have been found in non-coalescing ribonucleoprotein condensates[26] and in multiphase complex coacervates in vitro[38]. However, the underlying molecular mechanisms and biophysical driving forces behind the formation of multiple coexisting droplets, also known as emulsification[39], require further investigation.

Several explanations for why or how emulsions can be thermodynamically stable at biological relevant timescales are currently under debate[40]. One possible mechanism is that the presence of active ATP-dependent processes might conveniently regulate the conditions where droplets grow and coalesce[41–44]. Other studies suggest that proteins with various highly distinct interacting domains may form micelle-like condensates[31,45,46]. In multicomponent mixtures, another possibility could be that specific binding proteins act as powerful surfactants and, thus, reduce the droplet surface tension penalty leading to multicondensate coexistence[35,36,38,47]. Moreover, a recent alternative explanation suggests that the interplay between protein diffusion and saturation of protein binding sites can also induce size-conservation in condensates[29,30]. It is plausible that all these different mechanisms contribute to the size conservation of condensates inside cells under different conditions.

In this work, we use a minimal protein model, which recapitulates the experimentally-observed relationship between protein valency and critical parameters[23,48,49], to investigate the regulation of droplet size in binary mixtures of multivalent proteins (scaffolds and clients). We show that liquid–liquid phase separation of scaffolds and clients mixtures, where clients act as surfactants, can give rise to single droplets or multiple size-conserved droplets (Fig. 1a). Further, we reveal that the transition between the two scenarios can be regulated by the condensate scaffold/surfactant client ratio. Our simulations suggest how general molecular features such as protein connectivity, binding affinity, and droplet composition can critically modulate and stabilize the formation of size-conserved condensates[29,30], and might have also implications to understanding the phase behavior of multilayered condensates[23,38,50].

Results

A minimal protein model for scaffold and client mixtures. Coarse-grained potentials have emerged as powerful tools for describing the phase behavior of biomolecules, such as proteins and nucleic acids, and delineating the underlying physicochemical features that drive LLPS[51,52]. Various levels of molecular resolution can be achieved with coarse-grained models; encompassing mean field descriptions[53,54], lattice-based approaches[55,56], minimal models[23,48,49,56–60], and sequence-dependent representations[46,60–65]. Here, we employ our minimal protein model[48], which has been previously applied to unveil the role of protein multivalency in multicomponent condensates[49] and multilayered condensate organization[66], as well as to investigate the role of RNA in RNA-binding protein nucleation and stability[67]. In this model, proteins are described by a pseudo hard-sphere potential[68] that accounts for their excluded volume, and by short-range potentials for modeling the different protein binding sites, and thereby mimicking protein multivalency[48] (Fig. 1b). For computational efficiency, an implicit solvent is used; therefore, the condensed phase corresponds to a liquid phase, and the diluted phase to a vapor. In what follows, the unit of distance is σ, the molecular diameter of the proteins (both scaffolds and clients), and the unit of energy $k_B T$ (for further details on the model parameters and the employed reduced units see the "Methods" section).

Following the framework of Banani et al.[20], scaffolds are defined as proteins that can establish both homotypic interactions and heterotypic interactions with clients, while clients (or hereafter called surfactants) as proteins that are limited to bind only to scaffolds (i.e., they do not bind to other clients; except where otherwise stated). Within this scheme, phase separation is driven by scaffolds (high-valency proteins), whereas surfactants (proteins with lower valency) are recruited to condensates at the expense of depleting LLPS-stabilizing scaffold–scaffold interactions. We model scaffolds as 4-binding sites particles and surfactants as 3-binding sites particles. (Fig. 1b). This choice fulfills two important requirements: (1) it ensures that scaffolds have a higher valency than the clients, and (2) it allows us to easily establish a common simulation temperature for the system at which each type of patchy particle distinctly behaves as either a scaffold or a surfactant client. The two-phase coexisting densities as a function of temperature for our model of scaffold proteins (blue), a 50:50 binary mixture of scaffolds and surfactants (black), and a system composed of just surfactants (red) are depicted in Fig. 1c (Top panel). Note that while in the 50:50 mixture, surfactant proteins do not interact homotypically, in the pure surfactant system they do; the latter allows us to compare the effect of protein valency on the critical parameters of pure self-interacting protein systems. These simulations show that the addition of surfactant proteins that are strong competitors for the scaffold binding sites significantly hinders the ability of scaffolds to phase separate (i.e., clients lower the critical temperature)[20,49]. Moreover, the presence of surfactant clients drastically reduces the surface tension of the condensates (black curve) as shown in Fig. 1c (Bottom panel). In the following section, we elucidate the implications of the client-induced surface tension reduction on the behavior of phase-separated condensates.

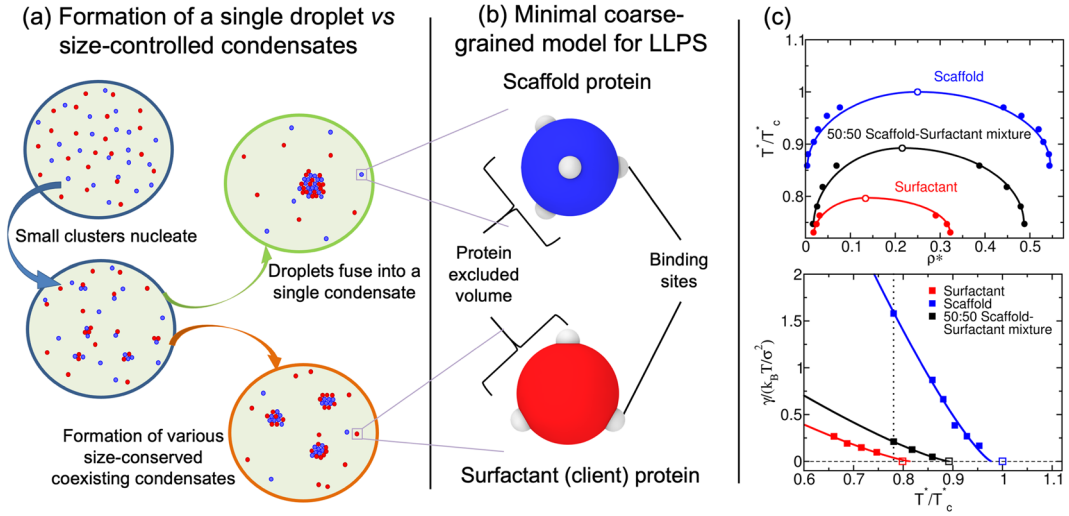

Figure 1. (**a**) Schematic representation of single-droplet formation versus size-conserved protein condensation. Scaffold proteins are depicted as blue spheres, while surfactant clients are shown as red spheres. (**b**) Minimal coarse-grained model for protein LLPS: Blue and red spheres represent the excluded volume of scaffold and surfactant (client) proteins respectively, while gray patches represent the binding sites of the proteins. Two different proteins are modeled: scaffold proteins, with 4 promiscuous binding sites in a tetrahedral arrangement, and surfactant proteins, with 3 binding sites in a planar equidistant arrangement that can only bind to scaffold binding sites (except where otherwise stated). Details on the model parameters are provided in the "Methods" section. (**c**) Top: Phase diagram in the temperature–density plane for a scaffold protein system (blue), a 50:50 scaffold–surfactant mixture (black) and for a hypothetical surfactant system in which client proteins can self-interact (red). The same self-interacting potential employed for scaffold proteins (blue) is also applied for surfactant proteins that can hypothetically self-interact (red). This system serves to further illustrate the effect of multivalency in LLPS. Filled circles indicate the estimated coexisting densities from Direct Coexistence simulations[68–71], and empty ones depict the critical points calculated via the law of rectilinear diameters and critical exponents[72] (see Supporting Information for details on these calculations). Bottom: Surface tension (γ) dependence on temperature for the scaffold system (blue), the 50:50 scaffold–surfactant mixture (black), and the hypothetical system in which surfactant proteins can self interact (red). Filled squares represent direct estimations of γ, continuous curves depict fits to our data of the following form: $\gamma \propto (T^* - T_c^*)^{1.26}$, and empty squares show the critical temperature of each system evaluated through the law of rectilinear diameters and critical exponents[72]. The vertical dashed line indicates the temperature at which the remainder of our study is performed. Note that temperature (in reduced units, T^*) is renormalized by the critical temperature T_c^* (also in reduced units) of the scaffold protein system ($T_c^* = 0.12$).

Size-conserved condensates can be regulated by surfactant proteins. Despite their seemingly subordinate role in condensate formation, client molecules (e.g. low valency proteins or self-avoiding nucleic acids) can significantly impact the organization of biomolecular condensates[20, 38, 73, 74]. Remarkable examples include the multilayered organization of stress granules[75, 76], the nucleoli[26], nuclear speckles[77], mixtures of RNA-binding proteins with RNA[78, 79], and in vitro complex coacervates[38, 73, 74]. While scaffolds mainly contribute to maximizing the condensate liquid network connectivity[20, 49], clients can severely reduce the droplet interfacial free energy cost of creating an interface[23]. The proliferation marker protein Ki-67, a component of the mitotic chromosome periphery, is an important example of a natural surfactant in intracellular compartmentalization[35]. In this context, our simulations show that low-valency clients that do not establish homotypic interactions behave as surfactants within condensates, as they can greatly decrease the condensate surface tension [Fig. 1c (Bottom)].

To further investigate the role of surfactant clients, we perform two sets of Direct Coexistence (DC) simulations[68–71]: (1) for the pure scaffold system, and (2) for the 50:50 binary mixture of scaffold and surfactants shown in Fig. 1c. At a constant temperature of $T^*/T_c^* = 0.75$ and a global (system) density of $\rho^* = 0.136$ (both in reduced units; see "Methods" section for further details on the employed reduced magnitudes), we create three different simulation box geometries with the dimensions summarized in Table 1. Using this approach, we can effectively modulate the surface/volume ratio of the condensed phase (hereafter called droplet): where the droplet surface is $S = 2 * L_y * L_z$ and its volume is $V = L_y * L_z * L_{x,slab}$; with $L_{x,slab}$ representing the width of the condensate in the x direction, and L_x, L_y and L_z being the different sides of the simulation box. Figure 2a–c, summarizes the phase behavior of the pure scaffold system (Top panels) and the 50:50 binary mixture (Middle panels) along the three designed simulation box geometries. The pure scaffold condensate exhibits distinct surface/volume ratios depending on the box geometry (see Table 1), while maintaining the same droplet density on all three cases. In other words, the scaffold condensate would be able to continuously grow as a single droplet at expense of the diluted phase until reaching equilibrium. In contrast, the scaffold–surfactant mixture yields various coexisting equilibrium droplets with a roughly constant surface/volume ratio of 0.21(2) σ^{-1} in all systems

Box geometry	(a)	(b)	(c)
Lx/σ	43.7	68.1	124.8
(Ly = Lz)/σ	17.5	13.7	10.3
Scaffold S/V ratio / σ^{-1}	0.14(2)	0.12(2)	0.09(2)
50:50 mixture S/V ratio / σ^{-1}	0.21(2)	0.21(2)	0.20(2)

Table 1. Simulation box dimensions and condensate surface/volume ratios (S/V) of the three box geometries represented in Fig. 2 for the pure scaffold system and the 50:50 binary mixture of scaffold and surfactant proteins. Geometries (a), (b) and (c) account for the (a), (b) and (c) panels shown in Fig. 2, respectively.

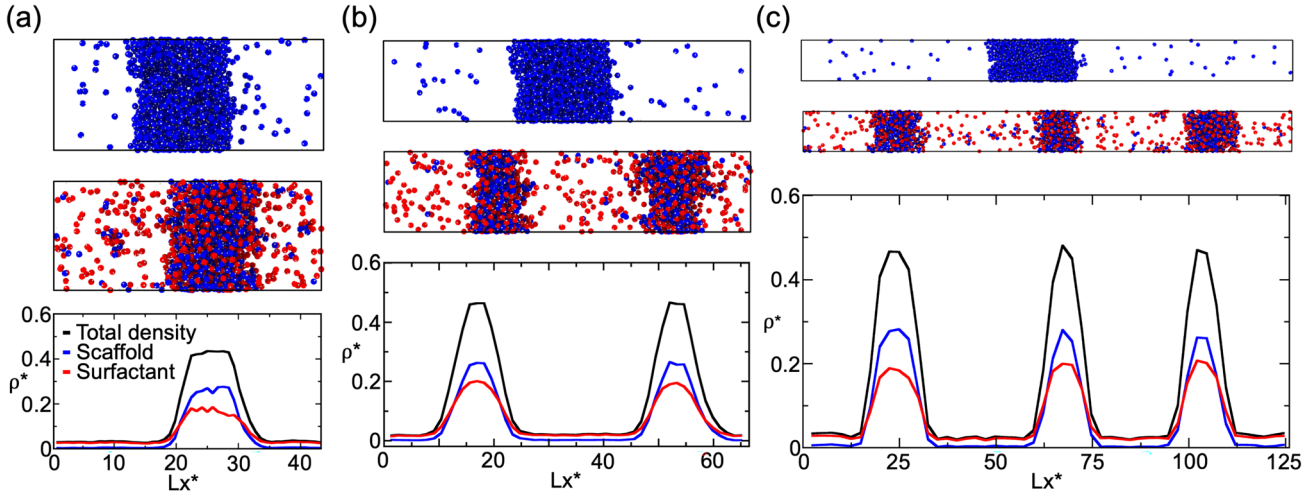

Figure 2. Direct Coexistence simulations for a scaffold protein system (Top panels) and a 50:50 binary mixture of scaffold and surfactant proteins (Middle panels) with different simulation box geometries (see Table 1 for the different simulation box sides employed in *a*, *b* and *c* geometries). All simulations were performed at $T^*/T_c^* = 0.75$ and a global density of $\rho^* = 0.136$. (**a**)–(**c**) (Top and Middle panels): Scaffold proteins are colored in blue and surfactant clients in red; the protein binding sites of both protein types are colored in gray. (**a**)–(**c**) (Bottom panels): Density profiles along the long side (x) of each simulation box L_x for the 50:50 scaffold–surfactant mixture. The (scaffold+surfactant) total density profile is depicted in black, the surfactant client density profile is shown in red, and that of scaffolds in blue. Densities of the different scaffold–surfactant coexisting droplets, as well as the scaffold molar fraction in the condensed and diluted phase, for each geometry are given in Table 2. The density of the pure scaffold condensates in all geometries is 0.54(1).

(Table 1). This result strongly suggests how size-conserved multidroplet phase behavior can be simply induced by the presence of low valency clients, which actively contribute to lowering the surface tension of the droplets [Fig. 1c (Bottom)]—thereby serving as natural surfactants[35,36].

We next analyze the composition of the different coexisting scaffold–surfactant condensates along the distinct box geometries [Fig. 2 (Bottom panels)]. In all cases, including the pure scaffold condensate, the properties of the droplets are remarkably similar; i.e., the density of all droplets, as well as their composition and surface tension are roughly constant (Table 2). Notably, we find that the surfactant density profile becomes higher than that of scaffolds at the droplet interface, showing how the partition coefficient of surfactants is greater than that of scaffolds in the outer region. Previous works suggest that accumulation of surfactants at the interface is preferable as it minimizes the condensate surface tension[35,38,66]. We also verify that the presence of multiple coexisting droplets is the thermodynamically stable state, rather than just metastable, by simulating over sufficiently long timescales to allow for multiple droplet fusion events and variations in droplet composition. Importantly, these tests reveal that even when in contact the droplets coexist without coalescing or altering their equilibrium composition. The multidroplet behavior of size-conserved condensates (in our case of $\sim 0.21\ \sigma^{-1}$ surface/volume ratio) is a consequence of the thermodynamic conditions of our system (i.e., mixture composition, temperature, and density). Note that droplet curvature effects such as Laplace internal pressure[80] or surface tension dependence on droplet curvature[81] have not been considered in our simulations, since we do not expect them to play an important role at biologically relevant droplet size scales ($\mathcal{O}\ \mu$m)[40]. Those effects are only expected to be dominant in the nanometer scale (i.e., up to droplet radii of tens of nanometers)[81–84].

The presence of surfactant clients within the condensate substantially lowers the liquid network connectivity[49,85,86] and, therefore, reduces the enthalpic gain sustaining LLPS. Consequently, the system minimizes its free energy by optimizing the number of surfactants that are incorporated into the condensed phase; i.e., by creating higher surface/volume ratios, where surfactants are preferentially located towards the interface rather than in the core. Such free energy optimization yields multiple coexisting condensates of a certain size, rather than a single-condensate system. Moreover, the emergence of multiple coexisting droplets stabilised by surfactant

Box geometry	(a), 1 droplet	(b), 2 droplets	(c), 3 droplets
$\gamma/(k_B T/\sigma^2)$	0.26(4)	0.23(4)	0.23(4)
Scaffold molar fraction in condensate	0.585(5)	0.576(5)	0.583(5)
Scaffold molar fraction in diluted phase	0.041(2)	0.041(2)	0.043(2)
$\rho^*_{\text{condensate}}$	0.47(1)	0.47(1)	0.46(1)
$\rho^*_{\text{diluted phase}}$	0.030(3)	0.027(3)	0.029(3)

Table 2. Properties of systems presented in Fig. 2 containing 50:50 scaffold–surfactant compositions in different simulation box geometries. In all cases, 6 independent simulations (with different initial velocity distributions), each starting from a pre-equilibrated configuration, were performed. For the geometries with more than one droplet, the values are averaged over the different droplets, although the variance between distinct droplets is significantly small as shown in Fig. 2 (Bottom panels).

proteins instead of a single condensate can also contribute to increase the entropy of the system, and thus, lowering its free energy. Nonetheless, this behavior is only thermodynamically favorable when the condensate surface tension is very low, as in the case of the 50:50 scaffold–surfactant condensates with a surface tension of $\gamma = 0.23\, k_B T\, \sigma^{-2}$—almost 7 times lower than that of the pure scaffold condensate ($\gamma = 1.58\, k_B T\, \sigma^{-2}$) at the same conditions ($T^*/T^*_c = 0.75$).

Surfactant concentration critically modulates droplet size. As discussed in the previous section, our minimal protein model shows that for a given composition of scaffold and surfactant clients, independent of the imposed box geometry, droplets can only grow to a certain size. This size-restricted growth is, in turn, determined by the optimal surface/volume ratio that minimizes the free energy of the coexisting liquid phases. Therefore, larger system sizes lead to higher number of coexisting size-conserved droplets. On the other hand, when the condensate is only composed of scaffold proteins, as the system size increases, the size of the condensate simply grows instead of yielding new multiple size-conserved droplets (Fig. 2). These results illustrate how a simple model for scaffold and surfactant proteins, merely controlled by protein valency and binding affinity, can recapitulate mesoscale features of in vivo and in vitro condensates that exhibit size-conserved growth[7,26,34–38]. Since this phase behavior only arises when the concentration of surfactants is not negligible, we now investigate how condensates can switch between both scenarios, and how their surface/volume ratio is modulated by their relative scaffold–surfactant composition.

By gradually increasing the surfactant client concentration of the scaffold–surfactant mixture (at a constant temperature and system density), the condensate size progressively decreases to accommodate the equilibrium droplet surface/volume ratio to the simulation box geometry; i.e., the condensate splits into two and, subsequently, into three coexisting liquid droplets (see Fig. 3). We note that the composition of the different coexisting droplets at a given surfactant concentration is remarkably similar, highlighting that all droplets are in equilibrium. In parallel, we evaluate the surface tension of the different coexisting droplets as a function of surfactant concentration. We find that γ monotonically decreases (but not linearly) as the surfactant concentration increases (Fig. 3). This result is not surprising (see Fig. 1c) given that one of the key molecular driving forces behind size-conserved multidroplet formation is the reduction of γ by surfactants coating the droplet surface. Above a certain surfactant concentration—exceeding 65 % for our system and at the given conditions—LLPS is inhibited. Beyond this limit, the condensate liquid network connectivity sustained by scaffold proteins can no longer compensate the mixing entropy of the system. We also analyze the surface/volume ratio of the droplets as a function of surfactant concentration. At infinitely low surfactant concentration, the condensate displays a ratio of $\sim 0.09\, \sigma^{-1}$. However, such a ratio is fully determined by the total number of proteins in the system and the box geometry, since as shown in Fig. 2, scaffold condensates can reach any droplet size when surfactants are absent. At low surfactant concentrations (i.e., % surfactant $< 27.5\%$), the maximum droplet size corresponds to a surface/volume ratio of $\sim 0.11\, \sigma^{-1}$. Beyond that threshold concentration, the condensate shrinks, and to achieve the equilibrium surface/volume ratio, it splits into smaller coexisting droplets. The maximum equilibrium droplet size in the two-droplet regime is that corresponding to ratios of $\sim 0.19\, \sigma^{-1}$. Finally, for surfactant compositions higher than 38%, three coexisting condensates emerge. The maximum surface/volume ratio that droplets can achieve is $\sim 0.27\, \sigma^{-1}$ at 60% client composition, which is only possible due to the extreme reduction in the surface tension ($\gamma = 0.15\, k_B T\, \sigma^{-2}$)—more than one order of magnitude lower than that of the pure scaffold condensate ($\gamma = 1.58\, k_B T\, \sigma^{-2}$) at the same temperature and system density.

Previous studies have highlighted the challenges associated with measuring condensate surface tensions due to the small size of protein droplets[3,87]. Nonetheless, there are available estimates of this magnitude for ribonucleoprotein condensates, and these measurements demonstrate that surfactant proteins can reduce γ by orders of magnitude[26]. With our minimal model, we qualitatively observe such behavior when surfactant proteins are recruited to the condensate, giving rise to emulsions of multiple coexisting droplets with very low surface tension. Surfactant proteins can lead to the formation of multidroplet emulsions by inducing multilayered condensate architectures[26,38]. Diverse biomolecular organelles, either in the cell nucleus, such as the nucleoli[26], or nuclear speckles[77], as well as in the cytosol, such as stress granules,[75,76] exhibit this type of organization. Moreover, different in vitro complex coacervates[38,50,73,74], bioengineered ribonucleoprotein condensates in living cells[88] and mixtures of RNA-binding proteins and RNA molecules[78,79] are also known to show multilayered assemblies. In Fig. 4a, we analyze the droplet architecture of a protein condensate with a 50:50 scaffold–surfactant composition

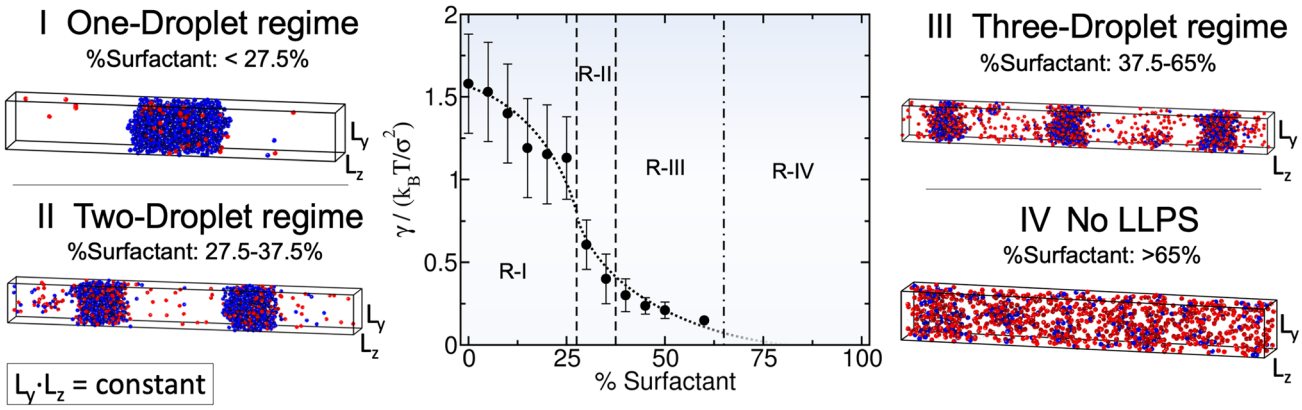

Figure 3. Surface tension (γ) dependence on the scaffold–surfactant composition (in %) at $T^*/T_c^* = 0.75$ and system (global) density of $\rho^* = 0.136$. Vertical dashed lines indicate the maximum surfactant concentration which allows LLPS for a given number-droplet regime in our system. Note that the maximum droplet size varies continuously with surfactant concentration even within the same number-droplet regime. The maximum condensate size in terms of surface/volume ratio along the different number-droplet regimes are: $\sim 0.11\,\sigma^{-1}$ for the one-droplet regime, $\sim 0.19\,\sigma^{-1}$ for the two-droplet regime, and $\sim 0.27\,\sigma^{-1}$ for the three-droplet regime. At surfactant client compositions exceeding 65% there is no LLPS. For each computed composition, the global density of the system and the simulation box cross-section ($L_z L_y$) are kept constant. Snapshots of the DC simulations along different surfactant composition droplet regimes are included. Please note that in the multidroplet regime, the different coexisting droplets (at the same surfactant concentration) exhibit similar compositions and densities.

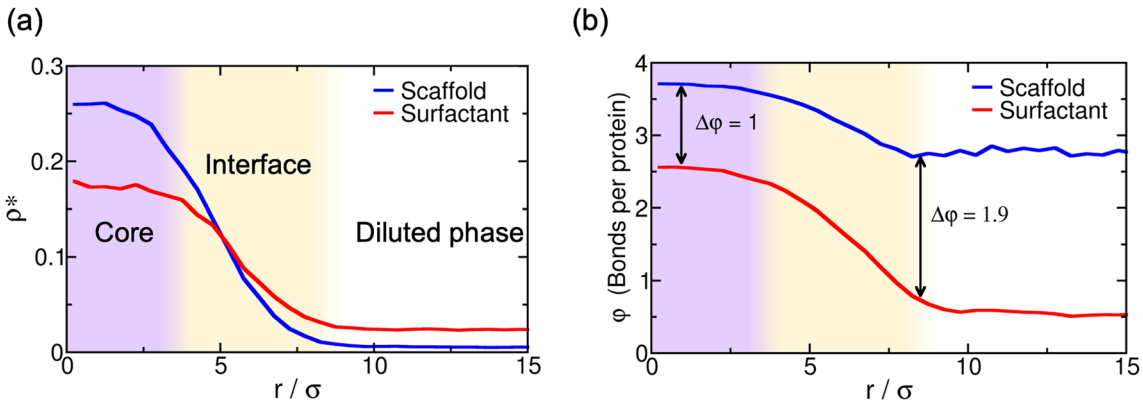

Figure 4. (**a**) Droplet density profile (in reduced units) for scaffold (blue) and surfactant proteins (red) from the droplet center of mass towards the surrounding diluted phase for the 50:50 binary mixture within the simulation box geometry shown in Fig. 2a at $T^*/T_c^* = 0.75$. (**b**) Average number of engaged binding sites per protein (φ) as a function of distance from the droplet center of mass for scaffold (blue) and surfactant proteins (red). One binding site is considered to be engaged to another if the distance between them is less than $0.145\,\sigma$ (i.e., the maximum bond length interaction between distinct protein binding sites; for further details on these calculations see Supporting Information).

[i.e., the system in Fig. 2a (Middle panel)]. We find that in the droplet core, the scaffold to surfactant ratio is much higher than along the interfacial region (Fig. 4a), where it drops to almost half that of the surfactant proteins. Nonetheless, the surfactant concentration within the droplet core is still remarkably high considering its destabilizing role in the condensate liquid network connectivity[49]. The observed non-homogeneous condensate organization stems from the higher valency of scaffold proteins, which allows them to establish higher molecular connectivity within the core condensate and, thus, induce higher enthalpic gain upon multilayered assembly.

Finally, to gain further insight on the droplet liquid network connectivity, we evaluate the average number of engaged binding sites per protein (φ) as a function of distance from the center of mass of the condensates (Fig. 4b). Scaffold proteins present a significantly higher amount of molecular connections per protein than surfactants (i.e., $\varphi \sim 3.7$ and $\varphi \sim 2.5$ for scaffolds and surfactants, respectively), at the droplet core. This observation highlights how surfactants negatively contribute to the stability of the condensate. However at the interface, such diminished connectivity of surfactant proteins ($\varphi \sim 1$) with respect to that of scaffolds ($\varphi \sim 3$) substantially reduces the condensate surface tension—by decreasing the enthalpic cost (Δh_i) of creating an interface. This energetically favorable protein arrangement, controlled in our model just by the variance in protein valency of the components, is expected to be contributed also by changes in relative binding affinities among proteins, and

be modulated by post-translational modifications in vivo[7,34]. Furthermore, such variations can be also relevant to understand the physical parameters controlling multilayered condensate organization and, ultimately, regulation of the formation of size-conserved multidroplet emulsions[29–31].

Conclusions

In this work, we employ our minimal protein model[48] to demonstrate how biomolecular multidroplet emulsification can be controlled by the subtle balance between liquid network connectivity and droplet surface tension, and how general molecular features such as protein valency and binding affinity can critically regulate this behavior. By using a binary mixture of scaffold and client proteins that act as surfactants (following the original definition proposed by Banani et al.[20]), we design a set of Direct Coexistence simulations in which we can conveniently modulate the simulation box geometry to assess the propensity of the condensates to accommodate different surface/volume ratios. The ability (or disability) of these mixtures to adopt different surface/volume ratios, imposed by the box geometry, can be regarded as an indirect measurement of the droplet propensity to grow beyond a certain size. We find that pure component scaffold condensates can easily adapt to distinct surface/volume ratios; in support of their ability to grow and fuse into a single droplet. However, 50:50 binary scaffold–surfactant mixtures stabilize instead several coexisting liquid condensates with roughly constant surface/volume ratios to accommodate the imposed system size geometry. Such behavior is a clear signature of size-conserved multidroplet emulsification, as found in the nucleolus[7], ribonucleoprotein condensates[26,88], micelle-like condensates[31,45], and in vitro complex coacervates[38].

We also elucidate the role of surfactant concentration in size-conserved droplet growth. By gradually decreasing the scaffold–surfactant ratio in our mixtures, we observe that the maximum droplet size is reduced, while simultaneously increasing the number of coexisting condensates. This trend continues until a sufficiently high surfactant concentration is reached, where LLPS is no longer possible. Moreover, as clients are added, the droplet surface tension dramatically decreases, facilitating the formation of multiple coexisting small liquid droplets at low interfacial energetic cost. Client proteins, besides decreasing the stability of the condensates[49], can effectively behave as natural droplet surfactants[35,36]. Due to their considerably lower molecular connectivity compared to that of scaffolds, surfactants preferentially migrate towards the droplet interface; thereby, minimizing the enthalpic cost of creating an interface[23]. Heterogeneous molecular organizations of condensates have been observed in stress granules[75,76], the nucleoli[26] and nuclear speckles[77]. We find that such heterogeneity enable the maximization of the condensate liquid network connectivity through scaffold–scaffold protein interactions within the droplet core.

Rationalizing the underlying mechanisms employed by cells to precisely regulate the size of their diverse membraneless compartments and processing bodies[7,34,36] represents a crucial step towards understanding intracellular spatiotemporal cell organization. Taken together, our coarse-grained simulations help to elucidate the relationship between single-droplet phase formation and size-conserved multidroplet architecture, and put forward general molecular features such as valency and binding affinity as chief drivers in these scenarios.

Methods

We model our coarse-grained multivalent proteins using the MD-Patchy potential proposed in Ref.[48], which is composed by two different set of potentials: a Pseudo Hard-Sphere (PHS) potential[89] to continuously describe the repulsive interaction and excluded volume between different protein replicas, and a continuous square-well (CSW)[90] potential to describe the patch-patch interactions among different protein binding sites. The u_{PHS} potential is described by the following expression:

$$u_{PHS} = \begin{cases} \lambda_r \left(\frac{\lambda_r}{\lambda_a}\right)^{\lambda_a} \varepsilon_R \left[\left(\frac{\sigma}{r}\right)^{\lambda_r} - \left(\frac{\sigma}{r}\right)^{\lambda_a}\right] + \varepsilon_R; & \text{if } r < \left(\frac{\lambda_r}{\lambda_a}\right)\sigma \\ 0; & \text{if } r \geq \left(\frac{\lambda_r}{\lambda_a}\right)\sigma \end{cases} \quad (1)$$

where $\lambda_a = 49$ and $\lambda_r = 50$ are the exponents of the attractive and repulsive terms respectively, ε_R accounts for the energy shift of the PHS interaction, σ is the molecular diameter (and our unit of length) and r is the center-to-center distance between different PHS particles. For the patch–patch interaction we use the following expression:

$$u_{CSW} = -\frac{1}{2}\varepsilon_{CSW}\left[1 - tanh\left(\frac{r - r_w}{\alpha}\right)\right] \quad (2)$$

where ε_{CSW} is the depth of the potential energy well, r_w the radius of the attractive well, and α controls the steepness of the well. We choose $\alpha = 0.005\sigma$ and $r_w = 0.12\sigma$ so that each patch can only interact with another single patch.

The mass of each patch is a 5% of the central PHS particle mass, which is set to 3.32×10^{-26} kg, despite being this choice irrelevant for equilibrium simulations. This 5% ratio fixes the moment of inertia of the patchy particles (our minimal proteins). The molecular diameter of the proteins, both scaffold and clients, is $\sigma = 0.3405$ nm, and the value of ε_R/k_B is 119.81K. All our results are presented in reduced units: reduced temperature is defined as $T^* = k_B T/\varepsilon_{CSW}$, reduced density as $\rho^* = (N/V)\sigma^3$, reduced pressure as $p^* = p\sigma^3/(k_B T)$, and the reduced unit of time as $\sqrt{\sigma^2 m/(k_B T)}$. In order to keep the PHS interaction as similar as possible to a pure HS interaction, we fix $k_B T/\varepsilon_R$ at a value of 1.5 as suggested in Ref.[89] (fixing $T = 179.71$K). We then control the effective strength of the binding protein attraction by varying ε_{CSW} such that the reduced temperature, $T^* = k_B T/\varepsilon_{CSW}$, is of the order of $\mathcal{O}(0.1)$.

Since both u_{PHS} and u_{CSW} potentials are continuous and differentiable, we perform all our simulations using the LAMMPS Molecular Dynamics package[91]. Periodic boundary conditions are used in the three directions of space. The timestep chosen for the Verlet integration of the equations of motion is $\Delta t^* = 3.7 \times 10^{-4}$. The cut-off radius of the interactions of both potentials is set to 1.17σ. We use a Nosé–Hoover thermostat[92,93] for the NVT simulations with a relaxation time of 0.074 in reduced units. For NpT simulations, a Nosé–Hoover barostat is employed with the same relaxation time[94].

The methodological details of the calculation of the phase diagram, surface tension and engaged binding sites per protein through a local order parameter, are provided in the Supporting Information document.

Received: 30 April 2021; Accepted: 6 July 2021
Published online: 27 July 2021

References

1. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid–liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
2. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease (2017).
3. Brangwynne, C. P. *et al.* Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).
4. Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
5. Marnik, E. A. & Updike, D. L. Membraneless organelles: P granules in caenorhabditis elegans. *Traffic* **20**, 373–379 (2019).
6. Lu, H. *et al.* Phase-separation mechanism for c-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558**, 318–323 (2018).
7. Brangwynne, C. P., Mitchison, T. J. & Hyman, A. A. Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci.* **108**, 4334–4339 (2011).
8. Wang, J. *et al.* A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699 (2018).
9. Krainer, G. *et al.* Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun.* **12**, 1–14 (2021).
10. Strom, A. R. *et al.* Phase separation drives heterochromatin domain formation. *Nature* **547**, 241–245 (2017).
11. Larson, A. G. *et al.* Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
12. Su, X. *et al.* Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352**, 595–599 (2016).
13. Yoo, H., Triandafillou, C. & Drummond, D. A. Cellular sensing by phase separation: Using the process, not just the products. *J. Biol. Chem.* **294**, 7151–7159 (2019).
14. Sabari, B. R. *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361** (2018).
15. Kroschwald, S. *et al.* Different material states of Pub1 condensates define distinct modes of stress adaptation and recovery. *Cell Rep.* **23**, 3327–3339 (2018).
16. Bouchard, J. J. *et al.* Cancer mutations of the tumor suppressor SPOP disrupt the formation of active, phase-separated compartments. *Mol. Cell* **72**, 19–36 (2018).
17. Brundin, P. *et al.* Prion-like transmission of protein aggregates in neurodegenerative diseases. *Nat. Rev. Mol. Cell Biol.* **11**, 301–307 (2010).
18. Ambadipudi, S., Biernat, J., Riedel, D., Mandelkow, E. & Zweckstetter, M. Liquid–liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein tau. *Nat. Commun.* **8**, 1–13 (2017).
19. Robberecht, W. & Philips, T. The changing scene of amyotrophic lateral sclerosis. *Nat. Rev. Neurosci.* **14**, 248–264 (2013).
20. Banani, S. F. *et al.* Compositional control of phase-separated cellular bodies. *Cell* **166**, 651–663 (2016). http://www.sciencedirect.com/science/article/pii/S0092867416307395.
21. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
22. Espinosa, J. R. *et al.* Liquid-network connectivity regulates the stability and composition of biomolecular condensates with many components.
23. Sanchez-Burgos, I., Espinosa, J. R., Joseph, J. A. & Collepardo-Guevara, R. Valency and binding affinity variations can regulate the multilayered organization of protein condensates with many components. *Biomolecules* **11**, 278 (2021).
24. Ditlev, J. A., Case, L. B. & Rosen, M. K. Who's in and who's out–compositional control of biomolecular condensates (2018).
25. Shin, Y. *et al.* Liquid nuclear condensates mechanically sense and restructure the genome. *Cell* **175**, 1481–1491 (2018).
26. Feric, M. *et al.* Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686–1697 (2016). http://www.sciencedirect.com/science/article/pii/S0092867416304925.
27. Liu, X. *et al.* Phase separation drives decision making in cell division. *J. Biol. Chem.* **295**, 15172 (2020).
28. Zhang, Y., Li, Z., Chen, N., Huang, Y. & Huang, S. Phase separation of arabidopsis EMB1579 controls transcription, mRNA splicing, and development. *PLoS Biol.* **18**, e3000782 (2020).
29. Ranganathan, S. & Shakhnovich, E. I. Dynamic metastable long-living droplets formed by sticker-spacer proteins. *Elife* **9**, e56159 (2020).
30. Dar, F. & Pappu, R. Phase separation: Restricting the sizes of condensates. *Elife* **9**, e59663 (2020).
31. Nakamura, H., DeRose, R. & Inoue, T. Harnessing biomolecular condensates in living cells. *J. Biochem.* **166**, 13–27 (2019).
32. Binder, K., Block, B. J., Virnau, P. & Tröster, A. Beyond the van der Waals loop: What can be learned from simulating Lennard–Jones fluids inside the region of phase coexistence. *Am. J. Phys.* **80**, 1099–1109 (2012).
33. Zhang, H. *et al.* RNA controls polyQ protein phase transitions. *Mol. Cell* **60**, 220–230 (2015).
34. Kilchert, C., Weidner, J., Prescianotto-Baschong, C. & Spang, A. Defects in the secretory pathway and high Ca2+ induce multiple p-bodies. *Mol. Biol. Cell* **21**, 2624–2638 (2010).
35. Cuylen, S. *et al.* Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. *Nature* **535**, 308–312 (2016).
36. Navarro, M.G.-J. *et al.* RNA is a critical element for the sizing and the composition of phase-separated RNA-protein condensates. *Nat. Commun.* **10**, 1–13 (2019).
37. Wegmann, S. *et al.* Tau protein liquid–liquid phase separation can initiate tau aggregation. *EMBO J.* **37**, e98049 (2018).
38. Lu, T. & Spruijt, E. Multiphase complex coacervate droplets. *J. Am. Chem. Soc.* **142**, 2905–2914 (2020).
39. Lovell, P. A. & Schork, F. J. Fundamentals of emulsion polymerization. *Biomacromolecules* **21**, 4396–4441 (2020).
40. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and challenges in studying liquid–liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 (2019).
41. Wurtz, J. D. & Lee, C. F. Chemical-reaction-controlled phase separated drops: Formation, size selection, and coarsening. *Phys. Rev. Lett.* **120**, 078102 (2018).

42. Zwicker, D., Seyboldt, R., Weber, C. A., Hyman, A. A. & Jülicher, F. Growth and division of active droplets provides a model for protocells. *Nat. Phys.* **13**, 408–413 (2017).
43. Wurtz, J. D. & Lee, C. F. Stress granule formation via ATP depletion-triggered phase separation. *N. J. Phys.* **20**, 045008 (2018).
44. Kirschbaum, J., & Zwickerhttps, D. Controlling biomolecular condensates via chemical reactions. *J. R. Soc. Interface.* **18**, 20210255. https://royalsocietypublishing.org/doi/full/10.1098/rsif.2021.0255 (2021).
45. Ruff, K. M., Harmon, T. S. & Pappu, R. V. Camelot: A machine learning approach for coarse-grained simulations of aggregation of block-copolymeric protein sequences. *J. Chem. Phys.* **143**, 12B607 (2015).
46. Tejedor, A. R., Garaizar, A., Ramírez, J. & Espinosa, J. R. Dual RNA modulation of protein mobility and stability within phase-separated condensates. *bioRxiv* (2021).
47. Kelley, F. M., Favetta, B., Regy, R. M., Mittal, J. & Schuster, B. S. Amphiphilic proteins coassemble into multiphasic condensates and act as biomolecular surfactants. *bioRxiv* (2021).
48. Espinosa, J. R., Garaizar, A., Vega, C., Frenkel, D. & Collepardo-Guevara, R. Breakdown of the law of rectilinear diameter and related surprises in the liquid-vapor coexistence in systems of patchy particles. *J. Chem. Phys* **150**, 224510 (2019).
49. Espinosa, J. R. *et al.* Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components. *Proc. Natl. Acad. Sci.* **117**, 13238–13247 (2020).
50. Fisher, R. S. & Elbaum-Garfinkle, S. Tunable multiphase dynamics of arginine and lysine liquid condensates. *Nat. Commun.* **11**, 1–10 (2020).
51. Dignon, G. L., Best, R. B. & Mittal, J. Biomolecular phase separation: From molecular driving forces to macroscopic properties. *Annu. Rev. Phys. Chem.* **71**, 53–75 (2020).
52. Choi, J.-M., Holehouse, A. S. & Pappu, R. V. Physical principles underlying the complex biology of intracellular phase transitions. *Annu. Rev. Biophys.* **49**, 107–133 (2020).
53. Lee, J., Popov, Y. O. & Fredrickson, G. H. Complex coacervation: A field theoretic simulation study of polyelectrolyte complexation. *J. Chem. Phys.* **128**, 224908 (2008).
54. Weber, C. A., Zwicker, D., Jülicher, F. & Lee, C. F. Physics of active emulsions. *Rep. Prog. Phys.* **82**, 064601 (2019).
55. Harmon, T. S., Holehouse, A. S. & Pappu, R. V. Differential solvation of intrinsically disordered linkers drives the formation of spatially organized droplets in ternary systems of linear multivalent proteins. *N. J. Phys.* **20**, 045002 (2018).
56. Choi, J.-M., Dar, F. & Pappu, R. V. Lassi: A lattice model for simulating phase transitions of multivalent proteins. *PLOS Comput. Biol.* **15**, e1007028 (2019).
57. Statt, A., Casademunt, H., Brangwynne, C. P. & Panagiotopoulos, A. Z. Model for disordered proteins with strongly sequence-dependent liquid phase behavior. *J. Chem. Phys.* **152**, 075101 (2020).
58. Jacobs, W. M. & Frenkel, D. Phase transitions in biological systems with many components. *Biophys. J.* **112**, 683–691 (2017).
59. Russo, J., Tartaglia, P. & Sciortino, F. Reversible gels of patchy particles: Role of the valence. *J. Chem. Phys.* **131**, 014504 (2009).
60. Blas, F. J., MacDowell, L. G., De Miguel, E. & Jackson, G. Vapor–liquid interfacial properties of fully flexible Lennard–Jones chains. *J. Chem. Phys.* **129**, 144703 (2008).
61. 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 (2018).
62. Das, S., Lin, Y.-H., Vernon, R. M., Forman-Kay, J. D. & Chan, H. S. Comparative roles of charge, π, and hydrophobic interactions in sequence-dependent phase separation of intrinsically disordered proteins. *Proc. Natl. Acad. Sci.* **117**, 28795–28805 (2020).
63. Alshareedah, I., Moosa, M. M., Raju, M., Potoyan, D. A. & Banerjee, P. R. Phase transition of RNA–protein complexes into ordered hollow condensates. *Proc. Natl. Acad. Sci.* **117**, 15650–15658 (2020).
64. Latham, A. P. & Zhang, B. Consistent force field captures homolog resolved hp1 phase separation. *bioRxiv* (2021).
65. Regy, R. M., Dignon, G. L., Zheng, W., Kim, Y. C. & Mittal, J. Sequence dependent phase separation of protein–polynucleotide mixtures elucidated using molecular simulations. *Nucleic Acids Res.* **48**, 12593–12603 (2020).
66. Sanchez-Burgos, I., Espinosa, J. R., Joseph, J. A. & Collepardo-Guevara, R. Valency and binding affinity variations can regulate the multilayered organization of protein condensates with many components. *Biomolecules* **11** (2021). https://www.mdpi.com/2218-273X/11/2/278.
67. Joseph, J. A. *et al.* Thermodynamics and kinetics of phase separation of protein–RNA mixtures by a minimal model. *Biophys. J.* **120**, 1–12 (2021).
68. Jover, J., Haslam, A., Galindo, A., Jackson, G. & Müller, E. Pseudo hard-sphere potential for use in continuous molecular-dynamics simulation of spherical and chain molecules. *J. Chem. Phys.* **137**, 144505 (2012).
69. Ladd, A. J. & Woodcock, L. V. Triple-point coexistence properties of the Lennard–Jones system. *Chem. Phys. Lett.* **51**, 155–159 (1977).
70. García Fernández, R., Abascal, J. L. & Vega, C. The melting point of ice ih for common water models calculated from direct coexistence of the solid–liquid interface. *J. Chem. Phys.* **124**, 144506 (2006).
71. Espinosa, J. R., Sanz, E., Valeriani, C. & Vega, C. On fluid-solid direct coexistence simulations: The pseudo-hard sphere model. *J. Chem. Phys.* **139**, 144502 (2013).
72. Rowlinson, J. S. & Widom, B. *Molecular Theory of Capillarity* (Courier Corporation, 2013).
73. Mountain, G. A. & Keating, C. D. Formation of multiphase complex coacervates and partitioning of biomolecules within them. *Biomacromolecules* (2019).
74. Boeynaems, S. *et al.* Spontaneous driving forces give rise to protein–RNA condensates with coexisting phases and complex material properties. *Proc. Natl. Acad. Sci.* **116**, 7889–7898 (2019).
75. Jain, S. *et al.* ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell* **164**, 487–498 (2016).
76. Guillén-Boixet, J. *et al.* RNA-induced conformational switching and clustering of G3BP drive stress granule assembly by condensation. *Cell* **181**, 346–361 (2020).
77. Fei, J. *et al.* Quantitative analysis of multilayer organization of proteins and RNA in nuclear speckles at super resolution. *J. Cell Sci.* **130**, 4180–4192 (2017).
78. Sanders, D. W. *et al.* Competing protein–RNA interaction networks control multiphase intracellular organization. *Cell* **181**, 306–324 (2020).
79. Kaur, T. *et al.* Sequence-encoded and composition-dependent protein–RNA interactions control multiphasic condensate topologies. *bioRxiv* (2020).
80. Liu, H. & Cao, G. Effectiveness of the Young–Laplace equation at nanoscale. *Sci. Rep.* **6**, 1–10 (2016).
81. Tolman, R. C. The effect of droplet size on surface tension. *J. Chem. Phys.* **17**, 333–337 (1949).
82. Espinosa, J. R., Vega, C. & Sanz, E. Homogeneous ice nucleation rate in water droplets. *J. Phys. Chem. C* **122**, 22892–22896 (2018).
83. Montero de Hijes, P., Espinosa, J. R., Sanz, E. & Vega, C. Interfacial free energy of a liquid-solid interface: Its change with curvature. *J. Chem. Phys.* **151**, 144501 (2019).
84. Sanchez-Burgos, I., de Hijes, P. M., Rosales-Pelaez, P., Vega, C. & Sanz, E. Equivalence between condensation and boiling in a Lennard–Jones fluid. *Phys. Rev. E* **102**, 062609 (2020).
85. Bianchi, E., Largo, J., Tartaglia, P., Zaccarelli, E. & Sciortino, F. Phase diagram of patchy colloids: Towards empty liquids. *Phys. Rev. Lett.* **97**, 168301 (2006).
86. Mazarakos, K. P., Ghosh, A., Zhang, X., Nguemaha, V. & Zhou, H.-X. Affinity between macromolecular regulators leads to precise control of liquid–liquid phase separation. *Biophys. J.* **118**, 202a (2020).

87. Alshareedah, I., Thurston, G. M. & Banerjee, P. R. Quantifying viscosity and surface tension of multicomponent protein-nucleic acid condensates. *Biophys. J.* (2021).
88. Cochard, A. *et al.* RNA at the surface of phase-separated condensates impacts their size and number (2021).
89. Jover, J., Haslam, A. J., Galindo, A., Jackson, G. & Müller, E. A. Pseudo hard-sphere potential for use in continuous molecular-dynamics simulation of spherical and chain molecules. *J. Chem. Phys.* **137**, 144505 (2012).
90. Espinosa, J., Vega, C. & Sanz, E. The mold integration method for the calculation of the crystal-fluid interfacial free energy from simulations. *J. Chem. Phys.* **141**, 134709 (2014).
91. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
92. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **81**, 511–519 (1984).
93. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697. https://doi.org/10.1103/PhysRevA.31.1695 (1985).
94. Hoover, W. G. Constant-pressure equations of motion. *Phys. Rev. A* **34**, 2499 (1986).

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No 803326). J. R. E. acknowledges funding from the Oppenheimer Fellowship and from Emmanuel College Roger Ekins Research Fellowship. I. S. B. acknowledges funding from the Oppenheimer Fellowship, EPRSC Doctoral Programme Training number EP/T517847/1 and Derek Brewer Emmanuel College scholarship. J. A. J. is a Junior Research Fellow at Kings College. This work has been performed using resources provided by the Cambridge Tier-2 system operated by the University of Cambridge Research Computing Service (http://www.hpc.cam.ac.uk) funded by EPSRC Tier-2 capital grant EP/P020259/1.

Author contributions

I.S.B. and J.R.E. conceived the project. I.S.B. conducted the simulations. I.S.B. and J.R.E. analyzed the results. J.A.J. and R.C.G. contextualized the results. I.S.B. and J.R.E. wrote the initial version of the manuscript. All the authors reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available athttps://doi.org/10.1038/s41598-021-94309-y.

Correspondence and requests for materials should be addressed to J.R.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.