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

Cite as: J. Chem. Phys. 156, 194902 (2022); https://doi.org/10.1063/5.0087769
Submitted: 09 February 2022 • Accepted: 27 April 2022 • Published Online: 16 May 2022
Published open access through an agreement with JISC Collections

Ivan Palaia and ANDela Šarić

COLLECTIONS

Paper published as part of the special topic on 2022 JCP Emerging Investigators Special Collection

ARTICLES YOU MAY BE INTERESTED IN

Machine learning implicit solvation for molecular dynamics
The Journal of Chemical Physics 155, 084101 (2021); https://doi.org/10.1063/5.0059915

Empty liquid state and re-entrant phase behavior of the patchy colloids confined in porous media
The Journal of Chemical Physics 156, 161102 (2022); https://doi.org/10.1063/5.0088716

Fully periodic, computationally efficient constant potential molecular dynamics simulations of ionic liquid supercapacitors
The Journal of Chemical Physics 156, 184101 (2022); https://doi.org/10.1063/5.0086986
Controlling cluster size in 2D phase-separating binary mixtures with specific interactions

Cite as: J. Chem. Phys. 156, 194902 (2022); doi: 10.1063/5.0087769
Submitted: 9 February 2022 • Accepted: 27 April 2022 • Published Online: 16 May 2022

Ivan Palaia\textsuperscript{1,2} and Anđela Šarić\textsuperscript{1,2,a)\textsuperscript{2}}

AFFILIATIONS
\textsuperscript{1} Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria
\textsuperscript{2} Department of Physics and Astronomy, Institute for the Physics of Living Systems, University College London, London WC1E 6BT, United Kingdom

Note: This paper is part of the 2022 JCP Emerging Investigators Special Collection.
a) Author to whom correspondence should be addressed: andela.saric@ist.ac.at

ABSTRACT
By varying the concentration of molecules in the cytoplasm or on the membrane, cells can induce the formation of condensates and liquid droplets, similar to 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 then becomes 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 phenomena are widely present in cells.\textsuperscript{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 groups.\textsuperscript{3–6} 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.\textsuperscript{7,8}

Condensates, or clusters, of biomolecules not only form in the cytoplasm, but also on two-dimensional membrane surfaces.\textsuperscript{9–13} Increasing density is often assumed to promote the creation of a dense phase, thus favoring phase separation. However, re-entrant behavior of the dense phase with respect to the concentration has been observed in several phase-separating systems, suggesting that increasing the 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 interactions,\textsuperscript{14–16} in others, the effect seems to be fully stoichiometric\textsuperscript{17–20} and most likely functional.\textsuperscript{21}

Computational work on biological condensates has built upon various models, ranging from simple lattice models\textsuperscript{22,23} to off-lattice coarse-grained polymers,\textsuperscript{24,25} stickers and spacers,\textsuperscript{26,27} or patchy colloids.\textsuperscript{18–20,23,25,27} 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 in the membrane; for instance, the clustering of transmembrane protein LAT (Linker for activation of T cells) activates signaling to produce cytokines when T-cell receptors are exposed to an antigen,\textsuperscript{10} while protein Whi3 in fungi condenses with RNA on the endoplasmic reticulum.\textsuperscript{11}
through homotypic interactions. Many systems correspond to this description, but others - especially in 2D - do not and feature mainly specific heterotypic interactions, which are possibly of considerable strength.

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, have been studied less. Predictions often rely on classical kinetic theory, 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.

In this paper, we study the kinetics of aggregation of a binary mixture with heterotypic interactions. The two species, which we name A and B, represent two kinds of proteins that can freely diffuse in a 2D environment, such 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 cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a binding cannot. By geometrical constraints, bonds are exclusive, meaning that a bond cannot connect more than two binding sites and a bonding site already participating in a bond cannot engage in a second bond. The number of binding sites of either species is the valence, which 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 compositions and sizes. We find that the 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,

$$U_0(R) = \begin{cases} \varepsilon_0 \left[ \left( \frac{q}{R} \right)^6 - 2 \left( \frac{q}{R} \right)^4 + 1 \right] & \text{if } R < \sigma, \\ 0 & \text{if } R \geq \sigma, \end{cases}$$

where $R$ is the distance between the centers of the two particles, $\sigma$ is their diameter, and $\varepsilon_0$ is chosen to be equal to $10kT$. The potential $U_0$ is purely repulsive and a standard choice when modeling volume exclusion. 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 of $0.475\sigma$ from the center of the particle. Patches on an A molecule interact only with patches on a B molecule, with the following attractive potential:

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

where $r$ is the distance between the centers of the two patches, $\varepsilon$ is the bond strength and is a parameter that we vary, $r_{in} = 0.05\sigma$ is the range within which the interaction energy is strongest, and $r_{out} = 0.15\sigma$ 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 $r < r_{out}$. The chosen geometry and parameter values for $\varepsilon_0, \sigma, r_{in}$, and $r_{out}$ finely prevent an A molecule from forming more than one bond with the same B molecule, and vice versa. To ensure the rigid motion of the patches with the center of the particle, the forces acting on the $q + 1$ atoms of the rigid body are summed to obtain a force acting on their center of mass and a torque relative to their center of mass, at each time step.

Simulations are run in the LAMMPS Molecular Dynamics Simulator (input scripts are available) and visualized with OVITO. 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 atoms$^{-2}$. This value of density is physiologically plausible for transmembrane proteins, which often bind to cytoplasmic proteins of variable density, localized close to the membrane. The time step 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 simulation is run for $5 \times 10^4$ time steps $\tau_s$, a time large enough to allow cluster growth and nucleation, when present. To mimic an implicit solvent, the system is subjected to a Langevin thermostat of friction coefficient $\gamma = 1m/\tau_0$. Note that for times much larger than $1\tau_0 = 100\tau_s$, the mass is effectively

![FIG. 1. Sketch of particles of type A (light blue) and type B (dark blue) for valence $q = 3, 4,$ and 5. The big disks, of radius $\sigma$, represent volume exclusion. Patches on A particles interact only with patches on B particles: when two interacting patches, of radius $0.1\sigma$, superimpose, the energy gain is maximum and equals $-\varepsilon$. The dashed lines delimit the interaction area, of radius $0.3\sigma$: two patches attract each other only if their interaction areas intersect. See Sec. II for details.](image-url)
merely rescaling time by a multiplicative factor; hence, comparing simulations for particles at different \( q \) and, therefore, different mass \((q + 1) m\) is safe, as for the probed valences, the rescaling factor is always close to 1. Finally, every point in the plots shown is an average over at least 20 realizations of the simulation, run with different random number generator seeds.

We analyze 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 realizations of the simulation. The fraction of clustered particles involved in one bond only (Fig. 4) is defined as the proportion of single-bound particles vs the total number of particles in a given cluster, averaged over all clusters bigger than 30 particles, from different realizations. For all these variables, error bars represent the standard deviation of the averages.

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:

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

where \( R \) is again the distance between the centers of the two particles, \( \varepsilon_a \) is the isotropic attraction strength, and \( R_a \) is the isotropic attraction range. As described in Sec. III D, the presence of such attraction allows for cluster formation at lower bond strengths \( \varepsilon \). Particles can then move around and change neighbors, by breaking bonds and forming new ones, without having to leave their cluster.

### 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 five 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 aggregates,\(^6\) 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 0.25\( n_A \) to 4\( n_A \).

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.

The extent of clustering is completely described by the cluster size distribution, a quantity that has recently received increasing attraction\(^7\),\(^8\) and that we report in the supplementary material. For practical reasons, we use here as a probe the mean of such distribution, computed at the last time frame of our simulations (see Sec. III). 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 \( \varepsilon \). First, the plots show that clusters do not form within the simulation time (or do not grow) if the binding energy is below a certain threshold, depending on \( q \). Such a threshold is \( \approx 8 \) and \( 6kT \) for \( q = 3 \) and 4, respectively. This is because, in the dense phase, the 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 phase.\(^9\)

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 \( \varepsilon \) above the clustering threshold. Indeed, if there are...
FIG. 3. Average cluster size, at different binding energy values $\epsilon$ (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 Sec. II). 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 $\epsilon$ (indicated by the symbols).

Particular attention needs to be devoted to the equal concentration case, where we observe a non-monotonicity in the cluster size with $\epsilon$ as well. Such a phenomenon is analogous to magic-number effects reported in previous studies of many 3D systems or models, where fine tuning of valence and/or concentration can cause major variations in the condensate properties. This is shown 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 $\epsilon$, but hampered for low or strong $\epsilon$. This is again because lower $\epsilon$ allows for 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 the cluster snapshots in Fig. 3). The 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 behavior is more pronounced at strong than at weak binding $\epsilon$. 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 reorganization and formation of new links; this favors 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 $\epsilon$ (at least far from the equal molar ratio): if $\epsilon$ 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 $\epsilon$ is too large, the system gets stuck in a state where capping prevents coarsening.
more bond reorganization; indeed, clusters at intermediate $\epsilon$ appear more compact and with fewer defects (see the clusters snapshots in Fig. 2 at different bond strengths). This is particularly important at equal molar ratios, when, to gain the limited 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 over-abundant and A sites can be saturated anyway, by forming small B-coated clusters as described above; in such a case, lower $\epsilon$ just means a lower bond lifetime 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 might wonder whether the fully saturated state we observe at large concentration ratios is due to kinetic arrest or is a thermodynamic minimum. To compare the free energy of two limit states, sketched in Fig. 5. For simplicity, we discretize space: while making predictions more qualitative than quantitative, this largely simplifies free energy estimates. We 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. We restrict to the case $q = 3$ or 4, corresponding to a regular lattice: this avoids the geometrical complications related to amorphous clusters ($q = 5$).

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

$$F_I = -q n_A - k T \left[ \ln \frac{V}{n_B} + (n_B - n_A) \ln \left( \frac{n_B - n_A}{V} \right) \right].$$

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 $n_B - q n_A$ free B particles. Clearly, this scenario is only possible if $n_B > q n_A$. The related free energy is approximately

$$F_{II} = -q n_A - k T \left[ n_A - n_A \ln \left( \frac{n_A}{V} \right) + (n_B - q n_A) \ln \left( \frac{n_B - q n_A}{V} \right) \right].$$

The first two terms in square brackets represent the entropy of the tiny clusters, while the last two represent 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 (the first term in both equations) 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 favored 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 $kT$ 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 $-q n_A$ 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 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, we can group configurations by cluster perimeter and find that the cluster of minimal perimeter is the most favorable 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. A kinetic lock likely helps making such a state the observed one at large concentration ratios.

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 Sec. 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 $\epsilon$, provided that enough time is allowed for nucleation at intermediate $\epsilon$. 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 initialized in a metastable state, namely, at a point of the phase diagram between the spinodal and the binodal; the absence of nucleation indicates, instead, that the system is unstable and the relaxation process is a standard spinodal decomposition.

In the latter case ($\epsilon > 7kT$), the average number of particles in a cluster...
scales approximately as $t^{0.6}$, corresponding to a cluster diameter growing with exponent $\alpha \approx 0.6/2 = 0.3$. Provided that $n_B/n_A$ stays close to unity, this is true also for $q = 3$ and 5, as shown in Fig. 7, where values of the cluster diameter growth exponent $\alpha$ obtained from fitting cluster size curves at late times are reported. A power law behavior is a sign of self-similarity, which typically characterizes relaxation to equilibrium of a quenched interacting system (here quenching means initializing the system in the vapor state, but running the simulation at temperatures lower than critical, where the equilibrium state exhibits phase separation).

For low or high concentration ratios (left and right panels in Fig. 6), the aggregation kinetics is slower. At small $\epsilon$, this might be due to larger nucleation times, but linear behavior at large times rather points at a slowdown in the power-law growth rate. This is consistent with the fact that at large $\epsilon$, coarsening halts after a given time, as confirmed by the time evolution of the cluster size distribution (see the supplementary material). This observation reflects the more pronounced non-monotonicity observed at final time for large bond strengths in Fig. 3. Figure 7 captures the decay of exponent $\alpha$ 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 addition (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 a steady state on a time scale smaller than those represented in Fig. 6 ($<10^5 \tau_r$).

A value of the exponent $\alpha \approx 0.3$ 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. 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. 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, although values of $\alpha$ between 1/4 and 1/3 have also been observed when hydrodynamics is suppressed. 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.

We provide a simple theoretical argument, analogous to the one proposed in Ref. 57, showing why $\alpha$ 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 $\rho_M$ (the 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, $\rho_M = \rho_1/M$, where $\rho_1$ is the total number of molecules per unit surface. The average distance between two clusters at a given time is then $t_{M} \approx \rho_1^{-1/2} = (M/\rho_1)^{1/2}$. 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 $t_M \approx r_M^2/D_M \sim M^2$. Neglecting

![Figure 6](image6.png)

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

![Figure 7](image7.png)

FIG. 7. Growth exponent $\alpha$ measured from a linear fit of cluster growth curves in time (Fig. 6). At concentration ratios close to 1, $\alpha \approx 0.3$ for most curves. Away from this regime, the growth is hindered: coarsening is slowed down for $\epsilon = 10kT$ (dashed lines) and completely halted for $\epsilon = 20kT$ (continuous lines).
fractal effects, the mass of the clusters scales as the square of its linear dimension \( \ell_M \) (its diameter, for instance): \( M \sim \ell_M^D \), hence \( \epsilon_M \sim \ell_M^D \). 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 analyzed 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 the material and condensation of the 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 diffusivity.18

To obtain fluidity, we add a generic isotropic attraction between any two particles (homo- and heterotypic), parameterized by its strength \( \epsilon_a \) and range \( R_a \) [see Sec. II, 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 diffusivity \( D_{\text{int}} \)) 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 \( \epsilon \) and isotropic attraction \( \epsilon_a \), for a chosen range \( R_a \). While in the large \( \epsilon \) regime particles are locked by unbreakable bonds, there is a window where diffusivity is large. This occurs at small enough \( \epsilon \), so that bonds can be broken, yet large enough \( \epsilon_a \), 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 behavior of the average cluster size with concentration ratio, in the case with fluid clusters. When the proportion of energy coming from isotropic attraction is too high (large \( \epsilon_a \)), this behavior 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 a 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 condensates17,18 that do not belong to the client-scaffold category.

Such a size-control process, although it can lead to stationary states within the observation time (Fig. 6), is not of obvious thermodynamic origin. Analogous processes have been recently identified, for different systems,8,20,40,46 where the origin of the non-monotonicity was attributed to equilibrium physics.20,46 We show that it can, in fact, 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 \( \sim 0.3 \) to 0, upon an increase in 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. Nonetheless, we still observe significant concentration-dependent size control for any value of bond strength. Our analysis confirms the importance of kinetics in phase-transitions, consistent with recent observations.19

Yet another way to control the cluster size is by modulating valence. This strategy allows cells to tune properties of physiological condensates,19,28 particularly on membranes,59 but is also the functioning mechanism of artificial optogenetic tools used to study phase-separation.60 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 the cluster size and growth exponent (Figs. 3 and 7).

Similarly to cluster sizes, concentration ratios and bond strengths 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 the presence of heterotypic interactions only.

In conclusion, we highlight how clustering curves as a function of the 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 the concentration ratio might as well be a signature of dominantly heterotypic interactions.

SUPPLEMENTARY MATERIAL

See the supplementary material for a discussion of the cluster size distribution.

ACKNOWLEDGMENTS

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

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The code to generate input LAMMPS scripts used for this paper is openly available on GitHub.\(^{53}\)

APPENDIX: CALCULATION OF CONSTRAINED FREE ENERGY

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

\[
Z = \frac{1}{n_A! n_B!} e^{\beta q n_A \mathcal{C}},
\]

(A1)

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

\[
\mathcal{C} = \sum_{n_{BC}} V n_A! n_{BC}! \frac{(V - n_A - n_{BC}) n_B! (n_B - n_{BC})}{n_A! n_{BC}!} \mathcal{V}_{n_{BC}}.
\]

(A2)

The seven factors of this sum represent, respectively, the positions of the cluster in a lattice with \( V \) sites, the permutations of the \( n_A \) A particles in the cluster, the permutations of the \( n_{BC} \) B particles in the cluster, the positions of the \( n_{BF} \) free B particles (to be chosen among the sites not occupied by the cluster), the permutations of the \( n_B \) free B particles, the ways to choose which B particles are free and which belong to the cluster, and finally \( \mathcal{V}_{n_{BC}} \) the 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 the 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. Equation (A2) is then interpretable as a sum over all possible perimeters, and \( \mathcal{V}_{n_{BC}} \) is the number of different shapes of semiperimeter \( p = n_{BC} - n_A \) that can be built by clustering \( 2n_A \) particles.

The statistics of connected shapes on lattices, the so-called lattice animals or polyominoes, has been a long studied subject.\(^{51}\) \( \mathcal{V}_{n_{BC}} \) can be written as the product of the total number \( a_s \) of shapes of size \( s = 2n_A \) (scaling as \( 4.01^s \))\(^{33,51} \) 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 behavior of the latter quantity does not exist, so we extracted \( f_p \) for \( s \) up to 48 from exact tabulated enumerations of lattice animals.\(^{48,49,61} \) Figure 9 shows that semiperimeters, rescaled by \( \bar{p} \approx 0.95 \sigma^{0.80} \), seem to follow approximately a normal distribution with mean \( \mu = 1 \) and variance \( \sigma = 0.06 \).

In summary, Eq. (A2) can be rewritten as

\[
\mathcal{C} = n_A! n_B! V a_{2n_A} \sum_p \left( \frac{V - 2n_A - p}{n_B - n_A - p} \right) f_p.
\]

(A3)
In the sum, notice the competition between the second term, representing "shape" entropy (more shapes are available at a larger perimeter), and the first term, representing the translational entropy of free B particles (a larger cluster perimeter means fewer B particles in solution). Using Stirling's approximation and passing to the continuous limit for $p$, we obtain

$$\mathcal{C} \sim n_A! n_B! V^{2n_A/p} e^{-V n_A \epsilon_0 (V - n_B \epsilon_0)} \sum_p \exp \left[ \left( \frac{V - 2n_A - p}{2} \right)^2 - (n_B - n_A - p) \ln(n_B - n_A - p) - \frac{(p/p - \mu)^2}{2\sigma^2} \right]$$

(A4)

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 $a_{2n_A/p} = \delta_{p,0}$ (Kronecker delta) in Eq. (A3). By doing so, one recovers Eq. (4), which refers to a circular cluster.

REFERENCES


See https://github.com/ipalai/Clusters2DGit/ for more information about the simulation setup and LAMMPS input files.


See http://wasd.urz.uni-magdeburg.de/mertens/research/animals/ for more information about lattice animals and associated data; accessed on 30 September 2021.