Supporting Information for: "Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components"

Jorge R. Espinosa,[1] Jerelle A. Joseph,[1] Ignacio Sanchez-Burgos,[1] Adiran Garaizar,[1] Daan Frenkel,[2] and Rosana Collepardo-Guevara[1, 2, 3, *]

[1] *Maxwell Centre, Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge, CB3 0HE, United Kingdom*
[2] *Department of Chemistry, University of Cambridge, Cambridge, CB2 1EW, United Kingdom*
[3] *Department of Genetics, University of Cambridge, Cambridge, CB2 3EH, United Kingdom*

* rc597@cam.ac.uk

www.pnas.org/cgi/doi/10.1073/pnas.1917569117

I. SINGLE-MOLECULE PROTEIN CHARACTERISTICS AFFECTING THE STABILITY OF BIOMOLECULAR CONDENSATES

Here, we investigate how the single-component phase diagrams of different multivalent proteins (Fig. S1; also see pure systems listed in Table S1) change as we vary the number of binding sites on the protein surface ('protein valency'), the strength of binding among proteins, the nature of the binding sites from promiscuous (i.e. able to interact with any other site) to selective (i.e. restricted to interact with specific sites only), and the topological distribution of the binding sites on the protein surface. To account for cooperative and competing effects, we assess the phase behavior of single-component systems formed by thousands of copies of the same protein.

We find that the single-component protein phase diagrams are most strongly affected by changes in the protein valency (Fig. S1b). The liquid–liquid coexistence region decreases significantly as the protein valency decreases until it disappears completely for proteins with a valency of two [1] (Fig. S1b). These results are consistent with experimental observations showing that the minimum concentration required to observe protein liquid–liquid phase separation (LLPS) decreases significantly as the valency of the protein increases [2–5]. The results also agree well with simulated phase diagrams of colloidal patchy particles showing that the liquid–vapor equilibrium region shrinks as valency decreases [1, 6, 7].

Besides multivalency, we find that the ability of a protein to phase separate depends strongly on the relative binding strength of its binding sites. Gradually decreasing the strength of one of the three binding sites in our trivalent proteins to 50% (here termed 'Valency: 2.5'), or even further to just 25% (termed 'Valency: 2.25'), of its original value destabilizes liquid demixing significantly (Fig. S1b).

As was shown in Ref. [7] for patchy colloids, the topological distribution of binding sites on the protein surface also influences droplet condensation significantly (Fig. S2; top). A good topology from the point of view of LLPS—i.e. one that enhances the size of the liquid–liquid coexistence region—is one that maximizes the angular distances between the various binding directions on the protein surface, thereby reducing steric clashes between different pair interactions, and facilitating a higher connectivity of the liquid network. However, besides sterically hindering multivalent interactions, certain arrangements of the binding sites can disfavor LLPS at the expense of forming fully-saturated (and hence inert) protein complexes that cannot interact with the network in the condensed liquid [7]. As the valency increases, we would expect that flexible biomolecules (such as intrinsically disordered proteins, RNA or DNA) can achieve a uniform angular distribution of binding directions more easily than their rigid counterparts. We speculate that this ability to overcome steric hindrance is one of the reasons why proteins containing intrinsically disordered regions (IDRs) are favored in biomolecular condensates.

Another feature that has been suggested to favor LLPS is promiscuous over specific binding [8]. In Ref. [7], it was shown that limiting the binding sites to engage in specific binding, while having an excess of the specific cognates, has almost a negligible effect; shrinking only modestly the liquid-liquid coexistence region when compared to promiscuous binding (Fig. S2, bottom) [7]. However, we expect suppression of LLPS by specific binding to be greatly amplified in vivo. This is because inside cells specific cognate species are likely to be present in limited amounts, as in our example, which would result in the inability of some proteins to recruit their binding partners and, hence, an effective reduction of the valency of these proteins due to shortage of adequate binding partners.

II. COMPOSITION OF SCAFFOLD–CLIENT BIOMOLECULAR CONDENSATES

We observe that in binary biomolecular condensates the partitioning coefficient of scaffolds is always higher than that of the clients (Fig. S3a), even when clients bind with high affinity to the scaffolds. It is reasonable that scaffolds, due to their higher valencies, are present at higher concentrations in the condensed phase because they can establish more attractive interactions per unit of volume than their lower-valency counterparts.

Next, we compare the percentage of clients inside the condensates for the different types of scaffold–client mixtures we study (Fig. S3b; Mixtures 1–6 are described in Table S2). We find that partitioning of clients is heavily determined by the scaffold–client binding affinity. High-affinity (HA) clients competing for scaffold–scaffold binding sites are present inside the condensates at concentrations very near to the original mixing ratios (see Mixtures 1, 2, 5 and 6 in Fig. S3b). As expected, as the interaction strengths weaken the concentration of clients in the condensate decreases concomitantly. However, when the high-affinity clients bind to alternate sites in the scaffold (Mixture 5 in Fig. S3b), they remain fully recruited even at weak protein–protein attractions.

For the case in which clients have a higher valency than scaffolds (Mixture 6 in Fig. S3b), their concentration in the droplet increase with respect to the initial mixing ratio when approaching to the critical temperature. Low-affinity clients, which are poor competitors for scaffold–scaffold interactions, are present at much lower concentrations inside the condensates than the initial mixing ratios. This result demonstrates that partitioning of clients relies on a

sensitive balance of many factors, including the ability of clients to compete for scaffold–scaffold binding sites, and the stoichiometry in the mixture, but seems to be most strongly affected by the scaffold–client binding affinity.

III. PATCHY PARTICLE MODEL FOR PROTEIN LLPS

A. Minimal model for multivalent protein mixtures

In our simulations, we investigate the phase behavior of thousands of multivalent proteins that interact weakly with one another and undergo liquid–liquid phase separation (LLPS) under certain conditions of temperature (or inverse interaction strength) and protein concentration (measured in terms of volume fractions). To be able to describe the collective behavior of thousands of interacting proteins, we represent each individual multivalent protein as a patchy particle using our recently proposed molecular dynamics (MD) based patchy particle model [7]. In this model, each protein is treated as a pseudo hard sphere with a diameter σ and is decorated with M attractive sites ("patches") on its surface. Each sticky patch represents a binding site on the protein—capturing general features of protein–protein interactions, such as electrostatic and hydrophobic attraction, and hydrogen bonding—that allows the protein to engage in one single intermolecular (protein–protein) interaction; thus, M is a measure of the valency of the protein.

To model the pseudo-hard spheres protein core, we use a potential (v_{core}) of the form proposed in Ref. [9].

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

where λ_a=49 and λ_r=50 are the exponents of the attractive and repulsive terms, respectively, chosen for computational convenience. The interaction strength ε_R accounts for the energy of the pseudo hard-sphere interaction and σ is our unit of length. The term r denotes the center-to-center distance between protein cores.

For the attractive sites or sticky patches on the protein surface, we use the following continuous attractive square-well interaction proposed in Ref. [10]:

$$v_{prot-prot}(r) = -\frac{1}{2}\varepsilon_{prot-prot}\left[1 - \quad \tanh\left(\frac{r - r_w}{\alpha}\right)\right]; \tag{2}$$

where r is the distance between the centers of two attractive patches, r_w is the radius of the attractive well, and α controls the steepness of the well. $\varepsilon_{prot-prot}$ is the unit of energy in our simulations.

An advantage of our approach is that all the potentials are continuous and differentiable and, hence, MD simulations can be easily employed to sample phase space. Accordingly, we perform massively parallel MD simulations via LAMMPS [11]. In the implementation, we use a multi-centered rigid body composed of M+1 particles per protein: one central particle to account for the pseudo-hard sphere protein core and M particles rigidly bound to the protein core to represent the protein–protein binding sites.

Although the choice of the mass is irrelevant for equilibrium simulations, we choose the masses of the patches to be equal to 1/20 of the mass of the protein core. We also define α=0.005σ and r_w=0.12σ. This small value of r_w limits each sticky patch to interact with at most one other patch at a time, ensuring that the maximum number of multivalent interactions that each protein can engage in is equal to M.

In order to keep the isotropic pseudo-hard sphere interaction in our simulations as similar as possible to a pure hard sphere interaction, we follow Ref. [7] and fix k_BT/ε_R at a value of 1.5 [9, 12], where ε_R=119.81K. We then control the effective strength of the attraction by varying $\varepsilon_{prot-prot}$, such that the reduced temperature of the system is defined as $T^* = k_BT/\varepsilon_{prot-prot}$. We also use reduced units for density (ρ^* =(N/V)σ^3), volume fraction ($\phi = \frac{\pi}{6}\rho^*$), pressure ($p^* = p\sigma^3/k_BT$), and time ($\sqrt{\sigma^2 m/(k_BT)}$).

B. Liquid–liquid coexistence simulations

To evaluate the phase diagrams, we use the Direct Coexistence (DC) simulation method [12–14] and MD simulations. The DC method simulates coexistence by preparing periodically extended slabs of two phases, i.e. the diluted liquid (protein depleted phase) and the condensed liquid (protein enriched phase), in the same simulation box. We prepare the initial configurations, using the procedure of Ref. [7]. Specifically, we first equilibrate the condensed-liquid phase in an NpT-ensemble MD simulation at a reduced pressure p^* equal to 0 and at high $\varepsilon_{prot-prot}$ value (or low T^*) (where the equilibrium vapor pressure is negligible) using a cubic box. After the condensed liquid has been properly equilibrated, we elongate the periodic box in one direction (e.g. in x) by flanking the original simulation box by two

empty cubic boxes. We refer to this elongated side as the "long side of the simulation box", and denote it by L. We then perform constant NVT-ensemble MD simulations at a fixed temperature and different values of the inverse protein–protein interaction strength $\varepsilon_{prot-prot}$ (or equivalently, since $T^* = k_BT/\varepsilon_{prot-prot}$, at a fixed value of the protein–protein interaction strength and varying temperatures).

For each value of $\varepsilon_{prot-prot}$, once the system has reached equilibrium, we determine whether the system has formed a single homogeneous phase or two coexisting liquid phases. For cases where two phases coexist, we estimate the equilibrium densities/volume fractions of each phase by computing a density profile along the long side of the box. An example of this procedure is shown in Fig. S4. This procedure allows us to define a coexistence curve in the space of inverse interaction strength or temperature versus volume fraction that separates the region of the phase diagram—below which the system undergoes liquid demixing, and above which no phase separation is observed (Fig. S1). We denote the critical temperature—the maximum temperature that still favors demixing—as T^C and the critical inverse interaction strength as $\varepsilon^C_{prot-prot}$.

To estimate the critical temperature (T^C) of our phase diagrams, we fit the volume fraction difference between the coexisting diluted liquid and condensed liquid to the expression:

$$(\phi_c(T^*) - \phi_d(T^*))^{3.06} = d(1 - \frac{T^*}{T^C}); \tag{3}$$

where ϕ_c and ϕ_d are the volume fractions of the condensed and diluted phases respectively, and d is a fitting parameter. The critical volume fraction, ϕ^C, is estimated by assuming that the law of rectilinear diameter holds close to T^C:

$$(\phi_c + \phi_d)/2 = \phi^C + s_2(T^C - T^*); \tag{4}$$

where s_2 is a fitting parameter. We note that our estimates of the critical point are approximate and could be subjected to finite size effects. However, we did not attempt to carry out a systematic finite-size scaling analysis, since we are mainly interested in the qualitative features of the phase diagrams. The critical temperatures and packing fractions estimated for all the studied systems are summarized in Tables S1–S4.

To evaluate convergence of our simulations, we plot the potential energy (U) as a function of time and check when U has converged. In Figure S5, we plot U against time for Mixture 1 at the corresponding temperatures shown in Fig. S4. In all cases, convergence is achieved; in general, simulations conducted at lower temperatures took longer to equilibrate.

To determine the phase diagrams of all binary mixtures, we have used systems of 2000 patchy particles in total, containing different amounts of each type of patchy particle according to the mixture composition. For Mixture 1, we also modeled a system of 3000 patchy particles, to test whether our simulations were affected by finite size effects. In all cases, the results for systems of 2000 and 3000 patchy particles were the same, within the uncertainty. For the 6-component mixtures, we used a system composed of 6000 patchy particles: for the mixture with equal proportions of all types we used 1000 particles each, and for the one with 67% M4 proteins we used 1330 M4 proteins and 134 of every other type.

In the MD simulations, the timestep chosen for the Verlet integration of the equations of motion was Δt=0.0004 in reduced units (corresponds to 0.5 fs in the LAMMPS input file, if we make the (arbitrary) choice $T = 179.71K$). The cut-off radius for both hard-sphere and protein–protein interactions is 1.175 σ. We use the Nosé-Hoover thermostat [15, 16] for NVT simulations with a typical relaxation time of 0.125 in reduced units. For NPT simulations, the Nosé-Hoover barostat was used [17] (in combination with the Nosé-Hoover thermostat) with a typical relaxation time of 0.45 in reduced units. To account for the rotational motion of the patchy particles, we described the colloidal particles as rigid bodies, using the method implemented in LAMMPS [18] (see Ref. [7]).

C. Order parameters to probe the mechanism of liquid-liquid phase separation

To monitor condensate formation, we define a condensation order parameter $Q^X(L)$ as a function of the position of the proteins in the long axis of the simulation box and evaluate it independently for scaffolds and clients (X denotes either scaffolds, S, or clients, C) at different times during our simulations. $Q^X(L)$ measures the local increase in protein concentration relative to a reference case where all the proteins are homogeneously distributed across the simulation box. When a protein (scaffold or client) begins to aggregate at a position x in the box, its local concentration increases (i.e. $Q^X(L) > 0$ indicates condensation). When the proteins are not aggregated, their concentrations are equal to or smaller than their reference concentration (ϕ_0), the order parameter is set to zero (i.e. $Q^X(L) = 0$ indicates no condensation):

$$Q^X(L) = \begin{cases} 0, & \text{for } \phi^X \leq \phi_0 \\ \frac{\phi^X}{\phi^0} - 1, & \text{for } \phi^X > \phi^0 \end{cases} \tag{5}$$

In addition, we compute $\chi_{S/C}$ the ratio of the integrals of the condensation order parameter across the box for scaffolds over clients, to measure the relative proportion of scaffolds and clients in the condense phase:

$$\chi_{S/C} = \frac{\int_0^L Q^S(L)dL}{\int_0^L Q^C(L)dL} \tag{6}$$

Thus, increasing values of $\chi_{S/C}$ describe a protein cluster with an increasing proportion of scaffold proteins.

IV. HIGH-RESOLUTION COARSE-GRAINED MODEL FOR PROTEIN LLPS

To estimate the phase diagrams of the low complexity domain (LCD) of FUS protein and two mutants (see sequences in § VI), we use the hydrophobic-scale (HPS) sequence-dependent protein coarse-grained model of the Mittal group [19]. This model represents each amino acid residue with a sphere, with associated mass, charge, and Lennard-Jones parameters based on the amino acid identity. An implicit solvent modeled via the Debye-Hückel potential is also included. The Lenard-Jones parameters were defined based on a relative hydrophobicity scale of amino acids [20] and optimized to reproduce experimental radii of gyration, R_g, for a large set of proteins.

To compute the phase diagrams we perform Direct Coexistence simulations at constant volume and temperature, as described in § III. For each simulation, we use 80 copies of the same protein, and sample temperatures between 300K and 370 K. For the production runs, we simulate each system for $0.5\,\mu$s using a Langevin thermostat with a damping parameter of $5\,ps^{-1}$ and a time step of 10 fs . We use LAMMPS [11] to carry out all the simulations. The critical temperatures for unmodified FUS LCD, FUS LCD with 7 tyrosine-to-serine mutations, and FUS LCD with 14 tyrosine-to-serine mutations, were 364, 351, and 343 K, respectively.

V. ESTIMATION OF FUS LCD VALENCY

We define the valency of a protein within the condensed phase as the average number of protein neighbors that are "in contact" with it. Two proteins are in contact if at least five of their amino acid residues are separated by a distance of 6.5 Å or closer. We have defined this criterion given that a single amino acid–amino acid contact within the Dignon et al. force field [19] typically contributes ∼0.15-0.3 k_BT to the energy (at T=320 K, which is the temperature at which we evaluate valency). Hence, the proximity between two proteins only results in a significant favorable contribution to the enthalpy of the system, with respect to the energy scale for thermal fluctuations, when at least five contacts are made.

VI. AMINO ACID SEQUENCES OF PROTEINS USED IN RESIDUE-RESOLUTION COARSE-GRAINED SIMULATIONS

1. *Unmodified FUS LCD protein*

MASNNYTQQA TQSYGAYPTQ PGQGYSQQSS QPYGQQSYSG
YSQSTNTSGY GQSSYSSYGQ SQNTGYGTQS TPQGYGSTGG
YGSSQSSQSS YGQQSSYPGY GQQPAPSSTS GSYGSSSQSS
SYGQPQSGSY SQQPSYGGQQ QSYGQQQSYN PPQGYGQQNQ
YNS

2. *FUS LCD with seven tyrosines in GYG, GYS, SYG or SYS sequences randomly selected and mutated (red) to serine.*

MASNNYTQQA TQSSGAYPTQ PGQGYSQQSS QPYGQQSYSG
SSQSTNTSGY GQSSYSSSGQ SQNTGYGTQS TPQGYGSTGG

SGSSQSSQSS YGQQSSYPGY GQQPAPSSTS GSSGSSSQSS
SYGQPQSGSS SQQPSYGGQQ QSSGQQQSYN PPQGYGQQNQ
YNS

3. FUS LCD with seven additional tyrosines (14 in total) in GYG, GYS, SYG or SYS sequences randomly selected and mutated (red) to serine.

MASNNYTQQA TQSSGAYPTQ PGQGSSQQSS QPYGQQSSSG
SSQSTNTSGS GQSSSSSYGQ SQNTGYGTQS TPQGSGSTGG
SGSSQSSQSS SGQQSSYPGY GQQPAPSSTS GSSGSSSQSS
SSGQPQSGSS SQQPSSGGQQ QSSGQQQSYN PPQGYGQQNQ
YNS

VII. STATISTICAL UNCERTAINTY IN THE PHASE DIAGRAMS

We have quantified the uncertainty of our calculations when evaluating the coexistence densities/packing fractions of the reported phase diagrams. Given that for computing the coexistence densities we evaluate directly the density/packing fraction profiles along the long axis of our simulation boxes and we average the bulk density of each phase along many independent configurations, our only font of error is a statistical one rather than a systematic one (i.e. if for evaluating those densities we would have used an order parameter). In Table S5 we show the typical values of the statistical uncertainty associated to our coexistence densities and packing fractions shown in this work.

VIII. LIQUID NETWORK CONNECTIVITY

We define the connectivity of the liquid-network as the average number of bonds per protein per unit of volume within the condensed phase. To estimate this number from our sequence-dependent simulations, we first compute the average number of other proteins that each protein binds to within the condensed phase (as described in § V). We then combine this average number of bonds per protein with the computed coexistence liquid densities for each variant of FUS LCD at T/T_c=0.9 in the high-density liquid branch of their phase diagrams, to obtain the connectivity over the average volume that a single protein shows. The results for the connectivity for the three FUS LCD variants are reported in Table S6. For our simulations of minimal proteins, we also estimate the connectivity of the condensed liquid network by measuring the average number of protein–protein interactions per unit of volume (cubic molecular diameter). The values are given in Tables S8 and S7.

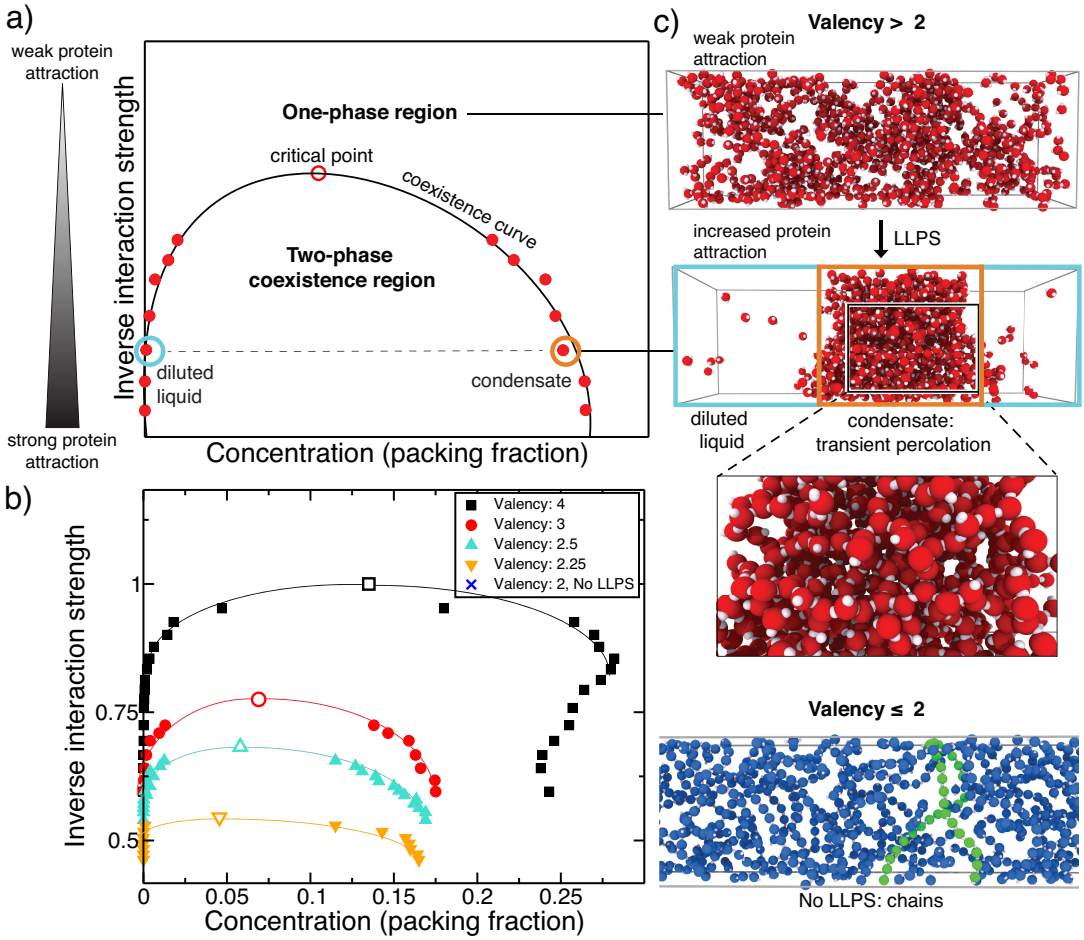

FIG. S1: Role of protein valency in the phase-behavior of single-component protein systems. (a) Illustration of the phase diagrams computed throughout this work. These diagrams explore the space of inverse interaction strength versus the packing fraction (a measure of protein concentration). (b) Liquid–liquid coexistence curves for proteins as a function of valency. The lines are shown as a visual guide to facilitate comparison of data points. The values of the inverse interaction strength presented throughout this work have been normalized by dividing them over the critical value for the 4-valency protein system. Note that the maximum and minimum in density for the 4-valency tetrahedral protein is a real feature of the system as shown in Ref. [7]. Because 2-valency proteins do not undergo LLPS, no coexistence curve is shown for such system. (c) Simulation snapshots for a 3-valency protein (red spheres) at supercritical values of the inverse interaction strength (top), coexistence of a condense phase with a dilute liquid at subcritical values of the inverse interaction strength for the same system (middle), and the well-mixed liquid obtained for the 2-valency protein (blue spheres); to highlight the formation of chains, we have colored a few 2-valency proteins within a chain in green (bottom).

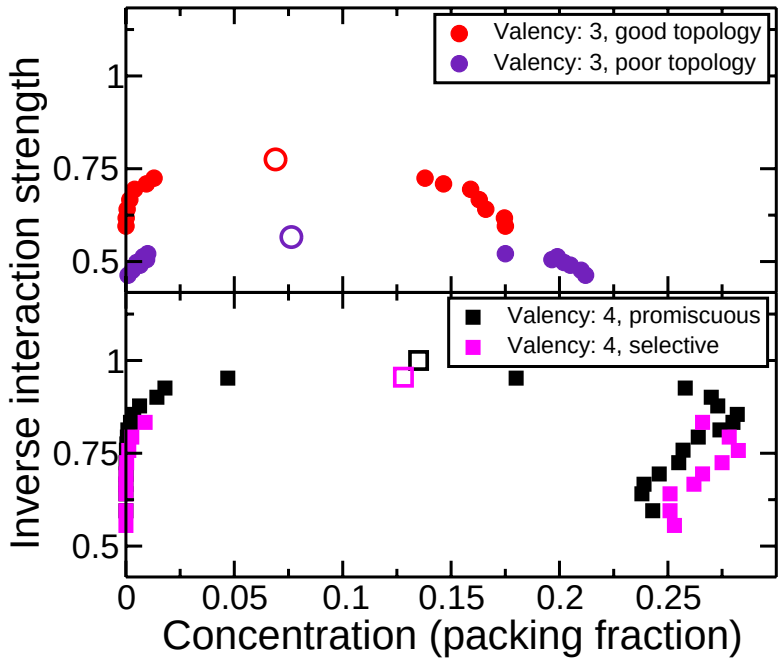

FIG. S2: Role of other protein binding-site characteristics in the phase-behavior of single-component protein systems. (Top) Comparison of the liquid–liquid coexistence curve for two three-valency protein systems with different topologies: (red circles) a "good" topology (i.e. that favours multivalency and LLPS) with three patches in an equatorial plane arrangement with angles between the vectors connecting the patches and the center of the protein of 120°, and (purple circles) a "poor" topology (i.e. imposing more steric hindrance of multivalent interactions) with angles between vectors of 90°. (Bottom) Comparison of the liquid–liquid coexistence curves for two different proteins with four attractive sites in a tetrahedral arrangement: (black squares) promiscuous binding (where all attractive sites can interact with one another) and (magenta squares) selective binding (with two different sites, A and B, and only sites of different types, i.e. A–B, are allowed to interact with one another.)

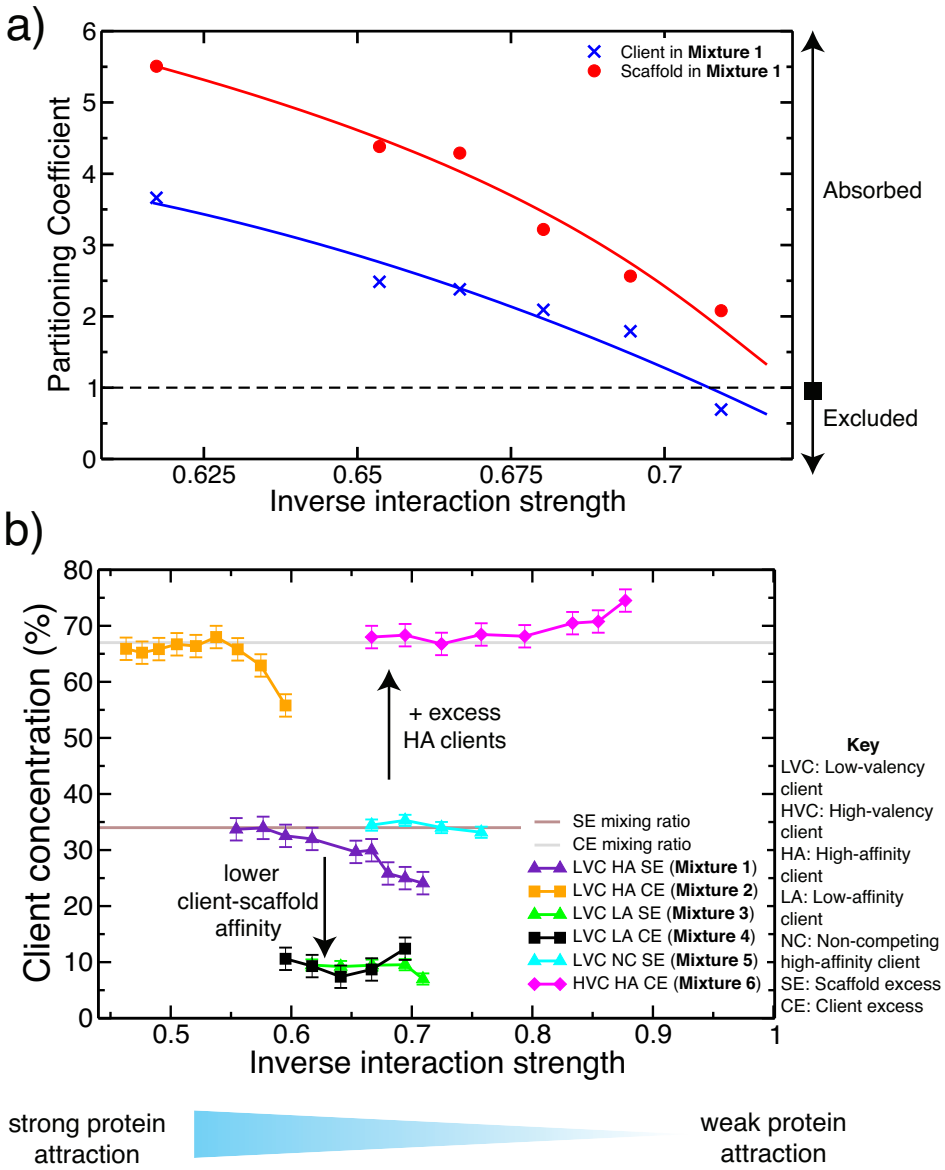

FIG. S3: Composition of scaffold–client condensates as a function of the client ability to compete for scaffold-scaffold interactions or bind to alternative binding sites. We compare six different types of binary mixtures. The description of the mixtures is given in Table S2. (a) Partitioning coefficient for scaffolds and clients for Mixture 1 as a function of the inverse interaction strength. (b) Concentration of clients inside the condensate as a function of the inverse interaction strength computed for different client–scaffold binding modes.

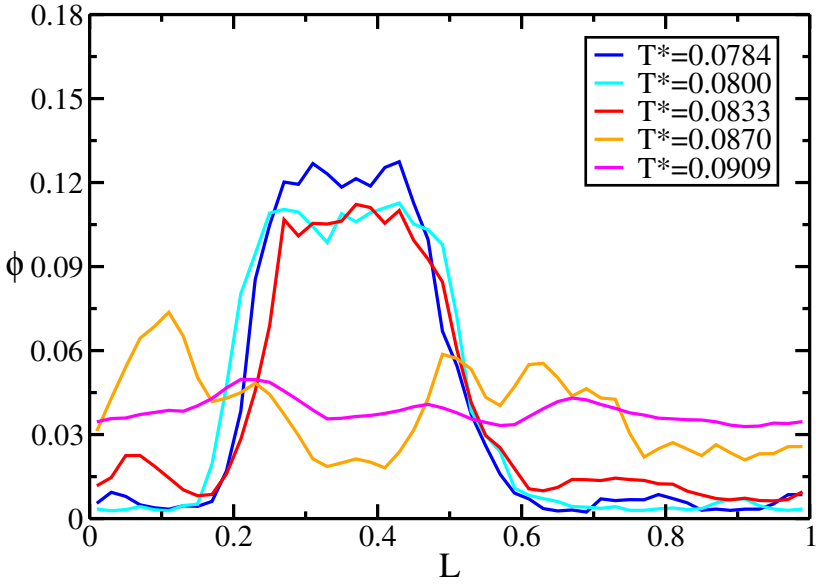

FIG. S4: Packing fraction profiles along the long side of the box for different temperatures of Mixture 1. No phase separation was observed for the two higher temperatures. The three lower ones clearly show phase separation. Packing fraction profiles have been computed without discriminating between the two different types of proteins.

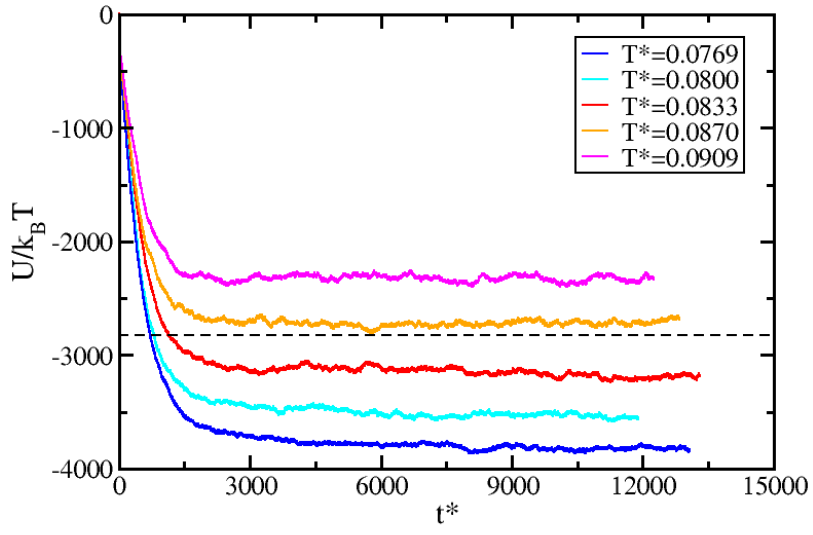

FIG. S5: Potential energy U as a function of time t^* for Mixture 1 at different temperatures. Curves above the black dashed line do not show phase separation, while those below do.

TABLE S1: Critical temperatures (T^C) and critical volume fractions (ϕ^c) estimated via equations (3) and (4) (described in § III) for all the single-component systems we study. Proteins are self-interacting, hence, LLPS is driven by homotypic interactions.

System type	Composition	T^C	ϕ^c
4-valency protein, promiscuous (M4)	100% M4	0.120	0.135
4-valency protein, selective (M4s)	100% M4s	0.113	0.13
3-valency protein, good topology (M3)	100% M3	0.093	0.069
3-valency protein, poor topology (M3p)	100% M3p	0.068	0.076
2.25-valency protein (M2.25)	100% M2.25	0.065	0.046
2-valency protein (M2)	100% M2	—no LLPS—	

TABLE S2: Critical temperatures (T^C) and critical volume fractions (ϕ^c) estimated via equations (3) and (4) (described in § III) for the scaffold–client system in Figures 3 and 4. Divalent proteins exhibit client-client interactions but do not undergo LLPS on their own due to their low valency. All other clients only exhibit scaffold–client interactions and cannot phase separate on their own. Key: S=scaffolds;C=Clients; HA=high-affinity scaffold–client; LA=low-affinity scaffold–client; NC=non-competing client.

System type	Composition (S in bold)	S-C binding	T^C	ϕ^c
Mixture 1	67% **M3** + 33% M2	HA	0.087	0.053
Mixture 2	33% **M3** + 67% M2	HA	0.074	0.041
Mixture 3	67% **M3** + 33% M2	LA	0.088	0.064
Mixture 4	33% **M3** + 67% M2	LA	0.087	0.069
Mixture 5	67% **M3** + 33% M2	NC	0.094	0.083
Mixture 6	67% **M3** + 33% M4	HA	0.098	0.074

TABLE S3: Critical temperatures (T^C) and critical volume fractions (ϕ^c) estimated via equations (3) and (4) (described in § III) for the multicomponent mixtures of self-interacting proteins in Figure 5. Key: D=Dominant LLPS component, i.e., the one with highest single-component critical point. All proteins interact with each other with high affinity.

System type	Composition (D in bold)	T^C	ϕ^c
Multi-component mixture 1	16.6% of each: **M4**, M4s, M3, M3p, M2.25, M2	0.096	0.106
Multi-component mixture 2	67% **M4** + 6.6% of each: M4s, M3, M3p, M2.25, M2	0.120	0.133

TABLE S4: Critical temperatures (T^C) and critical volume fractions (ϕ^c) estimated via equations (3) and (4) (described in § III) for the scaffold–client system in Figure 6. Clients only exhibit scaffold–client interactions; thus, they cannot phase separate on their own. Key: S=scaffolds;C=Clients; HA=high-affinity scaffold–client.

System type	Composition (S in bold)	S-C binding	T^C	ϕ^c
Ternary mixture 1	33.3% of each: **M4**, M3, M2	HA	0.090	0.123
Reduced mixture 1	67% **M4** + 33% M3	HA	0.116	0.132
Reduced mixture 2	67% **M4** + 33% M2	HA	0.105	0.126

TABLE S5: Typical statistical uncertainty of the coexistence densities/packing fractions for the different studied systems in this work. Note that ρ refers to the sequence dependent model, while packing fraction ϕ has been the magnitude plotted in the patchy particle phase diagrams. The lower bound value of the uncertainty is ascribed to the simulations at medium temperatures while the upper bound refers to the ones at higher and lower temperatures where sampling becomes harder, at high $T's$ because of finite size effects and at low $T's$ because of slow dynamics.

System	Uncertainty in ρ or ϕ
Sequence-dependent model	5-7%
Patchy particles pure systems	7-10%
Patchy particles binary mixtures	12-15%
Patchy particles 6-ary mixtures	10-15%

TABLE S6: Average number of bonds per cubic nanometer for the three FUS LCD variants studied in this work. These values has been measured for coexistence density of the high-density liquid branch at T/T_c=0.9.

FUS LCD variant	Connectivity (bonds/nm^3)
Unmodified	0.103
7 Tyr to Ser	0.077
14 Tyr to Ser	0.049

TABLE S7: Average number of bonds per cubic molecular diameter for pure condensed-phase protein systems at $\varepsilon_{prot-prot}$=10.5 k_BT and $p^* = 0.16$.

System	bonds/σ^3
2-valency protein	0.12
2.25-valency protein	0.13
3-valency protein (poor topology)	0.37
3-valency protein (good topology)	0.40
4-valency protein (selective)	0.96
4-valency protein (promiscuous	1.04

TABLE S8: Average number of bonds per cubic molecular diameter for a pure condensed-phase protein system and two binary scaffold–client systems. This value was computed at $\varepsilon_{prot-prot}$=13.33 k_BT and $p^* = 0$. Key: HA=high-affinity scaffold–client; NC=non-competing client.

System	bonds/σ^3
3-valency protein (S)	0.48
Mixture 1: 67% S + 33% 2-valency client; HA	0.30
Mixture 5: 67% S + 33% 2-valency client; NC	0.84

REFERENCES

[1] E. Bianchi, J. Largo, P. Tartaglia, E. Zaccarelli, and F. Sciortino, "Phase diagram of patchy colloids: Towards empty liquids," *Physical Review Letters*, vol. 97, no. 16, 2006.

[2] S. F. Banani, A. M. Rice, W. B. Peeples, Y. Lin, S. Jain, R. Parker, and M. K. Rosen, "Compositional Control of Phase-Separated Cellular Bodies," *Cell*, vol. 166, pp. 651–663, 7 2016.

[3] P. Li, S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani, P. S. Russo, Q.-X. Jiang, B. T. Nixon, and M. K. Rosen, "Phase transitions in the assembly of multivalent signalling proteins," *Nature*, vol. 483, p. 336, 3 2012.

[4] T. J. Nott, E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowietz, T. D. Craggs, D. P. Bazett-Jones, T. Pawson, J. D. Forman-Kay, and A. J. Baldwin, "Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles," *Molecular Cell*, vol. 57, pp. 936–947, 3 2015.

[5] J. Wang, J.-M. M. Choi, A. S. Holehouse, H. O. Lee, X. Zhang, M. Jahnel, S. Maharana, R. Lemaitre, A. Pozniakovsky, D. Drechsel, I. Poser, R. V. Pappu, S. Alberti, and A. A. Hyman, "A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins," *Cell*, vol. 174, pp. 688–699, 7 2018.

[6] H. Liu, S. K. Kumar, and F. Sciortino, "Vapor-liquid coexistence of patchy models: Relevance to protein phase behavior," *Journal of Chemical Physics*, vol. 127, no. 8, 2007.

[7] J. R. Espinosa, A. Garaizar, C. Vega, D. Frenkel, and R. Collepardo-Guevara, "Breakdown of the law of rectilinear diameter and related surprises in the liquid-vapor coexistence in systems of patchy particles," *Journal of Chemical Physics*, vol. 150, p. 224510, 6 2019.

[8] S. Kroschwald, S. Maharana, D. Mateju, L. Malinovska, E. Nuske, I. Poser, D. Richter, and S. Alberti, "Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules," *Elife*, vol. 4, 2015.

[9] J. Jover, A. J. Haslam, A. Galindo, G. Jackson, and E. A. Muller, "Pseudo hard-sphere potential for use in continuous molecular-dynamics simulation of spherical and chain molecules," *The Journal of Chemical Physics*, vol. 137, no. 14, p. 144505, 2012.

[10] J. R. Espinosa, C. Vega, and E. Sanz, "The mold integration method for the calculation of the crystal-fluid interfacial free energy from simulations," *J Chem. Phys.*, vol. 141, no. 13, p. 134709, 2014.

[11] S. Plimpton, "No Title," *J. Comput. Phys.*, vol. 117, p. 1, 1995.

[12] J. R. Espinosa, E. Sanz, C. Valeriani, and C. Vega, "On fluid-solid direct coexistence simulations: The pseudo-hard sphere model," *Journal of Chemical Physics*, vol. 139, no. 14, 2013.

[13] A. Ladd and L. Woodcock, "Triple-point coexistence properties of the lennard-jones system," *Chemical Physics Letters*, vol. 51, pp. 155–159, 10 1977.

[14] R. García Fernández, J. L. Abascal, and C. Vega, "The melting point of ice Ih for common water models calculated from direct coexistence of the solid-liquid interface," *Journal of Chemical Physics*, vol. 124, p. 144506, 4 2006.

[15] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," *Physical Review A*, vol. 31, p. 1695, 1985.

[16] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *The Journal of Chemical Physics*, vol. 81, no. 1, pp. 511–519, 1984.

[17] W. G. Hoover, "Constant-pressure equations of motion," *Phys. Rev. A*, vol. 34, pp. 2499–2500, 9 1986.

[18] H. Kamberaj, R. J. Low, and M. P. Neal, "Time reversible and symplectic integrators for molecular dynamics simulations of rigid molecules," *The Journal of Chemical Physics*, vol. 122, no. 22, p. 224114, 2005.

[19] G. L. Dignon, W. W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, "Sequence determinants of protein phase behavior from a coarse-grained model," *Plos Computational Biology*, vol. 14, p. e1005941, 1 2018.

[20] L. H. Kapcha and P. J. Rossky, "A simple atomic-level hydrophobicity scale reveals protein interfacial structure," *Journal of Molecular Biology*, vol. 426, pp. 484–498, 1 2014.