Controlling cluster size in 2D phase-separating binary mixtures with specific interactions

Ivan Palaia^{1,2} and Anđela Šarić^{1,2}

¹*Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria*

²*Department of Physics and Astronomy, Institute for the Physics of Living Systems, University College London, London WC1E 6BT, United Kingdom*

(*Electronic mail: andela.saric@ist.ac.at)

(Dated: 9 February 2022)

ABSTRACT: By varying the concentration of molecules in the cytoplasm or on the membrane, cells can induce the formation of condensates and liquid droplets through phase separation. Their thermodynamics, much studied, depends on the mutual interactions between microscopic constituents. Here, we focus on the kinetics and size control of 2D clusters, forming on membranes. Using molecular dynamics of patchy colloids, we model a system of two species of proteins, giving origin to specific heterotypic bonds. We find that concentrations, together with valence and bond strength, control both the size and the growth time rate of the clusters. In particular, if one species is in large excess, it gradually saturates the binding sites of the other species: the system becomes then kinetically arrested and cluster coarsening slows down or stops, thus yielding effective size selection. This phenomenology is observed both in solid and fluid clusters, which feature additional generic homotypic interactions and are reminiscent of the ones observed on biological membranes.

I. INTRODUCTION

Phase separation is a widely present phenomenon in cells^{1,2}. When concentrations of certain biochemical species exceed a threshold, condensates start to form, which bring together one or more macromolecules present in the cytoplasm. The stability of these condensates relies on the mutual interactions between the species involved: interactions can be either generic, as for proteins' intrinsically disordered tails coming together because of hydrophobicity, or specific, as for amino acid groups that form physico-chemical bonds only with given other groups³⁻⁵. In addition, attractive interaction or binding can occur between two specimens of the same species (homotypic interaction) or between different species (heterotypic interaction). Much progress has been made in understanding the phase diagrams of many of these systems, which determine whether the mixed or the demixed phases are stable at equilibrium, both from the experimental and the theoretical point of view^{6,7}.

Condensates, or clusters, of biomolecules not only form in the cytoplasm, but also on two-dimensional membrane surfaces⁸⁻¹². This occurs when at least one of the species involved is embedded in the membrane: for instance, the clustering of transmembrane protein LAT activates signalling to produce cytokines when T-cell receptors are exposed to an antigen⁹, while protein Whi3 in fungi condenses with RNA on the endoplasmic reticulum¹⁰.

Increasing density is often assumed to promote the creation of a dense phase, thus favoring phase separation. However, re-entrant behaviour of the dense phase with respect to concentration has been observed in sev-

eral phase-separating systems, suggesting that increasing concentration of a given species does not necessarily promote the formation of large assemblies. While in some cases there appears to be an electrostatic-related change in the interactions¹³⁻¹⁵, in others the effect seems to be fully stoichiometric¹⁶⁻¹⁹ and most likely functional²⁰.

Computational work on biological condensates has built upon various models, ranging from simple lattice models^{21,22}, to off-lattice coarse-grained polymers^{23,24}, stickers and spacers^{25,26}, or patchy colloids^{17-19,22,27}. In many studies, though, self-assembly is represented within the scaffold-client context, which assumes the existence of a set of proteins (scaffold) that can alone phase-separate through homotypic interactions^{2,19,27,28}. While many systems correspond to this descriptions, others, especially in 2D, do not and feature mainly specific heterotypic interactions^{3,5,17}.

While a large effort has been directed toward the understanding of the thermodynamics of phase-separating systems, the kinetics of aggregation, namely the mechanisms of growth and coarsening of clusters, has been studied less. Predictions often rely on classical kinetic theory²⁹⁻³⁴, oblivious of biological systems' specificity, and whose validity needs to be tested on a case-by-case basis. In a dynamic environment such as that of a cell, where thermodynamic equilibrium is rarely reached, a better understanding of aggregation kinetics could provide new tools to extract information on the microscopic components from macroscopic experimental observations^{10,35-37}.

In this paper, we study the kinetics of aggregation of a binary mixture with heterotypic interactions. The two species, that we name A and B, represent two kinds of proteins that can freely diffuse in a 2D environment, such

FIG. 1. Sketch of particles of type A (violet) and type B (blue) for valence $q = 3, 4$, and 5 . The big violet/blue disks, of radius σ , represent volume exclusion. Patches on A particles interact only with patches on B particles: when two interacting patches, of radius 0.1σ , superimpose, the energy gain is maximum and equals $-\varepsilon$. The dashed lines delimit the interaction area, of radius 0.3σ : two patches attract each other only if their interaction areas intersect. See Methods for details.

as a biological membrane. We describe such proteins as patchy colloids, endowed with binding sites that can form bonds with binding sites of the opposite species (Fig. 1): an A and a B particle can form a bond, but two A or two B particles cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding site already participating in a bond cannot engage in a second bond. The number of binding sites of either species is the valence, that we denote by q_A and q_B . For simplicity we restrict ourselves to the equal-valences case $q_A = q_B = q$. We study a range of integer valences q (between 3 and 5), a range of concentration ratios between B and A particles, and a range of different bond energies. For each system, at given values of valences, concentrations and bond energy, we perform Molecular Dynamics (MD) simulations and observe the self-assembly of clusters of A and B particles, of various composition and size. We find that concentration, valence and bond strength influence the rate of coarsening of the cluster, and in particular can drive the system to a kinetically arrested state, effectively resulting in cluster size selection.

II. METHODS

In our model of patchy particles, volume exclusion between particles is enforced through the Weeks-Chandler-Anderson potential

$$U_0(R) = \varepsilon_0 \left[\left(\frac{\sigma}{R} \right)^{12} - 2 \left(\frac{\sigma}{R} \right)^6 + 1 \right], \quad (1)$$

where R is the distance between the centres of the two particles, σ is their diameter, and ε_0 is chosen equal to

$10kT$. Patches are represented by ghost atoms that move rigidly with the particle; they are positioned at a fixed angular distance of $360^\circ/q$ from each other (as in Fig. 1) and at a radial distance 0.475σ from the centre of the particle. Patches on an A molecule interact only with patches on a B molecule, with the following attractive potential:

$$U_{AB\text{patch}}(r) = \begin{cases} -\varepsilon & \text{if } r < r_{\text{in}} \\ -\varepsilon \cos^2 \left(\frac{\pi}{2} \frac{r - r_{\text{in}}}{r_{\text{out}} - r_{\text{in}}} \right) & \text{if } r_{\text{in}} \leq r < r_{\text{out}} \\ 0 & \text{if } r \geq r_{\text{out}} \end{cases}, \quad (2)$$

where r is the distance between the centres of the two patches, ε is the bond strength and is a parameter that we vary, $r_{\text{in}} = 0.05\sigma$ is the range within which interaction energy is strongest, and $r_{\text{out}} = 0.15\sigma$ is the range of attraction. We consider that two particles of type A and B form a bond when they have two patches at a distance $r < r_{\text{out}}$. The chosen geometry and parameters finely prevent an A molecule from forming more than one bond with the same B molecule, and viceversa.

Simulations are run in LAMMPS^{38,39} and visualised with OVITO⁴⁰. We simulate a number of A particles n_A equal to 200 and vary the number of B particles n_B to produce the desired concentration ratio. At the beginning of each simulation, particles are positioned at random sites of a square lattice, defined such that the density of A particles is always equal to $0.03\sigma^{-2}$. This value of density is physiologically plausible for transmembrane proteins, which often bind to cytoplasmic proteins of variable density, localised close to the membrane¹⁷. The timestep chosen for the integration of the 2D equations of motion is $\tau_s = 0.01\tau_0$, where $\tau_0 = \sqrt{m\sigma^2/(kT)}$ is the simulation unit of time and m is the mass of an atom (a whole patchy particle has mass $(q+1)m$). Periodic boundary conditions are applied along both Cartesian directions. Every simulations is run for 5×10^7 time steps τ_s , a time large enough to allow cluster growth and nucleation, when present. To mimick an implicit solvent, the system is subjected to a Langevin thermostat of friction coefficient $\gamma = 1m/\tau_0$. Finally, every point in the plots shown is an average over at least 20 realisations of the simulation, run with different random number generator seeds.

We analyse selected time steps with a clustering method based on the computation of an adjacency matrix. The average cluster size (Figs. 3, 6 and 8) is defined as the number of A molecules present per cluster, averaged over clusters containing at least 2 A and 2 B molecules, from several realisations of the simulation. The fraction of single-bound clustered particles (Fig. 4) is defined as the proportion of single-bound particles versus total number of particles in a given cluster, averaged over all clusters bigger than 30 particles, from different realisations. For all these variables, error bars represent standard deviation of the averages.

We quantify compactness of clusters (plots not shown), by computing the moment of inertia of each cluster relative to an in-plane rotation about its center of mass; we then normalise such value by the analogous moment of inertia for an elliptic cluster of similar aspect ratio and of uniform density, equal to that of fully packed hard circles. The inverse of the obtained quantity defines our compactness. This procedure was already used and thoroughly described in Ref. 17.

Clusters are made fluid (this refers only to Fig. 8) by adding to the repulsive potential (1) an isotropic attraction between A-A, A-B, and B-B particles, with the following shape:

$$U_a(R) = \begin{cases} -\varepsilon_a & \text{if } R < \sigma \\ -\varepsilon_a \cos^2\left(\frac{\pi}{2} \frac{R - \sigma}{R_a - \sigma}\right) & \text{if } \sigma \leq R < R_a \\ 0 & \text{if } R \geq R_a \end{cases}, \quad (3)$$

where R is again the distance between the centre of the two particles, ε_a is the isotropic attraction strength, and R_a is the isotropic attraction range.

III. RESULTS

A. Valence, bond energy and molar ratio control cluster size

During the simulations, we observe the formation of clusters of A and B particles, some of which are represented at the bottom of Fig. 3. Clusters look roughly crystalline for $q = 3$ (triangular lattice) and $q = 4$ (square lattice), while they have an amorphous structure for $q = 5$, in which case volume exclusion forbids that all 5 bonds be satisfied. This gives us the opportunity to study the effect of valence both in crystalline condensates, which are of simpler understanding and are akin to bidimensional pathological aggregates⁴¹, and in amorphous condensates, more likely relevant for functional phase-separation phenomena occurring on the membrane.

We are interested in the effect of stoichiometry on phase separation, therefore we fix the number of A particles n_A in our simulation box (whose surface is constant) and vary the number of B particles n_B from $0.25n_A$ to $4n_A$.

The average connectivity of an A particle, defined as the number of bonds formed on average by an A particle and ranging from 0 to q , is represented in Fig. 2 for the last time frame of our simulations. For all valences, connectivity increases as we increase the relative concentration of B particles n_B/n_A . Indeed, A molecules cannot condensate alone and need B molecules to act as cross-linkers between them.

To better quantify the extent of clustering, which would be completely but unpractically described by the full cluster size distribution, we compute the mean of

such distribution, for the last time frame of our simulations (see Methods). The effect of the concentration of B linkers on the average cluster size is represented in Fig. 3, for different values of valence q and bond energy ε . Firstly, the plots show that clusters do not form within simulation time (or do not grow) if the binding energy is below a certain threshold, depending on q . Such threshold is approximately 8 and $6kT$ for $q = 3$ and 4, respectively. This is because in the dense phase bonds have to provide a certain energy to win the entropy of the dilute phase: in the bulk, q bonds contribute to such energy, so the necessary energy per bond is inversely proportional to q . For $q = 5$, the threshold bond strength increases back to $8kT$, because the amorphous dense phase is frustrated and exhibits on average less than 4 bonds per molecule. As in 3D, phase separation is therefore naturally enhanced by valence (provided that full valence is physically achievable) or at least by the effective valence exhibited in the dense phase²⁷.

Most interestingly, Fig. 3 shows that clustering is non-monotonic in the concentration of cross-linkers B. This is true at any q and at any ε above clustering threshold. Indeed, if there are too few cross-linkers, it is impossible to link all A molecules, some of which will remain free. Increasing the molar ratio n_B/n_A , up to 1, thus promotes both A connectivity (Fig. 2) and cluster size (Fig. 3). However, if cross-linkers are overabundant, they compete for binding sites on A molecules and end up filling all of them. It becomes then impossible for an already bound B molecule to find an empty binding site from another A molecule, so that A molecules become coated by B molecules and cross-links do not form (see cluster snapshots in Fig. 3). Average cluster size decreases, while A connectivity tends to saturate to q . A signature of this phenomenon is the average fraction of capping particles within a cluster, i.e. the proportion of clustered particles that are involved only in a single bond: this number is minimum at $n_B/n_A = 1$, and increases as the molar ratio deviates from 1 (Fig. 4).

In Figs. 3 and 4, the non-monotonic behaviour is more pronounced at strong than at weak binding ε . This is because at strong binding, bonds are irreversible and as soon as a full B cap forms around two given A molecules, no link will ever form between them. Instead, when binding is weak, caps have a finite lifetime and B molecules can detach, thus allowing, to some extent, for bond reorganisation and formation of new links; this favours the formation of larger clusters and of a proper dense phase. In other words, the cluster size is non-monotonic not only in the molar ratio n_B/n_A , but also in the bond energy ε (at least far from equal molar ratio): if ε is too low, clusters do not form because the energy gain is not sufficient to overcome the higher entropy of the dilute phase, but if ε is too large, the system gets stuck in a state where capping prevents coarsening.

Particular attention needs to be devoted to the

FIG. 2. Average connectivity of A molecules, at different binding energy values ϵ (in units of kT), for $q = 3, 4$ and 5 , at the end of the simulation.

FIG. 3. Average cluster size, at different binding energy values ϵ (in units of kT), for $q = 3, 4$ and 5 , at the end of the simulation. Only clusters bigger than 4 molecules are counted (see Methods). Below, for each valence, snapshots of typical clusters are shown for concentration ratios of 1/4, 1 and 4 (pointed at by vertical lines), and for chosen ϵ (indicated by the symbols).

equal concentration case, where we observe a non-monotonicity in cluster size with ϵ as well. Such phenomenon is analogous to magic-number effects reported in previous studies of many 3D systems or models^{21–24}, where fine tuning of valence and/or concentration can cause major variations in the condensate properties. This is seen in Fig. 2, particularly in the $q = 4$ and 5 cases, where connectivity is enhanced at $n_B / n_A = 1$ for intermediate bond strengths ϵ , but hampered for low or strong ϵ . This is again because lower ϵ allows for more bond reorganisation, indeed clusters at intermediate ϵ appear more compact and with fewer defects (see clusters snapshots in Fig. 2 at different bond strengths, and the Methods for a quantification of compactness). This is particularly important at equal molar ratio, when, due to the lim-

ited amount of B, full A connectivity can only be reached if no defects are present in the dense phase and the cluster is as compact as possible. At $n_B / n_A \gg 1$, this phenomenon is not relevant anymore, because the linker B is overabundant and A sites can be saturated anyway, by forming small B-coated clusters as described above; in such case, lower ϵ just means lower bond life time, and thus lower A connectivity (as shown by the right-hand-side parts of the panels in Fig. 2).

B. Arrested state or thermodynamic equilibrium?

One can wonder whether the fully saturated state we observe at large concentration ratios is due to kinetic ar-

FIG. 4. Fraction of clustered particles involved in one bond only, at the end of the simulation, for $q = 4$, at different binding energy values ϵ (in units of kT). Valences $q = 3$ and 5 exhibit similar behaviour.

rest or is a thermodynamic minimum. To gain physical intuition, we compare the free energy of two limit states, sketched in Fig. 5. For simplicity we discretise space and imagine V sites, filled by a number n_A of A particles and a number $n_B > n_A$ of B particles; the remaining sites are left empty.

A limit state (state I) is the one featuring one maximally connected, giant, round cluster with n_A A particles and $\sim n_A$ B particles, while $\sim n_B - n_A$ B particles are free (i.e. they form the dilute phase). Its free energy, for low densities and strong bonds, can be approximated as

$$F_I = -q\epsilon n_A - kT \left[\ln V + (n_B - n_A) - (n_B - n_A) \ln \left(\frac{n_B - n_A}{V} \right) \right]. \quad (4)$$

The first term in square brackets refers to the translational entropy of the cluster, while the last two terms represent the translational entropy of the free B particles.

The second limit state (state II) is the one where all A particles are fully capped by otherwise unbound B particles, so that the system has n_A tiny clusters, each made of 1 A and q B particles, and $n_B - qn_A$ free B particles. Clearly, this scenario is only possible if $n_B > qn_A$. The related free energy is approximately

$$F_{II} = -q\epsilon n_A - kT \left[n_A - n_A \ln \left(\frac{n_A}{V} \right) + (n_B - qn_A) - (n_B - qn_A) \ln \left(\frac{n_B - qn_A}{V} \right) \right]. \quad (5)$$

The first two terms in square brackets represent the entropy of the tiny clusters, while the last two the entropy of the free B particles.

Between states I and II, the one with the lower free energy is the thermodynamically stable one. Note that both states have the same energy (first term in both equations)

FIG. 5. Entropy S of the lattice model, at $q = 4$, for the two limit cases: I, where all A particles form a single cluster, and II, where each A particle forms an independent cluster with 4 B particles.

as all A particles have their binding sites filled, so that entropy tips the scale. Entropy for states I and II is represented in Fig. 5. The single cluster (solid line) is still substantially favoured at $n_B / n_A = q$. At larger n_B , the entropic advantage decreases and then reverses. However the free energies of the two states remain comparable on the scale of a kT per molecule.

In our definition of state I, we considered the giant cluster to have a defined (circular) shape. While making the calculation easier, this eliminates from the count many configurations of equal energy $-q\epsilon n_A$ and featuring one single cluster. If we redefine state I to account for all possible shapes, we can instead interpret F_I as a constrained free energy limited to minimal-energy configurations with one single cluster (just like F_{II} is a constrained free energy limited to configurations with n_A clusters). This leads to a new expression for F_I , worked out in the Appendix. In short, non-circular shapes require more B particles to coat the perimeter of the cluster: while enlarging the configuration space for the giant cluster, this reduces the number of free B particles and therefore the associated entropy. Using results from the statistics of lattice animals^{42–45}, we can group configurations by cluster perimeter and find that the cluster of minimal perimeter is the most favourable one, so that Eq. (4) is still approximately valid for our redefined state I.

Altogether, this suggests that the fully-capped state (II), while prevalent in our simulations, is not necessarily thermodynamically stable and that some kinetic lock is essential to make such a state the observed one at large concentration ratios.

FIG. 6. Time evolution of the average cluster size, as plotted in Fig. 3 and as defined in the text and in Methods, for $q = 4$. From left to right, the molar ratio n_B/n_A is $1/4$, 1 and 4 . Bond energies ϵ are in units of kT . Thick lines show the intervals of linear behaviour from which we extract growth exponents for Fig. 7. Energy not shown do not give clusters.

C. Kinetics of clustering

Prompted by the above considerations, we analyze the kinetics of clustering, which are likely relevant precisely for the size-control mechanism described in section III A.

The time evolution of the average cluster size is represented for $q = 4$ in Fig. 6, for three different molar ratios. At equal concentration (middle panel), the number of particles in the clusters (i.e. their surface) steadily grows with time for all interaction energies ϵ , provided that enough time is allowed for nucleation at intermediate ϵ . A nucleation process is particularly evident for $\epsilon = 7kT$, where the curve stays flat for the first 10^6 time steps. Nucleation indicates that the system is initialised in a metastable state, namely at a point of the phase diagram between the spinodal and the binodal; absence of nucleation indicates instead that the system is unstable and the relaxation process is a standard spinodal decomposition^{30,46}. In the latter case ($\epsilon > 7kT$) the average number of particles in a cluster scales approximately as $\sim t^{0.6}$, corresponding to a cluster diameter growing with exponent $\alpha \simeq 0.6/2 = 0.3$. Provided that n_B/n_A stays close to unity, this is true also for $q = 3$ and 5 , as shown by Fig. 7, where values of the cluster diameter growth exponent α obtained from fitting cluster size curves at late times are reported. A power law behaviour is a sign of self-similarity, which typically characterises relaxation to equilibrium of quenched interacting system (here quenching means initialising the system in the vapour state, but running the simulation at temperature lower than critical, where the equilibrium state exhibits phase separation).

For low or high concentration ratios, the aggregation kinetics is slower. At small ϵ , this might be due to larger nucleation times, but linear behaviour in Fig. 6 rather points at a slow down in the power-law growth rate, at least for the probed times. This is consistent with the fact that at large ϵ , coarsening halts completely after a given time. This observation reflects the more pronounced non-monotonicity observed at final time for large bond

strength in Fig. 3. Fig. 7 captures the decay of exponent α for different valences, as the concentration ratio deviates from unity.

We highlight that the process we see is mostly due to merging of existing clusters and not due to clusters growing by monomer addition²⁹ (i.e. condensation of free particles): the latter process occurs very rapidly, usually in the very first steps of the simulation. This is confirmed by the observation that the fraction of clustered A reaches steady state on a time scale smaller than those represented in Fig. 6 ($< 10^5 \tau_s$).

A value of the exponent $\alpha \simeq 0.3$ calls for a comparison with values reported in the literature. For 2D solid binary mixtures, an exponent $\alpha = 1/5$ was proposed when growth is dominated by cluster diffusion and coalescence, at temperatures well below the critical one^{31,32}. The so-called Lifshitz-Slyozov mechanism, akin to Ostwald ripening and due to evaporation and recondensation of single particles, should instead give $\alpha = 1/3$, irrespective of dimensionality^{30,33,46}. A richer phenomenology has been reported for fluid systems, for which (still in 2D) a crossover from $\alpha = 1/3$, to $1/2$, to $2/3$ is expected^{30,34,47,48}, although values of α between $1/4$ and $1/3$ have also been observed when hydrodynamics is suppressed^{49,50}. Our system is neither fully solid, because clusters can diffuse and coalesce, nor fully fluid, as hydrodynamic drag effects are negligible at our low densities. In addition the morphology of the clusters and the stoichiometry can affect growth^{30,47}.

We provide a simple theoretical argument, analogous to the one proposed in Ref. 51, showing why α in our case should be close to $1/4$. We assume that the system exhibits dynamic scaling (i.e. self-similarity), that the probability of coalescence upon collision does not depend on time (which is plausible when growth is not arrested), that all clusters have equal size M at a given time, for simplicity, and that the cluster density ρ_M (number of clusters per unit surface at a given time) is small compared to the density of molecules within clusters. Now, since the total number of molecules is conserved,

FIG. 7. Growth exponent α measured from linear fit of cluster growth curves in time (Fig. 6). At concentration ratio close to 1, $\alpha \simeq 0.3$ for most curves. Away from this regime, growth is hindered: coarsening is slowed down for $\epsilon = 10kT$ (dashed lines) and completely halted for $\epsilon = 20kT$ (continuous lines).

$\rho_M = \rho_1/M$ where ρ_1 is the total number of molecules per unit surface. The average distance between two clusters at a given time is then $r_M \simeq \rho_M^{-1/2} = (M/\rho_1)^{1/2}$. Given that the diffusion coefficient scales linearly with the mass, i.e. $D_M = D_1/M$, the typical time t_M it takes for two clusters of size M to collide under the only effect of diffusion can be estimated as $t_M \simeq r_M^2/D_M \sim M^2$. Neglecting fractal effects, the mass of the clusters scales as the square of its linear dimension ℓ_M (its diameter, for instance): $M \sim \ell_M^2$, hence $t_M \sim \ell_M^4$. Since the system is self-similar, this relation must hold for any length in the system and at any time, giving $\alpha = 1/4$.

D. Size control is retained in fluid clusters

The clusters of the simple model we analysed so far are essentially solid-like: especially for $q = 3$ and 4, they are crystalline and their internal conformation can only change by evaporation of material and condensation of new material at a different place. This means defects can be eliminated only if they occur at the cluster boundaries. We show that the size-controlling mechanism we propose can be physiologically relevant, by making our clusters fluid. Biological condensates are indeed liquid-like even in 2D – although the availability of cross-linking molecules has been shown to reduce their internal diffusivity¹⁷.

To obtain fluidity, we add a generic isotropic attraction between any two particles (homo- and heterotypic), parametrised by its strength ϵ_a and range R_a (see Methods, Eq. (3)). Such attraction allows for local rearrangement once A-B bonds break: in this case particles can move around without having to leave the cluster, as some cohesion energy is still provided by isotropic interactions. The diffusivity of particles within clusters (internal

FIG. 8. Top panel: the estimated internal diffusivity D_{int} , for $q = 5$ and $R_a = 2.0\sigma$, measured at $n_B/n_A = 1$. Black crosses indicate that clusters do not form or are too volatile. D_{int} is computed by fitting the mean square displacement of particle pairs belonging to the same cluster in the time range $4 - 4.1 \times 10^7 \tau_s$; pairs with one or both particles evaporating from the cluster within this time are discarded from the calculation. Bottom panel: an example of non-monotonic cluster size, measured at the end of the simulation, for a set of parameters where clusters are very fluid.

diffusivity D_{int}) is a measure of cluster fluidity. This is estimated, at late times, in the top plot of Fig. 8, as a function of bond strength ϵ and isotropic attraction ϵ_a , for a chosen range R_a . While in the large ϵ regime particles are locked by unbreakable bonds, there is a window where diffusivity is large. This occurs at small enough ϵ , so that bonds can be broken, yet large enough ϵ_a so as to guarantee cluster formation and stability.

If the fraction of cluster energy provided by bonds stays comparable to the one provided by isotropic attraction, meaning that clusters would fall apart without specific interactions, unbound particles are more likely to leave the cluster than bound ones. We then expect to retain a signature of the size selection mechanism we discussed above. This is shown in the bottom plot of Fig. 8, featuring non-monotonic behaviour of the average cluster size with concentration ratio, in a case with fluid clusters. When the proportion of energy coming from isotropic attraction is too high (large ϵ_a), this behaviour is lost.

IV. DISCUSSION

Using a model based on patchy colloids with simple geometries, we studied a 2D binary mixture where specific heterotypic interactions are dominant. When the concentrations of the two species are comparable, clusters form and can in principle grow indefinitely. On the contrary, when one species is overabundant cluster growth is hindered, resulting in smaller clusters (Fig. 3). This happens because particles from the minority species tend to have all their binding sites saturated by particles of the majority species. Such capping phenomenon results in cluster surfaces being coated by same-kind particles, which cannot bind, thus preventing cluster coarsening. This phenomenon has been experimentally observed in biological condensates^{16,17} that do not belong to the client-scaffold category.

We show that such size-control process, although it can lead to stationary states within observation time (Fig. 6), is not of obvious thermodynamic origin, as instead recently found for analogous processes in different systems^{19,52}. On the contrary, it can be driven by a kinetic trap whose depth increases with bond strength and unevenness in concentrations. This is reflected by a reduction in growth exponents, from $\simeq 0.3$ to 0, upon increase of the latter two quantities (Fig. 7). While large concentration unevenness decreases the probability for two clusters to coalesce, through the coating mechanism discussed above, strong bonds prevent bond rearrangement, necessary to reach equilibrium and eliminate defects. However, we still observe significant concentration-dependent size control for any value of bond strength.

Yet another way to control cluster size is by modulating valence. This strategy allows cells to tune properties of physiological condensates^{2,53}, particularly on membranes⁵⁴, but is also the functioning mechanism of artificial optogenetic tools used to study phase-separation^{7,55}. We see that increasing valence broadens the range of concentration ratios at which clusters form. Full capping can indeed only occur when $n_B/n_A \geq q$ (or $\leq 1/q$), as confirmed by our analysis of cluster size and growth exponent (Figs. 3 and 7).

Similarly to cluster size, concentration ratios and bond strength also affect the internal connectivity of the cluster: the connectivity of one species is obviously promoted by increasing the concentration of the cross-linking species, until it saturates (Fig. 2). The connectivity of the majority species decays with its concentration though. Optimal cluster connectivity is reached at equal concentration ratios. No matter the obvious limitations of the model, where spurious geometric effects are present, these considerations are mirrored by diffusivity measurements in presence of heterotypic interactions only¹⁷.

In conclusion, we highlight how clustering curves as

a function of concentration ratio might be a way to infer properties on the molecular scale. Our results suggest that if clustering is non-monotonic, specific heterotypic interactions are possibly responsible. The more pronounced the non-monotonicity, the stronger are specific bonds with respect to generic (isotropic) interactions. The latter, however, might be present if the cluster appears macroscopically fluid. Finally, the kinetics of cluster growth carries likewise relevant information: a decrease in growth exponent with concentration ratio might as well be a signature of dominantly heterotypic interactions.

ACKNOWLEDGMENTS

We thank Longhui Zeng and Xiaolei Su (Yale University) for bringing the topic to our attention and for useful comments. This work has received funding by the European Research Council under the European Union's Horizon 2020 research and innovation programme (ERC grant No. 802960 and Marie Skłodowska-Curie grant No. 101034413). We are grateful to the UK Materials and Molecular Modelling Hub for computational resources, which is partially funded by EPSRC (Grant No. EP/P020194/1 and EP/T022213/1). We acknowledge support from IST Austria and from the Royal Society (Grant No. UF160266).

APPENDIX: CALCULATION OF CONSTRAINED FREE ENERGY

The partition function of a macroscopic state where the maximum number of bonds are formed (qn_A) and there is only one giant cluster of whatever shape is the following:

$$Z = \frac{1}{n_A! n_B!} e^{\beta q n_A} \mathcal{C}, \quad (6)$$

where $\beta = (kT)^{-1}$ and $\mathcal{C}$ is the number of different microscopic configurations corresponding to the same macroscopic state we are constraining on.

We group configurations based on the number n_{Bc} of B particles belonging to the cluster; the free B particles are then $n_{Bf} = n_B - n_{Bc}$. We have

$$\mathcal{C} = \sum_{n_{Bc}} V n_A! n_{Bc}! \binom{V - n_A - n_{Bc}}{n_{Bf}} n_{Bf}! \binom{n_B}{n_{Bf}} \mathcal{N}_{n_{Bc}}. \quad (7)$$

The seven factors of this sum represent, respectively: positions of the cluster in a lattice with V sites, permutations of the n_A A particles in the cluster, permutations of the n_{Bc} B particles in the cluster, positions of the n_{Bf} free B particles (to be chosen among the sites not occupied by the cluster), permutations of the n_{Bf} free B particles,

FIG. 9. Distribution of semiperimeters, as a function of rescaled semiperimeter p . The rescaling factor $\bar{p}$ is defined in the text. The curves for sizes s from 20 to 48 (obtained from Ref. 56) approximately collapse on a single curve that resembles a Gaussian. Note that the vertical scale is logarithmic.

ways to chose which B particles are free and which belong to the cluster, and finally ($\mathcal{N}_{n_{Bc}}$) number of different cluster shapes (including rotations) that embed n_{Bc} particles in the cluster.

The B particles in the cluster can have either more than one bond (we say that they belong to the bulk) or only one bond (they form the boundary). The bulk contains always one B particle per A particle (see top-right sketch in Fig. 5), so there are n_A B particles in it. The number of boundary particles, instead, is proportional to the perimeter of the cluster bulk. In summary, the number of B particles in the cluster is $n_{Bc} = n_A + p$, where p is the semiperimeter of the shape formed by the $2n_A$ particles in the bulk. Eq. (7) is then interpretable as a sum over all possible perimeters, and $\mathcal{N}_{n_{Bc}}$ as the number of different shapes of semiperimeter $p = n_{Bc} - n_A$ that can be built clustering $2n_A$ particles.

The statistics of connected shapes on lattices, so-called lattice animals or polyominoes, has been a long studied subject⁴⁵. $\mathcal{N}_{n_{Bc}}$ can be written as the product of the total number a_s of shapes of size $s = 2n_A$ (scaling as $4.0^s/s$)^{44,45} and the perimeter distribution f_p (i.e. the fraction of shapes of size s that have semiperimeter p). To our knowledge a rigorous estimate of the asymptotic behaviour of the latter quantity does not exist, so we extracted f_p for s up to 48 from exact tabulated enumerations of lattice animals^{42,43,56}. Fig. 9 shows that semiperimeters, rescaled by $\bar{p} \simeq 0.95 s^{0.90}$, seem to follow approximately a normal distribution with mean $\mu = 1$ and variance $\sigma = 0.06$.

In summary, Eq. (7) can be rewritten as

$$\mathcal{C} = n_A! n_B! V a_{2n_A} \sum_p \binom{V - 2n_A - p}{n_B - n_A - p} f_p. \quad (8)$$

In the sum, notice the competition between the sec-

ond term, representing “shape” entropy (more shapes are available at larger perimeter), and the first term, representing translational entropy of free B particles (larger cluster perimeter means fewer B particles in solution). Using Stirling’s approximation and passing to the continuous limit for f_p , we obtain

$$\begin{aligned} \mathcal{C} \sim n_A! n_B! V \frac{a_{2n_A}}{\bar{p}} e^{-(V - n_A - n_B) \ln(V - n_A - n_B)} \\ \times \sum_p \exp \left[(V - 2n_A - p) \ln(V - 2n_A - p) \right. \\ \left. - (n_B - n_A - p) \ln(n_B - n_A - p) - \frac{(p/\bar{p} - \mu)^2}{2\sigma^2} \right]. \quad (9) \end{aligned}$$

We verified numerically that the Gaussian term only plays a role at very high densities, not relevant in our work. The dominant term of the sum is then the smallest perimeter one. To leading order, this amounts to setting $a_{2n_A} f_p = \delta_{p,0}$ (Kronecker delta) in Eq. (8). By doing so, one recovers Eq. (4), which refers to a circular cluster.

- ¹Y. Shin and C. P. Brangwynne, “Liquid phase condensation in cell physiology and disease,” *Science* **357**, eaaf4382 (2017).
- ²S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, “Biomolecular condensates: Organizers of cellular biochemistry,” *Nature Reviews Molecular Cell Biology* **18**, 285–298 (2017).
- ³L. B. Case, J. A. Ditlev, and M. K. Rosen, “Regulation of Transmembrane Signaling by Phase Separation,” *Annual Review of Biophysics* **48**, 465–494 (2019).
- ⁴G. L. Dignon, R. B. Best, and J. Mittal, “Biomolecular phase separation: From molecular driving forces to macroscopic properties,” *Annual Review of Physical Chemistry* **71**, 53–75 (2020).
- ⁵A. Musacchio, “On the role of phase separation in the biogenesis of membraneless compartments,” *The EMBO Journal*, e109952 (2022).
- ⁶M. F. Hagan and G. M. Grason, “Equilibrium mechanisms of self-limiting assembly,” *Reviews of Modern Physics* **93**, 25008 (2021), arXiv:2007.01927.
- ⁷S. Alberti, A. Gladfelter, and T. Mittag, “Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates,” *Cell* **176**, 419–434 (2019).
- ⁸Y. G. Zhao and H. Zhang, “Phase Separation in Membrane Biology: The Interplay between Membrane-Bound Organelles and Membraneless Condensates,” *Developmental Cell* **55**, 30–44 (2020).
- ⁹X. Su, J. A. Ditlev, E. Hui, W. Xing, S. Banjade, J. Okrut, D. S. King, J. Taunton, M. K. Rosen, and R. D. Vale, “Phase separation of signaling molecules promotes T cell receptor signal transduction,” *Science* **352**, 595–599 (2016).
- ¹⁰W. T. Snead, T. M. Gerbich, I. Seim, Z. Hu, and A. S. Gladfelter, “Membrane surfaces regulate assembly of a ribonucleoprotein condensate,” (2021), 10.1101/2021.04.24.441251.
- ¹¹K. J. Day, G. Kago, L. Wang, J. B. Richter, C. C. Hayden, E. M. Lafer, and J. C. Stachowiak, “Liquid-like protein interactions catalyze assembly of endocytic vesicles,” *Nature Cell Biology* **23**, 366–376 (2021).
- ¹²S. Mondal, S. Botterbusch, K. Narayan, I. Powers, J. Zheng, R. Jin, and T. Baumgart, “Multivalent Interactions between Molecular Components Involved in Clathrin Independent Endocytosis Drive Protein Phase Separation,” (2021), 10.1101/2021.12.27.474257.
- ¹³P. R. Banerjee, A. N. Milin, M. M. Moosa, P. L. Onuchic, and A. A. Deniz, “Reentrant Phase Transitions Drives Dynamic Substructure Formation in Ribonucleoprotein Droplets,” *Angewandte Chemie International Edition* **56**, 11354–11359 (2017).
- ¹⁴A. N. Milin and A. A. Deniz, “Reentrant Phase Transitions and Non-Equilibrium Dynamics in Membraneless Organelles,” *Biochemistry* **57**, 2470–2477 (2018).

- ¹⁵G. Krainer, T. J. Welsh, J. A. Joseph, J. R. Espinosa, S. Wittmann, E. de Csilléry, A. Sridhar, Z. Toprakcioglu, G. Gudiškytė, M. A. Czekalska, W. E. Arter, J. Guillén-Boixet, T. M. Franzmann, S. Qamar, P. S. George-Hyslop, A. A. Hyman, R. Collepardo-Guevara, S. Alberti, and T. P. Knowles, "Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions," *Nature Communications* **12**, 1–14 (2021).
- ¹⁶D. W. Sanders, N. Kedersha, D. S. Lee, A. R. Strom, V. Drake, J. A. Riback, D. Bracha, J. M. Eeftens, A. Iwanicki, A. Wang, M. T. Wei, G. Whitney, S. M. Lyons, P. Anderson, W. M. Jacobs, P. Ivanov, and C. P. Brangwynne, "Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization," *Cell* **181**, 306–324.e28 (2020).
- ¹⁷L. Zeng, I. Palaia, A. Šarić, and X. Su, "PLCγ1 promotes phase separation of T cell signaling components," *Journal of Cell Biology* **220**, e202009154 (2021).
- ¹⁸S. Sun, T. GrandPre, D. T. Limmer, and J. T. Groves, "Kinetic frustration by limited bond availability controls the LAT protein condensation phase transition on membranes," (2021), 10.1101/2021.12.05.471009.
- ¹⁹I. Sanchez-Burgos, J. A. Joseph, R. Collepardo-Guevara, and J. R. Espinosa, "Size conservation emerges spontaneously in biomolecular condensates formed by scaffolds and surfactant clients," *Scientific Reports* **11**, 1–10 (2021).
- ²⁰L. B. Case, X. Zhang, J. A. Ditlev, and M. K. Rosen, "Stoichiometry controls activity of phase-separated clusters of actin signaling proteins," *Science* **363**, 1093–1097 (2019).
- ²¹B. Xu, G. He, B. G. Weiner, P. Ronceray, Y. Meir, M. C. Jonikas, and N. S. Wingreen, "Rigidity enhances a magic-number effect in polymer phase separation," *Nature Communications* **11**, 4–11 (2020).
- ²²E. S. Freeman Rosenzweig, B. Xu, L. Kuhn Cuellar, A. Martínez-Sánchez, M. Schaffer, M. Strauss, H. N. Cartwright, P. Ronceray, J. M. Plitzko, F. Förster, N. S. Wingreen, B. D. Engel, L. C. Mackinder, and M. C. Jonikas, "The Eukaryotic CO₂-Concentrating Organelle Is Liquid-like and Exhibits Dynamic Reorganization," *Cell* **171**, 148–162.e19 (2017).
- ²³Y. Zhang, B. Xu, B. G. Weiner, Y. Meir, and N. S. Wingreen, "Decoding the physical principles of two-component biomolecular phase separation," *eLife* **10**, e62403 (2021).
- ²⁴P. Ronceray, Y. Zhang, X. Liu, and N. S. Wingreen, "Stoichiometry controls the dynamics of liquid condensates of associative proteins," (2021), arXiv:2101.01164.
- ²⁵J.-M. Choi, F. Dar, and R. V. Pappu, "LASSI: A lattice model for simulating phase transitions of multivalent proteins," *PLOS Computational Biology* **15**, e1007028 (2019).
- ²⁶J.-M. Choi, A. S. Holehouse, and R. V. Pappu, "Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions," *Annual Review of Biophysics* **49** (2020), 10.1146/annurev-biophys-121219-081629.
- ²⁷J. R. Espinosa, J. A. Joseph, I. Sanchez-Burgos, A. Garaizar, D. Frenkel, and R. Collepardo-Guevara, "Liquid network connectivity regulates the stability and composition of biomolecular condensates with many components," *Proceedings of the National Academy of Sciences* **117**, 13238–13247 (2020).
- ²⁸S. 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* **166**, 651–663 (2016).
- ²⁹J. Berry, C. P. Brangwynne, and M. Haataja, "Physical principles of intracellular organization via active and passive phase transitions," *Reports on Progress in Physics* **81**, 046601 (2018).
- ³⁰S. K. Das, S. Roy, and J. Midya, "Coarsening in fluid phase transitions," *Comptes Rendus Physique* **16**, 303–315 (2015).
- ³¹K. Binder and D. Stauffer, "Theory for the slowing down of the relaxation and spinodal decomposition of binary mixtures," *Physical Review Letters* **33**, 1006–1009 (1974).
- ³²K. Binder, "Theory for the dynamics of "clusters." II. Critical diffusion in binary systems and the kinetics of phase separation," *Physical Review B* **15**, 4425–4447 (1977).
- ³³I. Lifshitz and V. Slyozov, "The kinetics of precipitation from super-saturated solid solutions," *Journal of Physics and Chemistry of Solids* **19**, 35–50 (1961).
- ³⁴E. D. Siggia, "Late stages of spinodal decomposition in binary mixtures," *Physical Review A* **20**, 595–605 (1979).
- ³⁵S. Ranganathan and E. I. Shakhnovich, "Dynamic metastable long-living droplets formed by sticker-spacer proteins," *eLife* **9**, 1–25 (2020).
- ³⁶S. F. Shimobayashi, P. Ronceray, D. W. Sanders, M. P. Haataja, and C. P. Brangwynne, "Nucleation landscape of biomolecular condensates," *Nature* (2021), 10.1038/s41586-021-03905-5.
- ³⁷D. S. Lee, N. S. Wingreen, and C. P. Brangwynne, "Chromatin mechanics dictates subdiffusion and coarsening dynamics of embedded condensates," *Nature Physics* **17**, 531–538 (2021).
- ³⁸S. Plimpton, "Fast parallel algorithms for short-range molecular dynamics," *Journal of computational physics* **117**, 1–19 (1995).
- ³⁹P. in 't Veld, S. Plimpton, and G. Grest, "Accurate and efficient methods for modeling colloidal mixtures in an explicit solvent using molecular dynamics," *Comp. Phys. Comm.* **179**, 320–329 (2008).
- ⁴⁰A. Stukowski, "Visualization and analysis of atomistic simulation data with ovito—the open visualization tool," *Modelling and Simulation in Materials Science and Engineering* **18**, 015012 (2009).
- ⁴¹M. Fuxreiter and M. Vendruscolo, "Generic nature of the condensed states of proteins," *Nature Cell Biology* **23**, 587–594 (2021).
- ⁴²M. F. Sykes, D. S. Gaunt, and M. Glen, "Percolation processes in two dimensions. III. High density series expansions," *Journal of Physics A: General Physics* **9**, 715–724 (1976).
- ⁴³S. Mertens, "Lattice animals: A fast enumeration algorithm and new perimeter polynomials," *Journal of Statistical Physics* **58**, 1095–1108 (1990).
- ⁴⁴I. Jensen and A. J. Guttmann, "Statistics of lattice animals (polyominoes) and polygons," *Journal of Physics A: Mathematical and General* **33**, L257–L263 (2000).
- ⁴⁵A. J. Guttmann, "History and introduction to polygon models and polyominoes," in *Polygons, Polyominoes and Polycubes*, Lecture Notes in Physics, Vol. 775, edited by A. J. Guttmann (Springer, Dordrecht, 2009) Chap. 1, pp. 1–21.
- ⁴⁶K. Binder and P. Fratzl, "Spinodal decomposition," in *Phase Transformations in Materials*, edited by G. Kostorz (John Wiley & Sons, Ltd, 2001) Chap. 6, pp. 409–480, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/352760264X.ch6>.
- ⁴⁷A. Singh and S. Puri, "Phase separation in ternary fluid mixtures: a molecular dynamics study," *Soft Matter* **11**, 2213–2219 (2015).
- ⁴⁸G. Leptoukh, B. Strickland, and C. Roland, "Phase separation in two-dimensional fluid mixtures," *Physical Review Letters* **74**, 3636–3639 (1995).
- ⁴⁹E. Velasco and S. Toxvaerd, "Computer simulation of phase separation in a two-dimensional binary fluid mixture," *Physical Review Letters* **71**, 388–391 (1993).
- ⁵⁰J. Midya and S. K. Das, "Kinetics of domain growth and aging in a two-dimensional off-lattice system," *Physical Review E* **102**, 62119 (2020), 2009.11202.
- ⁵¹M. Kolb, "Unified description of static and dynamic scaling for kinetic cluster formation," *Physical Review Letters* **53**, 1653–1656 (1984).
- ⁵²M. Kar, F. Dar, T. J. Welsh, L. Vogel, R. Kühnemuth, A. Majumdar, G. Krainer, T. M. Franzmann, S. Alberti, C. A. M. Seidel, T. P. Knowles, A. A. Hyman, and R. V. Pappu, "Phase separating RNA binding proteins form heterogeneous distributions of clusters in sub-saturated solutions," (2022), 10.1101/2022.02.03.478969.
- ⁵³P. 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* **483**, 336–340 (2012).
- ⁵⁴S. Banjade and M. K. Rosen, "Phase transitions of multivalent proteins can promote clustering of membrane receptors," *eLife* **3**, 1–24 (2014).
- ⁵⁵D. Bracha, M. T. Walls, M.-T. Wei, L. Zhu, M. Kurian, J. L. Avalos,

J. E. Toettcher, and C. P. Brangwynne, “Mapping Local and Global Liquid Phase Behavior in Living Cells Using Photo-Oligomerizable

Seeds,” *Cell* **175**, 1467–1480.e13 (2018).

⁵⁶Data from <http://wasd.urz.uni-magdeburg.de/mertens/research/animals/>, accessed 30 September 2021.