Soft Matter

PAPER

View Article Online
View Journal | View Issue



Cite this: *Soft Matter*, 2026, **22**, 109

# Micellization as a connectivity transition: a topological Ising model with a hydrophobic constraint

Vicente Domínguez-Arca *[ab]

Received 14th August 2025,
Accepted 19th November 2025

DOI: 10.1039/d5sm00832h

rsc.li/soft-matter-journal

Micellization is commonly described as a collective response driven by the hydrophobic effect. Here we propose and study a topological Ising model that abstracts this effect as a solvent-exclusion constraint defined purely by local connectivity. On a square lattice with binary occupancy (amphiphiles/water), we characterize neighborhood by a topological kernel of radius $R$ and metric (Chebyshev or Manhattan). A water site becomes "restricted" when the local overlap with amphiphiles, computed *via* convolution with the kernel, exceeds a fixed threshold. The system energy is $F = N_{restr}$; we set $\alpha = 1$ by design, working in dimensionless units that prevent interpreting $\alpha$ as carrying any metric information. Dynamics are explored with Metropolis updates at temperature $T$. Control parameters are amphiphile density $\rho$, temperature $T$, the metric, and $R$. As an order parameter we use $S_{max}/N_a$, the fraction of amphiphiles in the largest connected cluster. In the surveyed ranges we observe, for more connective kernels (*e.g.*, Chebyshev with $R \geq 3$), the emergence of a giant component in finite regions of ($\rho$, $T$), while less connective configurations (*e.g.*, Manhattan with $R = 1$) do not aggregate in the same window. These results support the view that micellization, in this framework, is a connectivity transition governed by the topology of local interactions rather than by explicit metric scales. We discuss implications and routes for quantitative comparisons with experiments and more detailed simulations.



## 1 Introduction

Micelle formation is a paradigmatic collective phenomenon in amphiphilic systems.[1–7] It is traditionally rationalized as a mesoscopic manifestation of the hydrophobic effect and quantified *via* critical concentrations and aggregation curves.[8–11] A complementary route is to represent micellization with lattice-based spin or occupancy models, in which the relevant physics is distilled into local rules that reproduce the emergence of large-scale aggregated structures.[12–15] In this work we deliberately adopt a topological perspective, removing explicit metric dependencies, to ask to what extent micellization can be understood as a connectivity transition[16–19] controlled by the local degrees of freedom of the solvent and by the geometry of amphiphile neighborhoods.

Our starting point is a binary Ising-like model on a square lattice,[4,18,20–22] with two fields: amphiphile occupancy and water. The hydrophobic effect[23,24] is abstracted as a local restriction acting on water: a solvent voxel is deemed restricted when the amphiphile occupancy in its neighborhood, evaluated *via* convolution with a topological kernel of radius $R$, exceeds a fixed threshold. Neighborhood is defined without explicit physical distances, using discrete metrics (Chebyshev or Manhattan) that determine the kernel connectivity; operationally, this is a morphological dilation of the amphiphile set by a binary structuring element.[25] The energy of a configuration is the total number of restricted sites, $F = N_{restr}$, and we set $\alpha = 1$ by construction; this puts the model in natural (dimensionless) units and rules out any reading of $\alpha$ as encoding inter-site distances. The stochastic evolution follows a Metropolis scheme at temperature $T$.

Within this approach, the control parameters are: (i) the amphiphile density $\rho$, (ii) the temperature $T$, (iii) the topological metric, and (iv) the kernel radius $R$. The phase order parameter we monitor is $S_{max}/N_a$, the fraction of amphiphiles in the largest cluster, complemented by auxiliary measures (number of clusters, fraction of restricted water). Our working hypothesis is that, above critical thresholds of connectivity and occupancy, a giant component appears that we identify with the micellar state, and that the location of this transition depends primarily on the topology of the kernel (metric and $R$), rather than on any metric notion of distance. This picture is consistent with the view of hydrophobic assembly as a solvent-mediated, entropy-driven reorganization whose character

[a] *Biosystems and Bioprocess Engineering (Bio2Eng) Group, Institute of Marine Research of Spanish Research Council, IIM-CSIC, C/Eduardo Cabello 6, Vigo, 36208, Pontevedra, Spain. E-mail: vdominguez@iim.csic.es, vicente.dominguez@uni-bielefeld.de*
[b] *Faculty of Chemistry, Physical and Biophysical Chemistry, Bielefeld University, Universitätsstr. 25, Bielefeld, 33615, Nordrhein-Westfalen, Germany*

depends on connectivity and local crowding[9,10,26,27] and with the generic coarsening phenomenology expected for phase-ordering under curvature-driven dynamics.[28]

The contributions of this article are: (1) to formulate an Ising model with a purely topological energy (no explicit length scales) that captures aggregate formation; (2) to show, through systematic explorations in ($\rho$, $T$) and across kernels, that micellization emerges as a connectivity transition when the kernel is sufficiently connective; (3) to establish an analysis protocol based on $S_{\max}/N_a$ and phase maps that cleanly separate aggregated and non-aggregated regimes; and (4) to provide a basis for comparisons with continuum descriptions and experimental data while preserving an interpretation of micellization as a reorganization of solvent degrees of freedom induced by the local topology of amphiphiles.[12–14]

The remainder of the manuscript is organized as follows. In Section 2 we detail the kernel definition, the restriction threshold, and the Monte Carlo dynamics, as well as the observables used. Section 3 presents phase maps and representative cuts that illustrate the emergence of a giant component for different choices of metric and $R$. In Section 4 we discuss implications and limitations (finite size, threshold choice, dimensionality, polydispersity) and outline extensions (three-dimensional lattices, anisotropic kernels, couplings to continuum fields) together with strategies for quantitative confrontation with experiments.

# 2 Model and methods

## 2.1 Lattice, fields, and control parameters

We consider a square lattice of size $L \times L$ with binary occupancy

$$A \in \{0, 1\}^{L\times L}, \quad A_{ij} = 1 \text{ (amphiphile)}, \quad A_{ij} = 0 \text{ (water)}.$$

The amphiphile loading is $N_a = \sum_{ij} A_{ij}$ and the areal density is $\rho = N_a/L^2$. Unless otherwise stated, boundaries are open (zero-padded convolutions, no wrap-around).

Control parameters are: (i) $\rho$ (equivalently $N_a$ at fixed $L$), (ii) the reduced temperature $T$ of the Metropolis dynamics, (iii) a topological metric metric $\in$ {Chebyshev, Manhattan}, and (iv) a connectivity radius $R \in N$ that sets the neighborhood size.

## 2.2 Topological kernel and restriction operator

Let $A \in \{0, 1\}^{L\times L}$ denote the amphiphile occupancy ($A_{ij} = 1$ for amphiphile, $A_{ij} = 0$ for water) and $N_a = \sum_{i,j} A_{ij}$ the total number of amphiphiles. For a chosen metric and radius $R$, we define a binary kernel $K_R^{(\text{metric})} \in \{0, 1\}^{(2R+1)\times(2R+1)}$ by

$$K_R^{(\infty)}(\Delta x, \Delta y) = \mathbf{1}(\max\{|\Delta x|, |\Delta y|\} \le R),$$

$$K_R^{(1)}(\Delta x, \Delta y) = \mathbf{1}(|\Delta x| + |\Delta y| \le R),$$

corresponding to Chebyshev ($\ell_\infty$) and Manhattan ($\ell_1$) balls, respectively.

Let $*$ denote the discrete (zero-padded) convolution on the lattice. We compute the local amphiphile overlap as

$$C \equiv A \times K_R^{(\text{metric})}.$$

A water site $(i, j)$ is deemed restricted whenever the neighborhood contains at least one amphiphile; with threshold $\theta = 0.5$ this is captured by

$$\mathcal{R}_{ij}(A) = (1 - A_{ij})\mathbf{1}\,(C_{ij} > \theta), \quad \theta = 0.5.$$

The total number of restricted water sites is

$$N_{\text{restr}}(A) = \sum_{i,j} \mathcal{R}_{ij}(A).$$

This operator is equivalent to the morphological dilation of $A$ by $K$ restricted to water sites and provides a purely topological proxy for solvent cages around amphiphilic domains. The convolution kernel is not merely a mathematical device, but a physical proxy for the solvent's sensitivity to local amphiphile density: it defines the topological footprint through which amphiphiles collectively restrict water degrees of freedom. Aggregation then emerges as a cooperative strategy to minimize the union of these footprints—a solvent-driven, entropic mechanism encoded through connectivity alone.

## 2.3 Energy functional (fixed $\alpha = 1$)

The configuration cost is taken to be the count of restricted water sites,

$$F[A] = \alpha N_{\text{restr}}(A), \quad \alpha \equiv 1,$$

*i.e.*, we work in natural, dimensionless units where $\alpha$ carries no metric content. Lowering $F$ corresponds to sharing restriction shells *via* aggregation (reducing the union area of restricted water).

**2.3.1 Monte Carlo updates.** Given a trial move $A \to A'$, the Metropolis acceptance probability is

$$P_{\text{acc}} = \min\{1, \exp[-\Delta F/T]\}, \quad \Delta F \equiv F[A'] - F[A],$$

with $T$ the dimensionless temperature.

**2.3.2 Cluster connectivity and order parameter.** On the amphiphile subgraph we define connectivity to match the analysis topology: 4-neighbors for Manhattan and 8-neighbors for Chebyshev (unless stated otherwise). Let $\{s_\ell\}$ denote the sizes of the connected components (clusters) under that rule; then

$$S_{\max} = \max_\ell s_\ell, \quad \phi \equiv \frac{S_{\max}}{N_a}.$$

We use $\phi$ as the order parameter of aggregation and classify a state as aggregated whenever

$$\phi \ge \tau, \quad \text{with } \tau = 0.2 \text{ fixed in all analyses.}$$

**2.3.3 Restricted-water fraction.** As an auxiliary observable we monitor the fraction of water sites that are restricted,

$$f_{\text{restr}}(A) = \frac{N_{\text{restr}}(A)}{L^2 - N_a},$$

which quantifies the degree of solvent confinement. Typically $f_{\text{restr}}$ decreases when amphiphiles aggregate, reflecting the overlap (and hence reduction) of restricted zones around clustered amphiphiles.

 

### 2.4 Stochastic dynamics (metropolis with global swap proposals)

Each sweep proposes up to $\min(N_a, L^2 - N_a)$ exchanges between amphiphile and water sites chosen uniformly at random over the lattice: pick $p \in \{(i,j): A_{ij} = 1\}$ and $q \in \{(i,j): A_{ij} = 0\}$ and form $A'$ by swapping $A_p \leftrightarrow A_q$. Because proposals are symmetric, the Metropolis acceptance probability reads

$$P_{acc}(A \to A') = \min\left\{1, \exp\left(-\frac{\Delta F}{T}\right)\right\}, \quad \Delta F = F[A'] - F[A].$$

This move set conserves $N_a$, explores non-local rearrangements efficiently, and leaves the Boltzmann distribution $\propto e^{-F/T}$ invariant. Initial conditions are random placements of $N_a$ amphiphiles.

### 2.5 Cluster topology and observables

Connectivity among amphiphiles is defined on the occupancy graph induced by $A$ with either 4-neighborhood (von Neumann) or 8-neighborhood (Moore), set by an analysis flag. Unless otherwise noted, cluster connectivity matches the analysis topology: 4-neighborhood for Manhattan and 8-neighborhood for Chebyshev. We verified that using 8-neighborhood throughout yields the same qualitative phase maps. Let $\{s_\ell\}$ be the sizes of connected components (clusters). We record:

- Largest-cluster fraction

$$\frac{S_{max}}{N_a} = \frac{\max_\ell s_\ell}{N_a},$$

used as the order parameter for aggregation.

- Number of clusters $n_{cl}$ and the cluster-size histogram $n(s)$ (counts of clusters of size $s$).
- Restricted-water fraction

$$f_{restr} = \frac{N_{restr}(A)}{L^2 - N_a},$$

a geometric proxy for solvent cages.

For phase classification we define an aggregated state if $S_{max}/N_a \geq \tau$, with default threshold $\tau = 0.2$ used later in phase maps and summaries. Time-domain behavior is consistent with these classifications; see the kinetic traces in Fig. 3.

### 2.6 Scanning protocol and computational details

Parameter scans iterate over $(\rho, T)$ at fixed (metric, $R$) and $\alpha = 1$, averaging over seeds when applicable; typical grids use $L = 128$, $T \in \{0.25, 0.5, 1.0\}$, $R \in \{1, 2, 3\}$, and Chebyshev/Manhattan metrics. Each run records time series and summary statistics ($F$, $f_{restr}$, $S_{max}/N_a$, $n_{cl}$) and exports the final lattice and $n(s)$ for post-processing. Convolutions are computed *via* FFT (complexity $\mathcal{O}(L^2 \log L)$); cluster labeling is performed on CPU; optional GPU acceleration does not change model semantics.

### 2.7 Phase classification

Throughout this work we keep $\alpha = 1$ fixed. We classify a simulation as aggregated whenever the largest-cluster fraction satisfies $S_{max}/N_a \geq \tau$ with a fixed threshold $\tau = 0.2$; otherwise it is labelled non-aggregated.

## 3 Results

### 3.1 Overview

With $\alpha$ fixed to 1 throughout, we scanned amphiphile density $\rho$ (equivalently $N_a$ at fixed $L$), temperature $T$, the topological metric (Chebyshev or Manhattan), and the kernel radius $R$. Across this four-dimensional grid we find clear presence/absence regimes of aggregation as quantified by the largest-cluster fraction $S_{max}/N_a$. In particular, more connective kernels (*e.g.*, Chebyshev with $R \geq 3$) exhibit extended regions in $(\rho, T)$ where a giant component emerges, whereas less connective settings (*e.g.*, Manhattan with $R = 1$) do not show aggregation over the same scan (see figures).

This pattern substantiates the central hypothesis: amphiphile self-assembly in this model is primarily a connectivity phenomenon. Euclidean distance enters only implicitly—through the short-range neighborhood encoded by (metric, $R$)—and is otherwise secondary once $\alpha$ is fixed. Operationally, aggregation minimizes the effective cost $F = N_{restr}$ by sharing solvent-restriction shells, which is equivalent to maximizing the solvent's accessible microstates at fixed $(\rho, T)$. Once locality is fixed through the short-range neighborhood (metric, $R$), the controlling variable for assembly is connectivity rather than explicit Euclidean spacing; see Section 4.3 for a formal discussion.

### 3.2 Reading the aggregation maps ($\alpha = 1$, $\tau = 0.2$)

Unless otherwise noted, all maps use a fixed aggregation threshold $\tau = 0.2$. Fig. 1 and 2 summarizes the phase behavior at fixed $\alpha = 1$ as a function of neighborhood topology (metric and $R$), temperature $T$, and amphiphile count $N_a$. Each cell is colored by the final $S_{max}/N_a$; a checkmark denotes aggregated states ($S_{max}/N_a \geq \tau$), a cross non-aggregated states, and hatched cells were not scanned. The vertical dashed line marks $T = 1$.

**Chebyshev.** With Moore neighborhoods, aggregation is widespread as connectivity increases. For $R = 3$ all scanned states with $T \leq 1$ and $N_a \geq 800$ aggregate. For $R = 2$, aggregation already appears at $T = 1$ for $N_a = 400$–$800$, and also at $T = 0.5$ for $N_a = 400$. Thus, at fixed $\alpha$, enlarging the neighborhood promotes a giant component even at moderate density.

**Manhattan.** With taxicab neighborhoods, connectivity is more restrictive. For $R = 1$ no aggregation is observed in the explored window, whereas increasing the neighborhood to $R = 2$ recovers aggregation at lower $T$ and/or higher $N_a$ (*e.g.*, $T = 0.5$, $N_a = 400$ and $T = 1$, $N_a = 400$–$800$).

Taken together, these maps support the central claim of this work: once $\alpha$ is fixed, aggregation is primarily controlled by the topology of the neighborhood rather than by Euclidean spacing.

### 3.3 Phase structure in $(\rho, T)$ at fixed (metric, $R$)

For Chebyshev kernels with $R = 3$, the $S_{max}/N_a$ heatmaps display a sharp transition from dispersed to aggregated states upon increasing $\rho$ and/or lowering $T$ (see Fig. 1). The corresponding binary phase masks (using the criterion $S_{max}/N_a \geq \tau$ with $\tau = 0.2$) produce compact aggregated regions in $(\rho, T)$. By contrast, for Manhattan $R = 1$ the same analysis yields low $S_{max}/N_a$ across

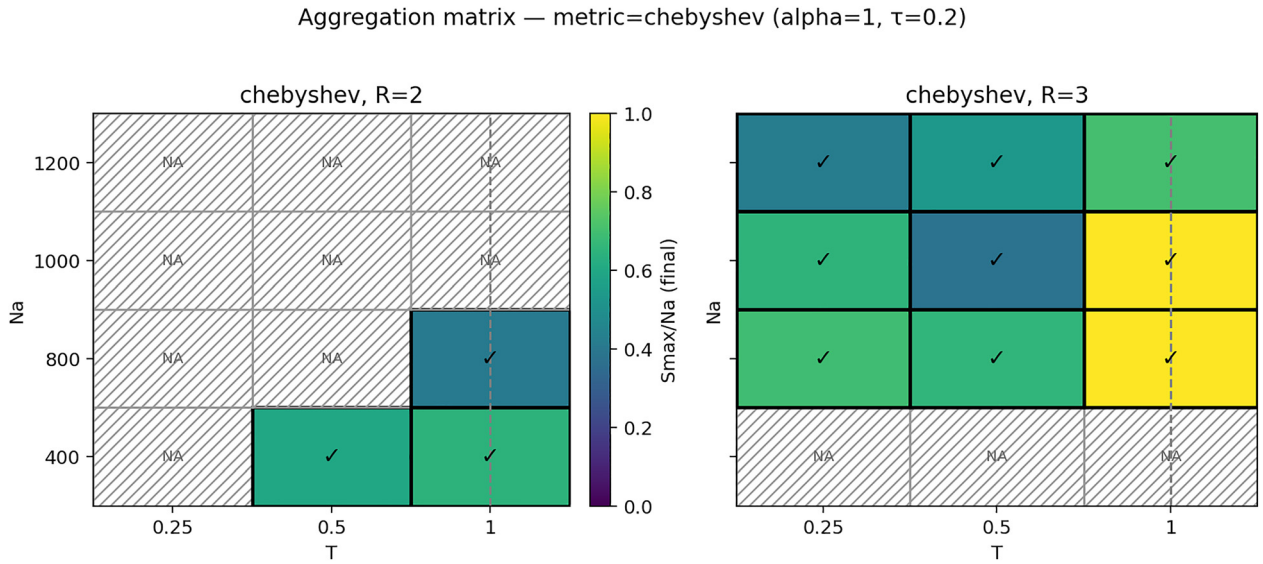


Fig. 1 Chebyshev (panels as drawn for $R = 2$ and $R = 3$). Aggregation matrices at fixed $\alpha = 1$ and $\tau = 0.2$. Cells show $S_{max}/N_a$; ✓ indicates aggregated ($\geq \tau$), × non-aggregated; hatched cells were not scanned. The dashed vertical line marks $T = 1$.

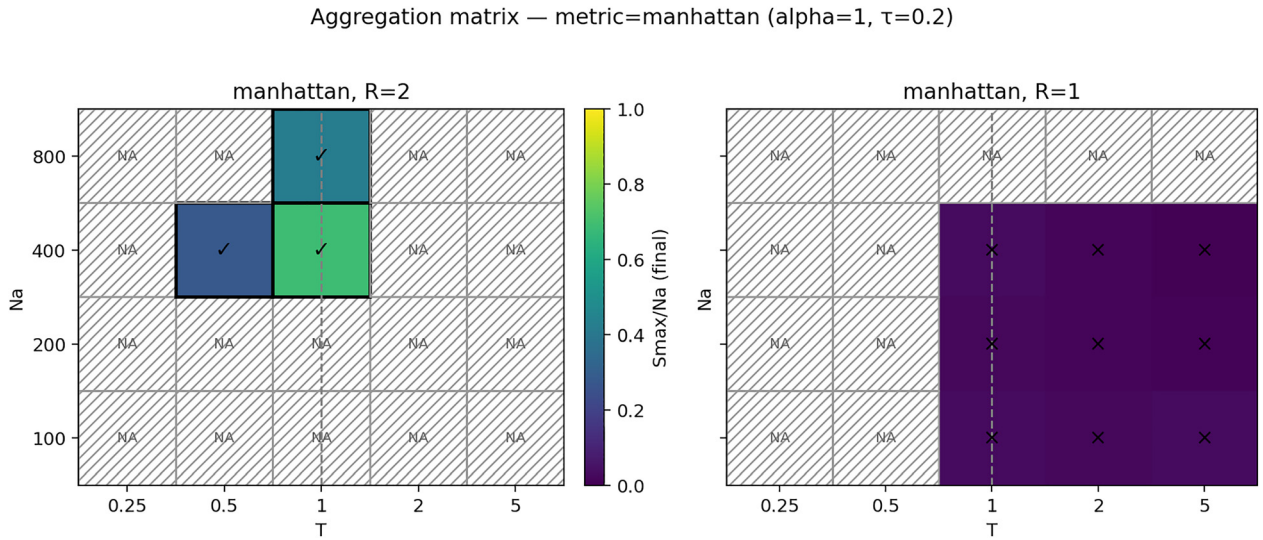


Fig. 2 Manhattan (panels as drawn for $R = 2$ and $R = 1$). Aggregation matrices at fixed $\alpha = 1$ and $\tau = 0.2$. Cells show $S_{max}/N_a$; ✓ indicates aggregated ($\geq \tau$), × non-aggregated; hatched cells were not scanned. The dashed vertical line marks $T = 1$.

the grid and essentially empty aggregated masks (see Fig. 2). Intermediate radii and metrics interpolate between these extremes, shifting the apparent threshold surface in ($\rho$, $T$) as connectivity increases with $R$ or with the choice of metric.

### 3.4 Order parameter and restricted-water coupling

The aggregated regime correlates with a marked decrease of the restricted-water fraction $f_{restr}$. This is the expected topological signature of an entropically driven assembly: when amphiphiles coalesce, the union of their restriction neighborhoods contracts, fewer water sites are restricted, and the solvent gains degrees of freedom (see figures comparing $S_{max}/N_a$ and $f_{restr}$ across the same grids). Importantly, this mechanism is encoded without invoking explicit metric distances; it follows purely from neighborhood connectivity and occupancy.

### 3.5 Cluster statistics near the transition

Cluster-size histograms $n(s)$ broaden in parameter regions where $S_{max}/N_a$ rises rapidly (see figures). Away from aggregation, $n(s)$ is dominated by small clusters; near the boundary, intermediate sizes proliferate; deep in the aggregated phase, a single giant component carries most amphiphiles and the remainder is a sparse gas of small clusters. These qualitative changes are consistent with a connectivity transition on a finite lattice. A finite-size susceptibility (*e.g.*, the standard second-moment proxy excluding $S_{max}$) is expected to peak near the boundary;

 

while we do not report it here, the recorded $n(s)$ enables such analysis.

### 3.6 Connectivity controls aggregation: metric and radius

At fixed ($\rho$, $T$), switching from Manhattan to Chebyshev and/or increasing $R$ effectively increases kernel connectivity. This shifts the phase boundary towards higher $T$ or lower $\rho$, expanding the aggregated domain (see comparative figures by metric and radius). Conversely, low-connectivity kernels suppress aggregation even at relatively high $\rho$, underscoring that topological reach, rather than Euclidean spacing *per se*, is the proximate control knob in this model.

### 3.7 Kinetic traces and equilibration (qualitative)

Time-domain behaviour corroborates the phase calls from the ($T$, $N_a$) maps. Fig. 3 displays three representative trajectories of the largest-cluster fraction $\phi \equiv S_{max}/N_a$ at fixed $\alpha = 1$. (i) In the aggregated regime (top) $\phi$ rises rapidly and saturates well above the classification threshold $\tau = 0.2$. (ii) A run not yet equilibrated (middle) exhibits a positive late-time slope while already exceeding (or approaching) $\tau$, consistent with ongoing coarsening. (iii) In a non-aggregated case (bottom) $\phi$ fluctuates around a flat baseline well below $\tau$, with slope compatible with zero. These kinetics justify using the final $\phi$ together with a fixed threshold $\tau$ to assign phases and illustrate the expected trends—slower growth at higher $T$ and stronger aggregation for more connective neighbourhoods (metric, $R$). Runs started from random placements converge to comparable final $\phi$ under the same ($\rho$, $T$, metric, $R$), consistent with equilibration under the Metropolis dynamics used here.

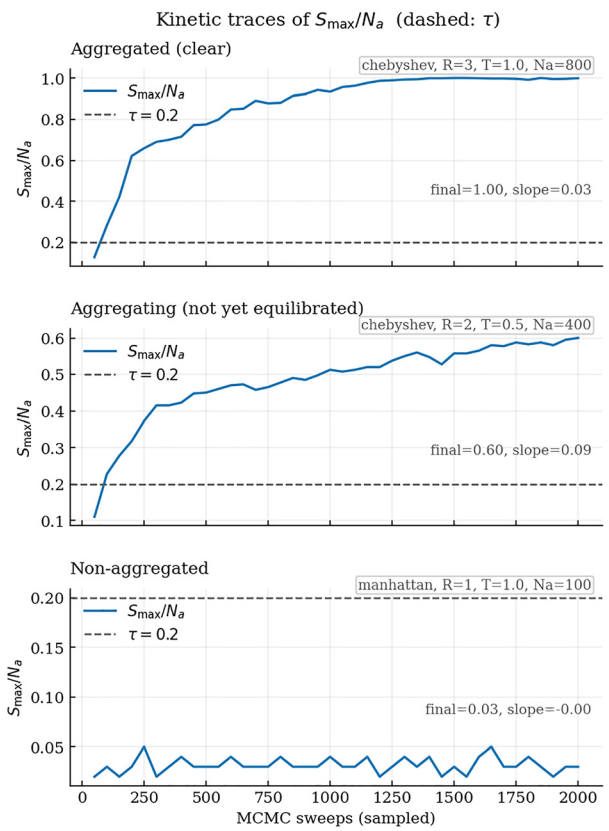


Fig. 3 Kinetic traces of $S_{max}/N_a$ (dashed: $\tau = 0.2$). Top: clear aggregation; middle: aggregation in progress (not fully equilibrated); bottom: non-aggregated. Panel annotations report (metric, $R$, $T$, $N_a$). The dashed horizontal line is the classification threshold $\tau$, and the in-panel text shows the final value and the trend slope (estimated from the last portion of the trace).

Consistent with a solvent-entropy mechanism, the kinetic signature in energy mirrors the behaviour of the order parameter. Fig. 4 shows that, at fixed $\alpha = 1$, the normalized energy $F/(L^2 - N_a) = f_{restr}$ decreases steadily as aggregation proceeds, reflecting the overlap of restricted-water regions around clustered amphiphiles. Under non-aggregating conditions it remains flat within noise. Together with Fig. 3, this dynamical evidence supports the view that aggregation reduces the union of restriction footprints, thereby lowering $F$ while increasing $S_{max}/N_a$.

Complementing the order parameter and energy views, the cluster-count kinetics in Fig. 5 display the expected coarsening signature: $n_{clusters}(t)$ decays in aggregated and still-aggregating runs, while it remains approximately constant in non-aggregated conditions. Together with Fig. 3 and 4, this consolidates the dynamic picture: aggregation ($S_{max}/N_a\uparrow$) proceeds by reducing the number of domains ($n_{clusters}\downarrow$) and the solvent restriction ($f_{restr}\downarrow$).

The real-space evolution mirrors these diagnostics. Fig. 6 juxtaposes three snapshots of an aggregating run (top) and a non-aggregating run (bottom), illustrating domain coarsening *versus* persistent dispersion at matched time points.

### 3.8 Reconciling the entropic narrative with the model's cost function

A potential source of confusion is the use of the symbol $F$ for the model's effective cost $N_{restr}$. In thermodynamic terms, the coarse-grained effect we encode as "energy" is an entropic penalty on the solvent: restricting water reduces its microstate count. Minimizing $F$ at low $T$ therefore corresponds to maximizing solvent entropy subject to the occupancy constraints. The Metropolis weight $\exp(-F/T)$ makes this precise in the model's ensemble: lower $F$ configurations are favored, as they imply fewer restricted water sites and thus more solvent degrees of freedom. Within this framing, our findings—that aggregation reduces $F$ and increases $S_{max}/N_a$ for sufficiently connective kernels—are exactly the lattice counterpart of the continuum intuition that micelles form to reduce hydrophobic interfacial "cost", *i.e.*, to liberate as much solvent entropy as possible.

**3.8.1 Clarification.** In this ensemble, $F$ is an effective solvent-centered cost that coarse-grains the loss of water microstates into a scalar penalty. It is not the microscopic free energy of amphiphile–amphiphile interactions. The Boltzmann weight $\exp[-F/T]$ therefore favors configurations with fewer restricted water sites (higher solvent entropy) without positing a direct enthalpic attraction among amphiphiles.

### 3.9 Controls and robustness

We fixed $\alpha = 1$ in all analyses so that no scale factor could be (mis)interpreted as encoding metric information. Sensitivity to (metric, $R$) is expected and desired, as these parameters define

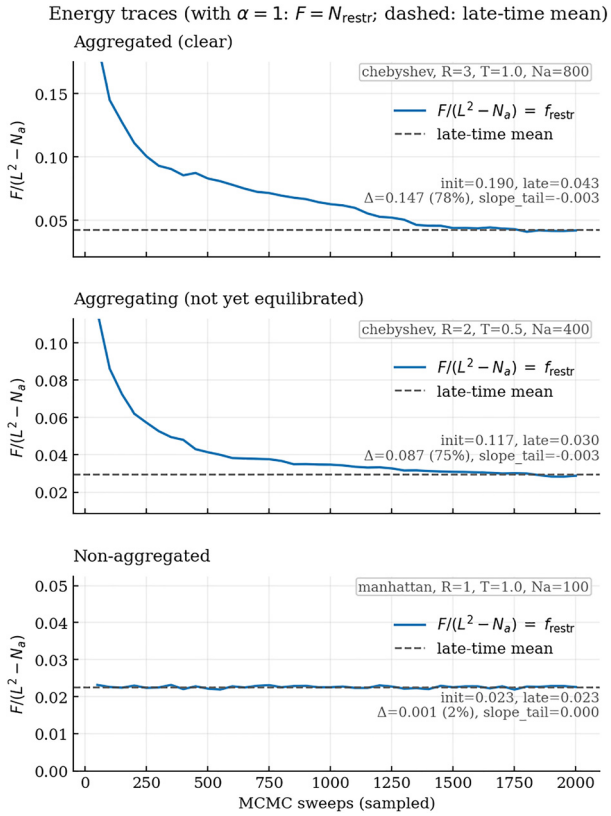


Fig. 4 Kinetic traces of the energy. With $\alpha = 1$ the energy equals the count of restricted water sites, $F = N_{restr}$. We plot the size-normalized energy $F/(L^2 - N_a) = f_{restr}$ to compare across runs with different $N_a$. Top: clear aggregation exhibits a marked decrease towards a low late-time mean; middle: a run still coarsening shows a sustained downward trend; bottom: non-aggregated conditions fluctuate around a flat baseline. The dashed horizontal line indicates the late-time mean estimated from the last fraction of the trace.

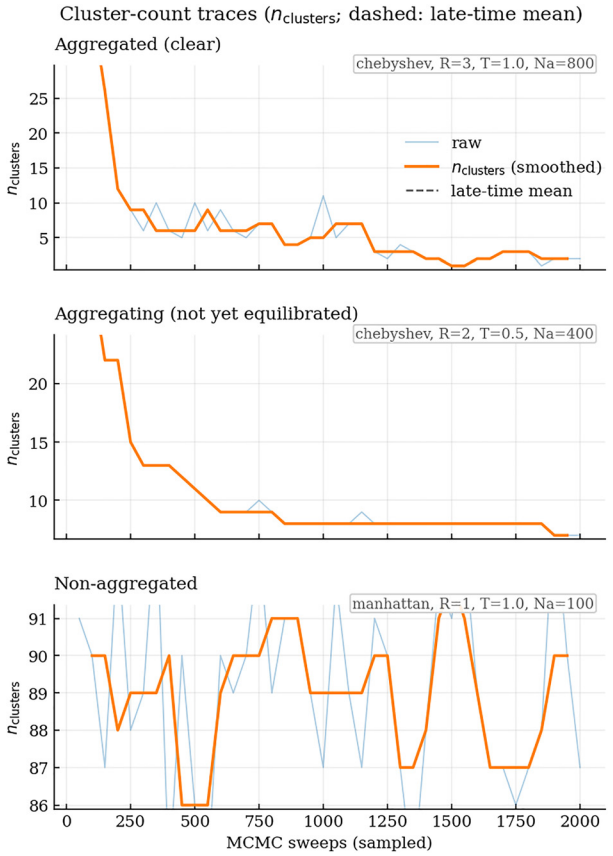


Fig. 5 Kinetics of the number of clusters. Top: clear aggregation shows a rapid decrease of $n_{clusters}$ as domains merge; middle: a sustained downward trend evidences ongoing coarsening; bottom: non-aggregated conditions fluctuate around a flat baseline. Thin lines show raw series, thick lines a light median smoothing; the dashed line marks the late-time mean.

the topological neighborhood. Results are qualitatively robust across seeds and system sizes used here; finite-size and boundary effects are discussed in Section 4.6; a full quantification is left for future work.

### 3.10 Summary of preliminary findings

- Aggregation emerges as a connectivity transition when the kernel is sufficiently connective (*e.g.*, Chebyshev, larger $R$), producing a giant component over finite regions of ($\rho$, $T$) (see figures).
- The aggregated regime coincides with a reduction in $f_{restr}$, consistent with the release of solvent degrees of freedom that drives micellization in this model.
- Low-connectivity kernels (*e.g.*, Manhattan with small $R$) suppress aggregation across the same ($\rho$, $T$) sweep, demonstrating that Euclidean distances are not required once topological neighborhood is specified and $\alpha$ is fixed.

## 4 Discussion

### 4.1 Rationale for the scanned domain

The goal here is not to chart an exhaustive phase boundary, but to test a causal claim: at fixed $\alpha$, aggregation is governed by the topology of the local neighborhood. We therefore adopted a discriminative slicing strategy in parameter space, targeting orthogonal extremes that maximally separate the two regimes: (i) low-connectivity/high-$T$/low-$N_a$ slices (Manhattan $R = 1$) that suppress aggregation, and (ii) high-connectivity/low-$T$/higher-$N_a$ slices (Chebyshev or Manhattan $R = 2$) that promote it. This design demonstrates, with $\alpha = 1$ held fixed, the existence of both aggregated and non-aggregated conditions and attributes the transition to connectivity rather than to distance-weighted interactions.

Coverage is therefore not complete in every quadrant (*e.g.*, Chebyshev at very high $T$, or Manhattan $R = 1$ at very low $T$ and very high $N_a$), because such cases are not required to validate the topological hypothesis and would divert focus from the main point. If desired, a minimal set of "bridging" points (*e.g.*, Manhattan $R = 1$ at $T = 0.5$, $N_a = 400$ and $T = 1$, $N_a = 800$; Chebyshev $R = 2$ at $T = 2$, $N_a = 400$) can be added without altering the qualitative conclusion: neighborhood connectivity is the operative control knob for micellization when $\alpha$ is held constant.

### 4.2 Phenomenology and stance

The central claim supported by our results is that amphiphile self-assembly, in the regime represented by our model, is best

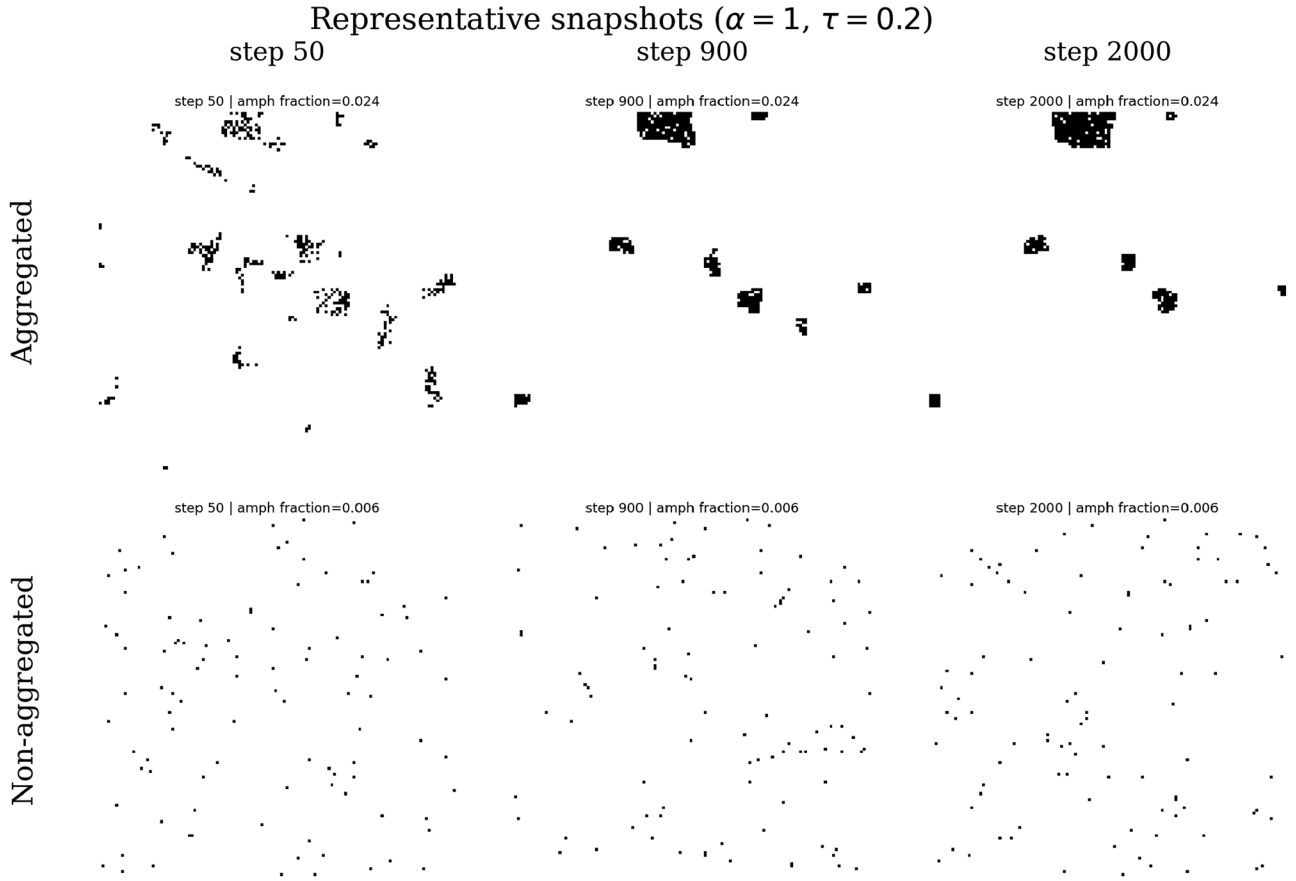


Fig. 6 Real-space snapshots of amphiphile occupancy. Top row: an aggregating run at early, intermediate, and late times shows domain coarsening into a few large clusters. Bottom row: a Aon-aggregating run remains as a dispersed gas of small clusters across the same time points. These snapshots visually confirm the kinetic trends reported for $S_{max}/N_a$, energy $F/(L^2 - N_a)$, and $n_{clusters}$.

viewed as a connectivity-driven and entropy-driven phenomenon. We deliberately fix $\alpha = 1$ so that no scale factor can be (mis)read as encoding metric distances; the only locality is the short-range neighborhood prescribed by the topological kernel (metric, $R$). Within this ensemble the statistical weight $\propto \exp[-F/T]$ with $F = N_{restr}$ makes explicit that aggregation is favored because it reduces the count of restricted water sites, thereby increasing the solvent's accessible microstates at fixed ($\rho$, $T$). In other words, the hydrophobic effect is represented here as a solvent-centered entropic penalty, not as a pairwise enthalpic attraction between amphiphiles.

### 4.3 Connectivity *versus* Euclidean distance

A recurring ambiguity in the literature is to speak of the hydrophobic effect as if it were a fundamental field-like interaction acting across Euclidean space. Our construction emphasizes a different organizing principle: once a short-range neighborhood is specified, the topology of occupancy controls whether a giant component (micellar state) emerges. Euclidean distances enter only implicitly *via* the choice of neighborhood (metric and $R$), which sets how many water sites become restricted around amphiphile domains. The fact that aggregation appears for Chebyshev $R \geq 3$ and is suppressed for Manhattan $R = 1$ at the same ($\rho$, $T$) is precisely a statement about connectivity thresholds, not about explicit metric ranges. This resolves a potential fallacy: our claim is not that geometry is irrelevant but that, once locality is fixed, the controlling variable for assembly is the topology of connections induced by that locality.

### 4.4 Surface adsorption *versus* micellization

Classical narratives often conflate two distinct entropic reorganizations in water-amphiphile systems: (i) interfacial adsorption (*e.g.*, to an air/water surface) and (ii) bulk micellization. In the first case, an external boundary acts as a topological "sink" that immediately reduces restricted-water volume by placing amphiphiles at an interface. In the second, reduction of restricted-water volume is achieved by coalescing amphiphiles into finite aggregates in the bulk. Our model isolates mechanism (ii) without invoking surface saturation pictures; the observed connectivity transition is therefore the bulk (interior) counterpart of the same solvent-entropic imperative that underlies surface activity. This distinction helps explain why a CMC inferred from surface-tension curves need not coincide with the percolation-like threshold for bulk aggregation in our phase maps. Crucially, the onset of surface activity and the bulk

connectivity threshold are boundary-condition dependent and need not coincide: interfacial adsorption reduces restricted-water volume by leveraging an external boundary, whereas bulk micellization achieves the same entropic relief through interior coalescence.

### 4.5 Reconciling the entropic narrative with the model's cost function

Calling the hydrophobic effect a "force" can be semantically misleading if one imagines a field with a metric range comparable to electromagnetism. In our formulation the effective cost $F$ is an entropic term: each restricted water site represents a local loss of configurational freedom. Aggregation lowers $F$ because restriction shells are shared and their union shrinks. While we use $F$ as an "energy" in the Boltzmann weight, its physical origin in this coarse-grained picture is the solvent entropy, not a direct amphiphile-amphiphile potential. This framing is consistent with the observed tight anticorrelation between $S_{max}/N_a$ and the restricted-water fraction.

### 4.6 Scope and limitations

Our minimal model omits several ingredients that are known to shape aggregate morphology and size in real systems: (i) electrostatics and multipolar interactions within the polar headgroups; (ii) tail flexibility, chain length polydispersity, and packing frustration; (iii) solvent compressibility and finite correlation lengths beyond a hard neighborhood cutoff; (iv) off-lattice positional degrees of freedom and curvature energetics in 3D. We set these aside by design to interrogate the essence of self-assembly: the solvent's drive to maximize accessible microstates *via* topological reorganization. Consequently, our phase boundaries should be read as entropic-connectivity thresholds; adding electrostatics and specific enthalpic terms is expected to shift these boundaries and select morphologies (spherical *versus* cylindrical *versus* lamellar) without invalidating the connectivity mechanism itself.

Methodologically, two modeling choices deserve attention. First, the binary restriction rule uses a fixed threshold; while our scans show robust qualitative behavior, a sensitivity analysis of this threshold and its possible temperature dependence is warranted. Second, the present study is two-dimensional on a square lattice; lattice anisotropies and finite-size effects can smear or shift the apparent threshold. We therefore treat the reported phase maps as preliminary guides to the topology of the transition, to be consolidated by (i) size-scaling analyses and (ii) 3D extensions.

Although micelles are three-dimensional entities, the present model does not aim to reproduce their geometry but to formalize the topological origin of their formation. The connectivity transition described here is independent of dimensionality, since it arises from the reorganization of adjacency relations rather than metric distances. A 3D lattice implementation would not alter the nature of the transition, as the hydrophobic constraint operates identically in any dimension. Therefore, the use of a 2D lattice is a conceptual choice intended to isolate the phenomenon of entropy-driven connectivity, not to project a higher-dimensional morphology.

It is also worth clarifying the physical interpretation of the $\approx$1 nm cutoff discussed by Chandler *et al.*[9] In that framework, this limit does not correspond to a geometrical threshold but to a thermodynamic one: below this scale, electrostatic, dipolar, and ionic interactions act as energy-fluctuating agents that preserve the isotropy of the solvent's degrees of freedom. The system remains dominated by local noise, and the second law does not manifest as a directional entropic drive. Only when the excluded volume reaches a scale at which these fluctuations become correlated—beyond roughly one nanometer—does the solvent lose sufficient configurational freedom to generate an effective entropic gradient. In that regime, the second law ceases to compete with noise and instead organizes it.

The present model isolates precisely that moment of transition, abstracting the hydrophobic effect as the topological reorganization of solvent freedom under exclusion. Electrostatic and ionic interactions are therefore not dismissed but treated as subsequent modulators that define micellar size, morphology, and stability once the entropic connectivity transition has occurred.

It should be emphasized that this deliberate simplification is not a weakness but a conceptual decision. The omission of parameters such as ionic strength, chain length, or hydrophilic-hydrophobic ratio is meant to isolate the entropic and topological essence of micellization. The internal properties of the aggregate are expected to be dominated by van der Waals attraction arising from transient dipolar moments, whereas the interfacial properties depend on electrostatic, ionic, or protonation effects. These interactions determine morphology and colloidal stability, allowing the system to explore a broad region of its phase space, yet always under the aggregation constraint described in the present model.

Parametrizing the lattice nodes to include such effects could indeed provide finer structural resolution, but would shift the model toward a level already accessible by molecular dynamics. The present approach, conversely, aims not to go beyond molecular dynamics but beneath it—to its non-quantum foundation—showing that self-assembly is, in its origin, a purely thermodynamic phenomenon.

### 4.7 Falsifiable predictions and how to test them

The connectivity account leads to concrete, testable predictions:

1. Connectivity threshold. At fixed ($\rho$, $T$), increasing kernel connectivity (*e.g.*, switching Manhattan $\rightarrow$ Chebyshev or increasing $R$) should *lower* the aggregation threshold, observable as a leftward/downward shift of the boundary in ($\rho$, $T$) phase maps.

2. Entropic signature. Across the boundary, the restricted-water fraction should drop sharply and anticorrelate with $S_{max}/N_a$; the second moment of the cluster-size distribution (excluding $S_{max}$) should peak near the boundary.

3. Scale-factor invariance. With $\alpha$ fixed, re-scaling any putative "energy" unit should not affect the location of the boundary; any apparent dependence on $\alpha$ would indicate hidden metric content, which our construction avoids.

4. Decoupling from surface activity. In geometries that permit an external interface (not modeled here), the onset of

 

surface activity (slope change in surface tension) need not coincide with the bulk connectivity threshold; they can be tuned independently by changing boundary conditions *versus* neighborhood connectivity.

These statements can be scrutinized by (i) explicit-water simulations where one extracts solvent mobility maps around amphiphile clusters to proxy $N_{\mathrm{restr}}$, and (ii) experiments that manipulate effective connectivity (*e.g.*, *via* cosolvents or salt that modulate water structuring length scales) without dramatically altering amphiphile chemistry.

Although exemplified here through micellization, the model's formalism is general to any system where solvent exclusion imposes a topological constraint—including peptide amphiphiles, block copolymers, or other soft-matter assemblies governed by entropy-driven connectivity.

### 4.8 Outlook

Two extensions are especially natural. First, a 3D lattice with anisotropic kernels could connect more directly to spherical or cylindrical micelles and permit explicit curvature costs while preserving the solvent-entropic driver. Second, coupling the restriction operator to a continuous solvent order parameter would allow the neighborhood kernel to emerge from a finite correlation length rather than be imposed, bridging the gap between topological connectivity and continuum free-energy approaches. In either case, the organizing principle remains: micellization as a topology-controlled route to liberate solvent degrees of freedom under the constraints of the second law.

In summary, our findings support a coherent narrative in which amphiphile self-assembly is not anchored in long-range metric interactions but in the local connectivity of amphiphile neighborhoods that govern solvent restriction and release. This perspective clarifies the relation between surface adsorption and bulk micellization, explains the observed phase structure under variations of ($\rho$, $T$, metric, $R$), and provides a set of falsifiable predictions for simulation and experiment.

## Conclusions

We introduced a topological Ising-like model of amphiphile self-assembly in which the hydrophobic effect is encoded as a solvent-centered restriction count, $F = N_{\mathrm{restr}}$, evaluated over a short-range neighborhood specified by a metric and radius (metric, $R$). By fixing $\alpha = 1$ throughout, we removed any spurious metric content from the cost function and isolated the role of connectivity. Scans over amphiphile density $\rho$, temperature $T$, and (metric, $R$) showed that aggregation emerges as a connectivity transition: sufficiently connective kernels (*e.g.*, Chebyshev with larger $R$) yield a giant component over finite regions of ($\rho$, $T$), while low-connectivity settings suppress aggregation. The aggregated regime is consistently accompanied by a reduction of the restricted-water fraction, aligning the lattice phenomenology with the entropic picture in which micelles form to liberate solvent degrees of freedom.

Methodologically, our analysis establishes $S_{\max}/N_{\mathrm{a}}$ as a robust order parameter for the transition and highlights the tight anticorrelation between $S_{\max}/N_{\mathrm{a}}$ and $f_{\mathrm{restr}}$ as an operative entropic signature. Conceptually, the results clarify that Euclidean distances matter only through the chosen local neighborhood; once locality is fixed, the topology of occupancy governs assembly. Limitations—electrostatics, chain architecture, and off-lattice effects—are acknowledged and constitute levers expected to shift boundaries and select morphologies without undermining the connectivity mechanism.

Looking ahead, 3D extensions with anisotropic kernels and couplings to continuum solvent fields offer a principled path to bridge connectivity-controlled assembly with curvature energetics and experimentally observed morphologies, providing falsifiable predictions for simulations and experiments.

## Author contributions

The author confirms being the sole contributor to the conception, design, investigation, data analysis, writing, and revision of this manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Data availability

No primary research results, software or code have been included and no new data were generated or analysed as part of this review.

## Acknowledgements

Financial support by Xunta de Galicia through grant IN606B-2023/006 is gratefully acknowledged. Facilities provided by the Galician Supercomputing Centre (CESGA) are also acknowledged.

## Notes and references

1 V. Dominguez-Arca, J. Sabin, P. Taboada, L. Garcia-Rio and G. Prieto, *J. Mol. Liq.*, 2020, **308**, 113100.
2 Z. Cernochova, A. Bogomolova, O. V. Borisova, S. K. Filippov, P. Cernoch, L. Billon, O. V. Borisov and P. Stepanek, *Soft Matter*, 2016, **12**, 6788–6798.
3 R. Bordes and K. Holmberg, *Adv. Colloid Interface Sci.*, 2015, **222**, 79–91.
4 R. Alvares, S. Gupta, P. M. Macdonald and R. S. Prosser, *J. Phys. Chem. C*, 2014, **118**, 5698–5706.
5 Y. Li, K. Holmberg and R. Bordes, *J. Colloid Interface Sci.*, 2013, **411**, 47–52.
6 H. Comas-Rojas, C. Enriquez-Victorero, S. J. Roser, K. J. Edler and A. Perez-Gramatges, *Soft Matter*, 2013, **9**, 4003–4014.
7 I. W. Hamley, A. Dehsorkhi, P. Jauregi, J. Seitsonen, J. Ruokolainen, F. Coutte, G. Chataigne and P. Jacques, *Soft Matter*, 2013, **9**, 9572–9578.
8 Y. Moroi, *Micelles: Theoretical and Applied Aspects*, Springer, New York, NY, 1992.

9 D. Chandler, *Nature*, 2005, **437**, 640–647.
10 L. R. Pratt, Y. S. M. Michaud-Agrawal and S. B. Rempe, *J. Phys. Chem. B*, 2016, **120**, 6455–6460.
11 R. L. Baldwin, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 13052–13056.
12 S. J. Marrink, A. H. de Vries and A. E. Mark, *J. Phys. Chem. B*, 2004, **108**, 750–760.
13 A. Jusufi and A. Z. Panagiotopoulos, *Langmuir*, 2015, **31**, 3283–3292.
14 R. Nagarajan and E. Ruckenstein, *Langmuir*, 1991, **7**, 2934–2969.
15 A. G. Schlijper, L. A. M. Rupert, N. M. van Os and B. Smit, *J. Phys. Chem.*, 1990, **94**, 6933–6935.
16 S. Glotzer and A. Coniglio, *Comput. Mater. Sci.*, 1995, **4**, 325–333.
17 P. Lai and Y. Goldschmidt, *J. Phys. A: Math. Gen.*, 1989, **22**, 399–411.
18 R. Guo, X. Li, R. Wang, S. Chen, Y. Wu and Z. Li, *Chin. Phys. C*, 2025, **49**, 054103.
19 S. Lin, L. Zhao, S. Liu, Y. Wang and G. Fu, *Soft Matter*, 2023, **19**, 502–511.
20 L. V. Dumancas, D. E. Simpson and D. T. Jacobs, *J. Chem. Phys.*, 2015, **142**, 174902.
21 R. Behjatmanesh-Ardakani, M. A. Karimi and F. G.-M. Abadi, *J. Mol. Liq.*, 2009, **149**, 1–8.
22 J. D. Schmit, K. Ghosh and K. Dill, *Biophys. J.*, 2011, **100**, 450–458.
23 H. F. Xu and K. A. Dill, *J. Phys. Chem. B*, 2005, **109**, 23611–23617.
24 K. A. Dill, T. M. Truskett, V. Vlachy and B. Hribar-Lee, *Annu. Rev. Biophys. Biomol. Struct.*, 2005, **34**, 173.
25 in *Mathematical Morphology and Its Applications to Image Processing*, ed. J. Serra and P. Soille, Springer, Dordrecht, 1994.
26 K. Lum, D. Chandler and J. D. Weeks, *J. Phys. Chem. B*, 1999, **103**, 4570–4577.
27 K. A. Dill, *Biochemistry*, 1990, **29**, 7133–7155.
28 A. J. Bray, *Adv. Phys.*, 1994, **43**, 357–459.