Soft Matter

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/d5sm01246e

A solvent-topology perspective on hydrophobic aggregation

Vicente Domínguez-Arca [ab]

Hydrophobic aggregation is often described in terms of effective attractive forces between apolar units. In aqueous and soft-matter environments, however, aggregation necessarily involves a collective reorganization of the solvent, whose connectivity and confinement properties may play a central role. Here we investigate a deliberately minimal lattice model in which hydrophobic (apolar) segments interact with the solvent only through a local restriction rule, without introducing explicit solute–solute attractions. Each lattice site is occupied either by solvent or by an apolar segment at a fixed composition. Solvent sites are classified as restricted when the local density of nearby apolar segments exceeds a prescribed threshold, and configurations are sampled with a Boltzmann-like weight that penalizes the total number of restricted-solvent sites. Within this framework, aggregation can only arise through the reorganization of solvent restriction fields. By decomposing the solvent into free and restricted subsets, we analyze their connectivity using standard percolation diagnostics alongside conventional measures of solute clustering. Baseline simulations reveal that while amphiphile aggregation evolves smoothly with concentration, the restricted-solvent subset can undergo a sharp, kernel-dependent connectivity crossover: for sufficiently isotropic interaction neighborhoods, restricted water becomes a system – spanning only over an intermediate range of solute fractions. Free solvent, by contrast, remains spanning across the explored parameter range. Notably, these solvent-topology signatures are significantly sharper and more structured than the corresponding solute-ordering metrics. A systematic extension to larger system sizes refines this picture. The onset of restricted-solvent spanning remains localized and kernel selective, while high-density behavior becomes smoother and more fluctuation dominated, consistent with finite-size effects. Importantly, the amount of restricted solvent grows monotonically and robustly with system size, whereas its global connectivity reorganizes nontrivially. Taken together, these results support a restrained interpretation of hydrophobic aggregation as a solvent-driven topological crossover rather than a sharp thermodynamic phase transition, highlighting solvent topology as a sensitive diagnostic even when solute ordering remains gradual.

Received 16th December 2025,
Accepted 25th February 2026

DOI: 10.1039/d5sm01246e

rsc.li/soft-matter-journal

1 Introduction

Hydrophobic aggregation is often presented as if it were driven by an effective attraction between apolar moieties.[1–5] However, a long tradition of experimental and theoretical work has emphasized that, in water-rich soft-matter environments, aggregation cannot be understood independently of the collective response of the solvent: the dominant thermodynamic bookkeeping frequently reflects what the solvent must reorganize to accommodate hydrophobic material rather than any direct solute–solute energetic term.[6–9] In this view, solvent-mediated effects are intrinsically nonlocal and difficult to reduce to a single parameter: connectivity, confinement, and the spatial arrangement of hydration-like domains can all matter simultaneously, particularly near interfaces and in the presence of extended hydrophobic patches.[10–12]

In previous work, we introduced a deliberately minimal lattice framework aimed at isolating the hydrophobic effect from explicit chemical detail,[13] treating it instead as an entropy-dominated, connectivity-driven phenomenon mediated by solvent constraints. In that setting, amphiphile aggregation emerged without invoking direct solute–solute attractions, but rather as a consequence of how local solvent restrictions propagate through the system. This emphasis on collective solvent response is consistent with modern molecular theories in which solute attractive forces can modulate hydrophobic interactions while not exhausting their explanation, and where hydrophobic

[a] Biosystems and Bioprocess Engineering (Bio2Eng) Group, Institute of Marine Research of Spanish Research Council, IIM-CSIC, C/Eduardo Cabello 6, Vigo, Pontevedra, 36208, 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, Nordrhein-Westfalen, 33615, Germany

association retains characteristic temperature dependence even when attractive contributions are treated explicitly.[14,15] At the same time, analytical and integral-equation approaches have stressed that hydrophobicity can be captured without presupposing a simple ''hydrogen-bond ordering'' narrative, highlighting the role of packing, fluctuations, and water's anomalous liquid-state response.[16–18]

The present study builds directly on that conceptual foundation but shifts the emphasis from solute connectivity to solvent topology. Rather than asking when amphiphiles aggregate, we ask how the solvent reorganizes as aggregation proceeds, and whether the solvent itself undergoes qualitative changes in global connectivity. This change of perspective is motivated by the fact that solute ordering can evolve smoothly with concentration, whereas solvent connectivity is a genuinely global property that can reorganize abruptly even without a sharp solute-ordering transition. Such solvent-mediated crossovers are naturally framed in terms of interfacial fluctuations and collective rearrangements of the liquid, which have been argued to control hydrophobic assembly across length scales.[8–10,19]

To address this question, we consider a closely related minimal lattice model in which hydrophobic (apolar) segments interact with the solvent only through a local restriction rule. Each lattice site is occupied either by an apolar solute segment or by water, at fixed solute area fraction r and effective temperature T. A local neighborhood operator counts the number of solute segments within a prescribed kernel, and water sites are labeled as restricted when this count exceeds a threshold. The global cost of a configuration is defined as the total number of restricted-water sites, and Monte Carlo sampling favors configurations that reduce this cost while preserving overall composition. As in the previous work, no explicit pairwise attraction between solute segments is introduced; any aggregation or ordering must therefore arise indirectly, through the reorganization of the solvent restriction field. This construction is not intended as a chemically faithful model of amphiphiles (polar head groups and specific interactions are deliberately omitted), but as a controlled setting in which solvent-mediated organization can be studied with minimal assumptions.[7,20]

Within this construction, the central question becomes explicitly topological. As the solute fraction r increases, does the restricted-water subset remain a collection of disconnected patches, or can it form a system-spanning connected network? To probe this issue, we decompose the solvent into free water and restricted water and monitor their percolation (spanning) probabilities under standard nearest-neighbor connectivity. In parallel, we quantify solute aggregation through the largest-cluster fraction $f = S_{\max}/N_a$ and complementary statistics derived from the distribution of f. This dual monitoring is designed to separate (i) gradual changes in solute connectivity from (ii) potentially sharp, collective rearrangements in the solvent network, in the spirit of solvent coarse-graining strategies that explicitly treat water as a fluctuating field capable of controlling hydrophobic collapse pathways.[11]

Because connectivity is a global property, its characterization in lattice models necessarily raises questions of system size and finite-size effects. For this reason, the analysis is organized in two layers. First, we present results for a reference lattice size, which allows a clear identification of qualitative regimes and connectivity crossovers. Building on this baseline, we then introduce an explicit system-size extension, increasing the lattice linear dimension and refining the sampling in r, in order to assess the robustness of the observed phenomenology and to clarify which features persist as genuinely collective effects rather than finite-size artifacts. The goal is therefore not to map parameters directly to experimental concentrations, but to isolate a single mechanism—the emergence, modulation, and possible loss of global connectivity in a solvent subset defined by local restriction—and to test whether solvent topology can exhibit sharp, interpretable signatures even when solute ordering evolves more gradually.[8–10]

2 Model and methods

2.1 Lattice model: degrees of freedom and ensembles

We consider a two-dimensional square lattice of linear size L with $N = L^2$ sites, and with periodic boundary conditions in both directions. Each site $i \in \{1,\ldots,N\}$ is occupied either by a hydrophobic (apolar) solute segment (hereafter ''amphiphile segment'') or by solvent (water). The microscopic configuration is encoded by a binary field:

$$A_i \in \{0, 1\}, \tag{1}$$

where $A_i = 1$ denotes an amphiphile segment and $A_i = 0$ denotes solvent. The total number of amphiphile segments is fixed (canonical composition),

$$N_a \equiv \sum_{i=1}^{N} A_i, \quad r \equiv \frac{N_a}{N}, \tag{2}$$

where r is the solute (amphiphile-segment) area fraction. All simulations are performed at fixed (L, r, T).

Unless otherwise stated, results are first presented for a reference system size $L = 48$. To assess the robustness of the observed phenomenology with respect to system size and to better resolve connectivity trends, an explicit extension to a larger lattice ($L = 64$) is also considered, using the same model, parameters, and analysis pipeline.

2.2 Local neighborhood kernel and solvent restriction field

The key modeling ingredient is a local neighborhood operator that counts, for each site i, how many amphiphile segments lie within a prescribed interaction neighborhood. This neighborhood is specified by a binary kernel $K(\Delta x, \Delta y) \in \{0, 1\}$ with finite range R, defined on offsets $(\Delta x, \Delta y)$ such that

$$K(\Delta x, \Delta y) = \begin{cases} 1, & (\Delta x, \Delta y) \in \mathcal{N}_K, \\ 0, & \text{otherwise}, \end{cases} \qquad \Delta x, \Delta y \in \{-R, \ldots, R\}, \tag{3}$$

where $\mathcal{N}_K$ denotes the metric-dependent neighbourhood defined by eqn (4)–(7). We consider four kernels (metrics) commonly used in the literature:

• Chebyshev neighborhood (square of radius R):

$$K(\Delta x, \Delta y) = \mathbb{1}[\max(|\Delta x|, |\Delta y|) \le R]. \quad (4)$$

• Euclidean neighborhood (disk of radius R):

$$K(\Delta x, \Delta y) = \mathbb{1}\left[\Delta x^2 + \Delta y^2 \le R^2\right]. \quad (5)$$

• Manhattan neighborhood (diamond of radius R):

$$K(\Delta x, \Delta y) = \mathbb{1}[|\Delta x| + |\Delta y| \le R]. \quad (6)$$

• Cross neighborhood (axial plus-shape of radius R):

$$K(\Delta x, \Delta y) = \mathbb{1}[(\Delta x = 0 \text{ or } \Delta y = 0) \text{ and } \max(|\Delta x|, |\Delta y|) \le R]. \quad (7)$$

The kernel support (neighborhood size) is

$$S_K \equiv \sum_{\Delta x, \Delta y} K(\Delta x, \Delta y). \quad (8)$$

For any configuration A, we define the local amphiphile count field C_i as a discrete convolution:

$$C_i \equiv \sum_{j=1}^{N} K_{i-j} A_j, \quad (9)$$

where K_{i-j} denotes the kernel evaluated at the lattice offset from j to i (with the convention that $K_0 = 1$ if the neighborhood includes the center; in our implementation it does). Operationally, C_i counts the number of amphiphile segments in the neighborhood of i.

A solvent site i is defined as restricted if it is solvent ($A_i = 0$) and its neighborhood contains more than a threshold number θ of amphiphile segments:

$$R_i \equiv \mathbb{1}[(A_i = 0) \wedge (C_i > \theta)]. \quad (10)$$

Here $R_i \in \{0, 1\}$ is the restricted-solvent indicator field and θ is a model parameter controlling how easily solvent becomes "constrained" by nearby hydrophobic segments.

2.3 Energy (cost) function and statistical weight

We define the scalar cost (energy-like) function as the total number of restricted-solvent sites,

$$F(A) \equiv \sum_{i=1}^{N} R_i = \sum_{i=1}^{N} = \mathbb{1}[(A_i = 0) \wedge (C_i > \theta)]. \quad (11)$$

Configurations are sampled according to a Boltzmann-like weight at effective temperature T,

$$\pi(A) \propto \exp\left[-\frac{F(A)}{T}\right], \quad (12)$$

subject to the constraint $\sum_i A_i = N_a$.

This choice encodes the idea that solvent sites in highly constrained local environments contribute an effective free-energy penalty. Accordingly, Monte Carlo sampling favors configurations that minimize $F(A)$ while preserving the solute fraction.

2.4 Monte Carlo dynamics at fixed composition

We employ the Metropolis Monte Carlo algorithm that preserves N_a *via* random swaps between amphiphile and solvent sites. Each Monte Carlo *sweep* comprises n_{prop} swap proposals, where

$$n_{prop} \equiv \min(N_a, N - N_a), \quad (13)$$

so that at most the smaller phase is extensively updated per sweep.

2.5 Simulation protocol

For each kernel type, system size L, temperature T, threshold θ, and composition r, we perform independent simulations over n_{seed} random seeds. Unless otherwise stated, simulations at L = 48 serve as the reference dataset, while simulations at L = 64 constitute a systematic size extension carried out with identical parameters and analysis procedures.

For each run

1. Initialization: a random configuration with exactly N_a = round(rN) amphiphile sites placed uniformly without replacement.

2. Burn-in: N_{burn} Monte Carlo sweeps discarded to reduce dependence on the initial state.

3. Production: N_{prod} sweeps during which observables are sampled every s_{sample} sweeps, yielding $n_{samp} = N_{prod}/s_{sample}$ measurements per run.

Reported means and uncertainties correspond to averages over production samples and independent seeds. Figures corresponding to the extended system size are labeled with the same identifiers as their L = 48 counterparts, augmented by the suffix "64".

2.6 Outputs

For each (kernel, L, r, seed) we store

• $\langle F \rangle$, Std(F): mean and standard deviation of restricted solvent count.

• $\langle f \rangle$, Std(f) and $\chi_f = N_a \mathrm{Var}(f)$: amphiphile connectivity statistics.

• $P_{span}^{(free)}$ and $P_{span}^{(restr)}$: spanning probabilities of free and restricted solvent.

• (P_M, P_U, P_C, ΔB): barrier-like quantities derived from the f distribution.

All of the code was executed in deterministic settings with fixed random seeds per run, and CSV files containing raw per-run results were generated for subsequent aggregation and plotting.

3 Results

3.1 Overview: two coupled observables (solute ordering vs solvent topology)

Across all runs we monitor (i) an amphiphile ordering proxy, the largest-cluster fraction

$$f \equiv \frac{S_{max}}{N_a}, \quad (14)$$

where S_{max} is the size of the largest amphiphile cluster and N_a is the total number of amphiphile segments, and (ii) a solvent-topology proxy based on percolation of two disjoint solvent subsets:free water

$$W_{free} \equiv \{(i,j):A_{ij} = 0 \wedge \neg restr_{ij}\}, \quad (15)$$

and restricted water

$$W_{restr} \equiv \{(i,j):restr_{ij}\}, \quad restr_{ij} \equiv (A_{ij} = 0) \wedge (C_{ij} > \theta), \quad (16)$$

where $A_{ij} \in \{0, 1\}$ encodes solvent/solute occupation and C_{ij} is the local kernel count around site (i,j). Percolation is quantified as the probability that a given mask spans the box (horizontal or vertical) under 4-neighbour connectivity,

$$P_{span}(X) \equiv \langle \mathbb{I}[X \text{ spans}] \rangle, \quad X \in \{W_{free}, W_{restr}\}, \quad (17)$$

where $\langle \cdot \rangle$ denotes averaging over the production samples (and then over seeds).

The results are presented in two layers. First, we report the baseline dataset at $L = 48$, which allows a clear qualitative identification of the regimes where solute connectivity evolves and where the solvent topology exhibits sharp changes. Second, we introduce a systematic size extension at $L = 64$ with a denser sampling in r (same code, parameters, and analysis), in order to assess which features are robust and which are sensitive to finite-size effects in global-connectivity observables.

3.2 Percolation of free water: robustness of a spanning bulk solvent

Fig. 1 shows $P_{span}(W_{free})$ as a function of amphiphile density r for the $L = 48$ runs. For all metrics and radii considered, free water remains almost always spanning across the explored range: $P_{span}(W_{free}) \approx 1$. Only mild reductions appear at the largest r in some kernels (visible as small deviations from unity), which are consistent with finite-size fragmentation effects rather than a genuine loss of a system-spanning solvent backbone. Therefore, within this parameter range, the model does not predict a transition where the bulk solvent disconnects.

This observation is important because it separates two possibilities that are often conflated: (i) "less available water" as a geometrical reduction of free solvent volume, *versus* (ii) "loss of connectivity" of the free-water network. Here we mostly observe the former (a gradual reduction of free-water sites), not the latter (a true percolation breakdown).

3.2.1 System-size extension ($L = 64$). The same conclusion persists at $L = 64$. Fig. 6 shows $P_{span}(W_{free})$ for the size-extended runs (here shown for Chebyshev_R2 and Euclidean_R2, for which the full $L = 64$ sweep was performed). Across the densified r grid, free water continues to span with probability essentially unity throughout most of the range. Where deviations from 1 appear, they occur only at the highest densities and are accompanied by large uncertainty bars, consistent with the intuitive picture that, at sufficiently high solute loading, the remaining free-solvent set becomes narrow and susceptible to occasional fragmentation events. Crucially, even at $L = 64$ these deviations do not organize into a sharp percolation threshold: the free-solvent network remains the dominant system-spanning backbone in the model across the studied window.

3.3 Percolation of restricted water: kernel-dependent spanning and its size dependence

A qualitatively different behaviour emerges for restricted water. Fig. 2 shows $P_{span}(W_{restr})$ at $L = 48$. For Chebyshev- and Euclidean-type kernels with $R = 2$, the restricted-water subset exhibits a non-monotonic response:

- At low r, $P_{span}(W_{restr}) \simeq 0$.
- Over an intermediate density window, $P_{span}(W_{restr})$ becomes appreciable and the restricted-water domain forms a system-spanning network.

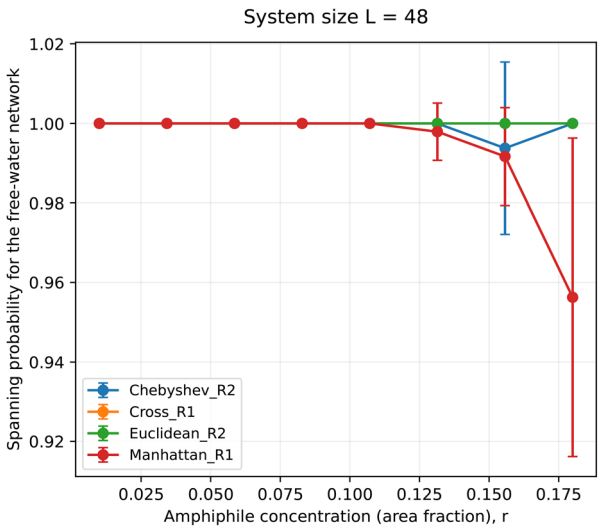

Fig. 1 Spanning probability of free water, $P_{span}(W_{free})$, as a function of amphiphile density r for the different kernels ($L = 48$ baseline).

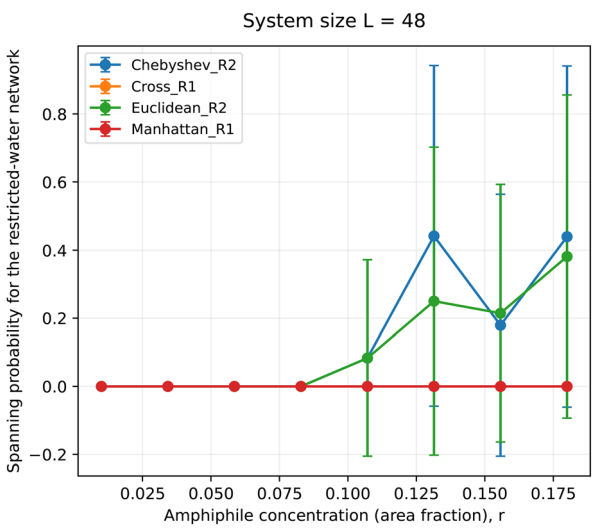

Fig. 2 Spanning probability of restricted water, $P_{span}(W_{restr})$, as a function of r for the different kernels ($L = 48$ baseline), showing a kernel-dependent transient spanning window for extended/isotropic neighbourhoods.

• At higher r, the spanning probability decreases again, indicating loss of global connectivity despite the continued presence of restricted sites.

This behaviour defines a finite-density spanning window that is strongly kernel dependent. In contrast, for Manhattan_R1 and Cross_R1 the restricted-water spanning probability remains essentially zero throughout (Fig. 2), indicating that system-scale connectivity of restricted water is not generic but is controlled by neighbourhood geometry.

3.3.1 System-size extension (L = 64). When the system size is increased to L = 64 and the r grid is densified, the same qualitative mechanism persists (Fig. 7). Restricted water remains non-spanning at low r, while spanning events emerge over an intermediate density range. However, compared to L = 48, the onset becomes more gradual and is accompanied by larger fluctuations, consistent with the binary nature of the spanning observable.

In the explored range, the L = 64 curves do not show as clean a return to zero at high r as in the smaller system. Instead, spanning can persist intermittently at larger r values. Overall, the size extension supports a topology-driven crossover whose detailed high-r behaviour is modulated by finite-size effects.

3.4 Solvent restriction level: growth of restricted-water population with r

Fig. 3 reports the average restriction measure $\langle F \rangle$ (the number of restricted-water sites, up to the sign convention used in the plot). Independently of the percolation results, the magnitude of restriction increases strongly with r for all kernels: more amphiphile content implies more solvent sites satisfying $C_{ij} > \theta$. This indicates that the amount of restricted water grows monotonically (or nearly so), even though its connectivity does not.

Taken together with Section 3.3, this establishes a crucial separation:

"More restricted water" does not imply "restricted water percolates".

The percolation peak (or onset/crossover region) is therefore a topological effect (connectivity/geometry), not merely a concentration effect.

3.5 Amphiphile ordering: evolution of the largest-cluster fraction

Fig. 4 shows the mean largest-cluster fraction $\langle f \rangle$ *versus* r at L = 48. Across kernels, $\langle f \rangle$ tends to increase with r, consistent with a progressive dominance of larger aggregates as amphiphile density rises. The precise baseline and slope depend on the interaction metric: kernels with more isotropic/larger neighbourhood support (*e.g.*, Chebyshev/Euclidean at larger R) favour higher $\langle f \rangle$ already at comparatively small r, whereas more constrained neighbourhoods (*e.g.*, Manhattan_R1/Cross_R1) typically yield lower $\langle f \rangle$.

Importantly, the solvent-topology signal (Section 3.3) is not a trivial mirror of $\langle f \rangle$: restricted water percolation appears as a localized crossover (or window) in r, while $\langle f \rangle$ varies more smoothly. This supports the interpretation that the strong structure in $P_{\mathrm{span}}(W_{\mathrm{restr}})$ is a property of solvent-network rearrangement, not simply "the moment when amphiphiles aggregate."

3.5.1 System-size extension (L = 64). Fig. 8 reports $\langle f \rangle(r)$ for L = 64 (Chebyshev_R2 and Euclidean_R2). The qualitative trend is unchanged: $\langle f \rangle$ increases with r, indicating increasingly dominant amphiphile clusters. What changes is primarily the resolution with which this increase is observed. With the denser r grid, one can see that the growth of $\langle f \rangle$ is not necessarily a single sharp jump but can proceed through gradual increments, with kernel-to-kernel differences emerging in the slope and in the size of fluctuations. This reinforces the key comparative point of the paper: the amphiphile ordering proxy evolves in a comparatively smooth manner as a function of r, while the solvent topology (especially restricted-water spanning) can show sharper, more structured behavior.

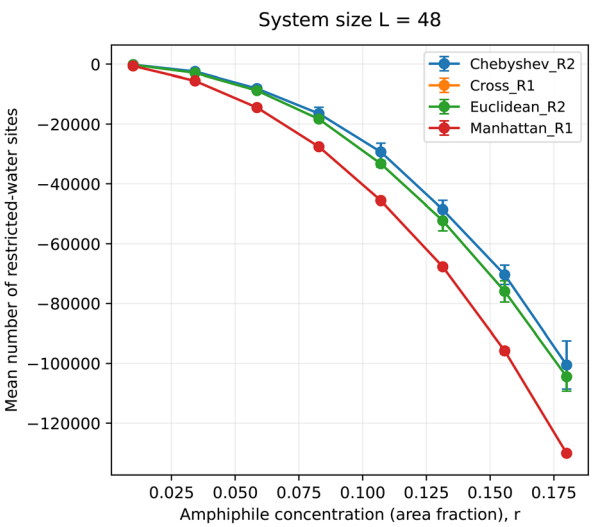

Fig. 3 Average restriction measure $\langle F \rangle$ *versus* r (plot sign convention as used in the figure), showing a strong increase in the magnitude of restricted-water population with amphiphile density (L = 48 baseline).

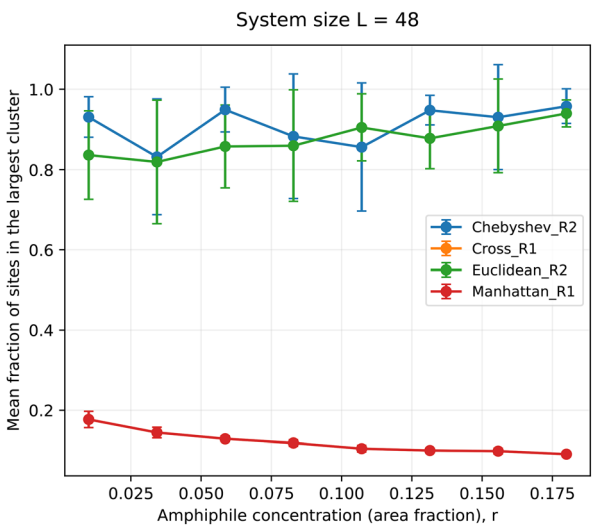

Fig. 4 Mean largest-cluster fraction $\langle f \rangle$ *versus* amphiphile density r for the available kernels/parameters (L = 48 baseline), illustrating progressive aggregation and its dependence on neighbourhood geometry.

To complement the mean trend, Fig. 9 shows the susceptibility-like proxy:

$$\chi_f \equiv N_a \mathrm{Var}(f), \tag{18}$$

for the $L = 64$ dataset. The $\chi_f(r)$ curves remain relatively noisy, with localized peaks whose precise positions and amplitudes depend on the kernel. This is consistent with finite-size connectivity crossovers where fluctuations in the largest-cluster fraction are enhanced over a limited density interval. Notably, these peaks do not generically coincide in a one-to-one way with the solvent spanning changes, underscoring that solvent topology is not merely a re-expression of solute clustering statistics.

3.6 Barrier-like statistic from the order-parameter neck probability

To further probe whether the ordering dynamics suggests metastability, we evaluate the barrier-like statistic ΔB constructed from the probabilities of being in "unordered" (U), "micellized" (M), and "neck" (C) regions in f:

$$P_U = \langle \mathbb{I}[f < t_{ord}] \rangle, \quad P_M = \langle \mathbb{I}[f \geq t_{ord}] \rangle, \quad P_C = \langle \mathbb{I}[|f - t_{ord}| \leq \delta] \rangle, \tag{19}$$

$$\Delta B \equiv -T \ln\left(\frac{P_C}{\sqrt{P_U P_M}}\right). \tag{20}$$

Fig. 5 shows $\Delta B(r)$ for the available kernel(s) at $L = 48$. In the displayed case, ΔB remains approximately constant across r, suggesting that within the explored range and under the chosen thresholds (t_{ord}, δ), there is no sharp barrier growth indicative of a strong first-order-like coexistence in the order parameter. This is consistent with the smoother evolution of $\langle f \rangle$ (Section 3.5).

3.6.1 System-size extension ($L = 64$). Fig. 10 reports the corresponding $\Delta B(r)$ curves for the $L = 64$ dataset. As in the baseline, ΔB does not exhibit a systematic growth with r into a large-barrier regime. Instead, the curves remain relatively flat with scatter, consistent with the interpretation that the ordering proxy f evolves through a finite-size crossover rather than through a strongly metastable, first-order-like jump. This strengthens the separation emphasized throughout the Results: the most prominent "transition-like" signature in these simulations is associated with solvent topology (restricted-water connectivity changes), whereas the solute ordering statistics do not display a comparably sharp barrier structure under the present model definition.

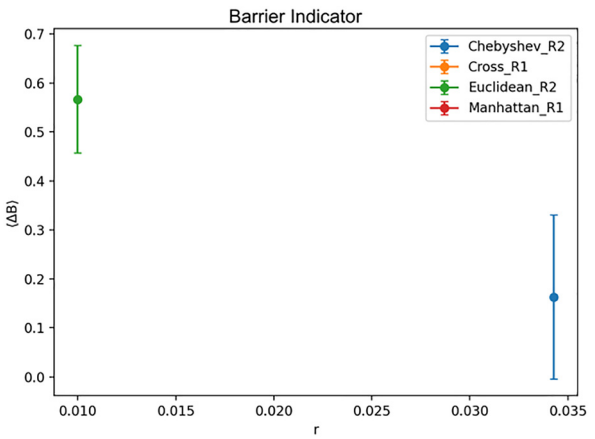

Fig. 5 Barrier-like statistic ΔB computed from the neck probability around a reference threshold in f, shown *versus* r for the available kernel(s) ($L = 48$ baseline).

3.7 Synthesis: a solvent-driven interpretation of hydrophobic aggregation and its scale dependence

Combining Fig. 1–5 (baseline, $L = 48$) and their size-extended counterparts Fig. 6–10, the simulations support a coherent picture with an explicit scale-dependent refinement:

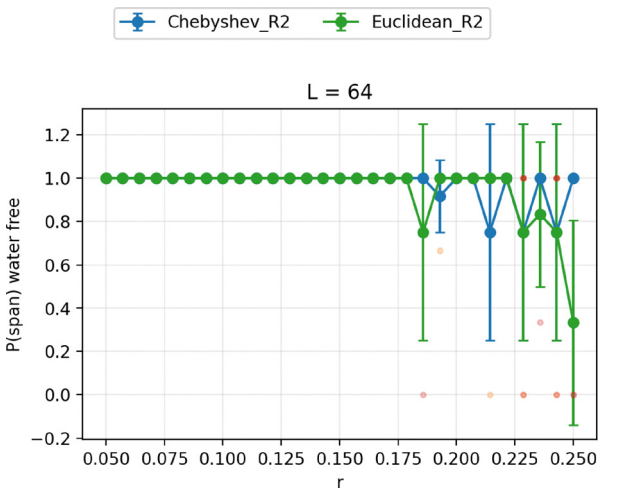

Fig. 6 Spanning probability of free water, $P_{span}(W_{free})$, as a function of amphiphile density r for the size-extended system ($L = 64$), shown for the kernels evaluated in the extension (*e.g.*, Chebyshev_R2 and Euclidean_R2).

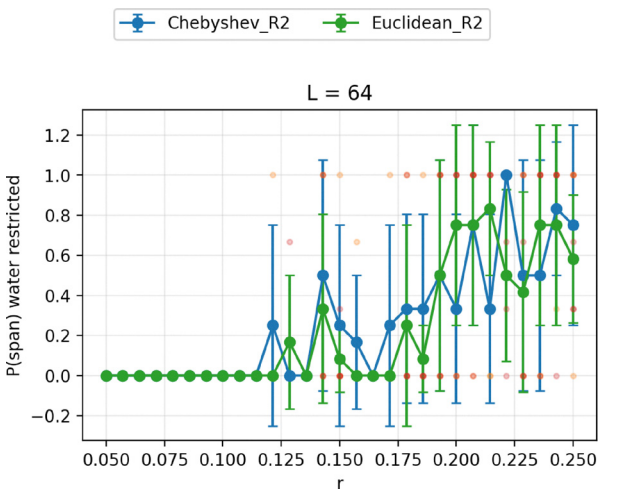

Fig. 7 Spanning probability of restricted water, $P_{span}(W_{restr})$, for the size-extended system ($L = 64$). Compared to the $L = 48$ baseline, the onset of spanning is smoother and spanning can persist intermittently at higher r within the explored range, highlighting finite-size effects in global-connectivity observables.

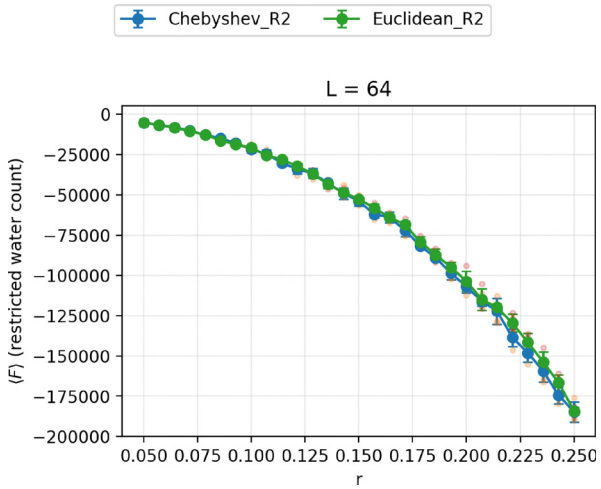

Fig. 8 Mean largest-cluster fraction $\langle f \rangle$ *versus* r for the size-extended system (L = 64), illustrating that solute connectivity increases smoothly while solvent-topology observables can show sharper structured changes.

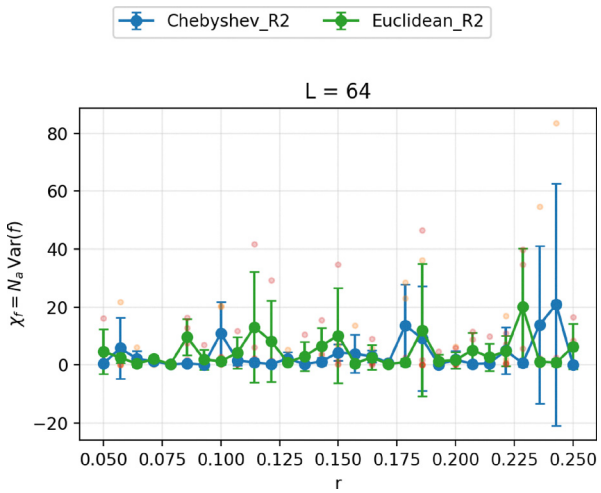

Fig. 9 Susceptibility-like proxy $\chi_f = N_a \mathrm{Var}(f)$ *versus* r for the size-extended system (L = 64), showing localized fluctuation enhancements consistent with finite-size connectivity crossovers.

1. The free solvent remains system-spanning throughout most of the explored density range, both at L = 48 and L = 64.

2. The amount of restricted water increases strongly with amphiphile density, and this monotonic growth is robust to system size.

3. Restricted water exhibits a kernel-dependent connectivity crossover: for sufficiently isotropic/extended neighborhoods (Chebyshev/Euclidean at R = 2), spanning becomes likely over a specific density interval. At L = 48 this appears as a clearly delimited (''transient'') window; at L = 64 the onset is smoother and spanning can persist intermittently to higher r in the explored range, indicating a finite-size modulation of the high-r behavior.

4. Amphiphile clustering increases with r in a comparatively smooth fashion, and neither $\langle f \rangle$ nor barrier-like statistics in

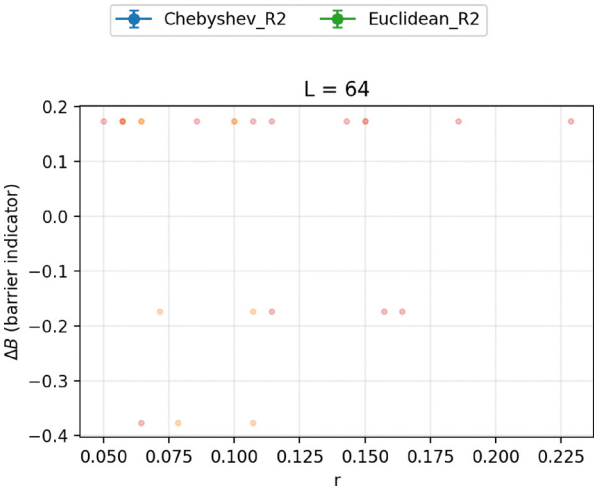

Fig. 10 Barrier-like statistic ΔB *versus* r for the size-extended system (L = 64), computed from the neck probability around the reference threshold in f.

f provide a universal sharp signature that mirrors the structured solvent-topology response.

Within this model, hydrophobic aggregation is therefore naturally framed as a solvent-connectivity problem: aggregation correlates with the regime where restricted-water shells overlap sufficiently to realize system-scale connectivity in the restricted subset. The L = 64 extension clarifies that the restricted-water spanning signal is best interpreted as a topological *crossover* whose detailed shape (including whether spanning is strictly ''transient'' at high r within a given scanned range) can depend on finite-size morphology and sampling resolution, while the broader conclusion—that solvent topology can change sharply even when solute ordering statistics evolve smoothly—remains robust.

4 Discussion

4.1 From aggregation to topology: reframing the hydrophobic effect

The central result of this work is that the most pronounced nontrivial signature in the simulations is not a sharp transition in amphiphile aggregation itself, but a topological crossover in the solvent. Across kernels and densities, the amphiphile largest-cluster fraction $\langle f \rangle$ typically evolves smoothly with r (and the barrier-like proxy ΔB remains comparatively featureless), whereas the solvent—when decomposed into free and restricted subsets—can display abrupt changes in global connectivity as measured by spanning probabilities. This separation supports the guiding viewpoint of the paper: within the present model, aggregation is not most naturally diagnosed by a discontinuity in a solute order parameter, but by a reorganization of the solvent into topologically distinct connectivity regimes.

The size-extension analysis strengthens this reframing. At L = 48 the solvent topology exhibits a sharply delimited spanning regime of restricted water for sufficiently isotropic kernels.

At $L = 64$ (with a denser sampling in r) the same qualitative mechanism persists—restricted-water spanning remains kernel-selective and emerges over a localized density range—but the detailed shape of the spanning curve becomes smoother and more fluctuation-dominated, as expected for a global, binary observable in finite systems. Thus, the refined conclusion is not merely that ''restricted water percolates transiently'', but that restricted-water connectivity undergoes a strong, kernel-controlled crossover whose apparent sharpness and high-r behaviour are modulated by finite-size morphology and sampling resolution.

This observation motivates a conceptual shift. Rather than interpreting hydrophobic aggregation as driven primarily by an effective attraction between apolar units, our results support a picture in which aggregation is a consequence of solvent reorganization. In this lattice model, the solvent does not simply become ''less available'' as amphiphile concentration increases; instead, its connectivity properties reorganize over a finite density range, producing a topological signature that can be sharper than the evolution of solute clustering statistics.

4.2 Restricted-water spanning as a solvent-driven crossover and its finite-size modulation

A central feature in the baseline dataset ($L = 48$) is that restricted water spans the system only over an intermediate range of amphiphile densities. This behaviour differs both from a strictly monotonic emergence of spanning and from a classical percolation transition with a single permanent threshold.

The $L = 64$ extension refines this picture. With increased system size and denser sampling in r, the onset of restricted-water spanning remains localized, but the high-r decrease becomes less sharply resolved and spanning can persist intermittently. This behaviour is consistent with finite-size effects in binary percolation diagnostics.

Accordingly, the most robust conclusion supported by both system sizes is that the restricted-water subset undergoes a strong connectivity crossover—from disconnected patches to frequent system-scale connectivity—whose detailed post-onset behaviour depends on system size and sampling resolution. In the explored size range, the data do not indicate clear critical sharpening with increasing L; rather, the onset becomes smoother and more fluctuation dominated, consistent with a finite-size topological crossover rather than a sharp asymptotic transition.

Within the model, the physical interpretation remains unchanged. At low densities, restricted solvent forms localized shells around isolated amphiphiles. As r increases, overlap between these domains enables system-scale connectivity. At larger r, morphological reorganization can either suppress a single backbone (as seen at $L = 48$) or sustain intermittent spanning (as observed at $L = 64$), indicating that restricted-water topology reorganizes in a structured, system-scale manner.

4.3 Kernel geometry and the role of interaction isotropy

A key outcome is the strong dependence of solvent topology on the interaction metric. Only kernels with sufficiently isotropic and extended neighbourhood support (Chebyshev_R2 and Euclidean_R2 in the studied set) generate a regime in which restricted water frequently spans the system. More anisotropic or lower-support kernels (Manhattan_R1, Cross_R1) fail to generate a spanning restricted-water network, with $P_{\mathrm{span}}(W_{\mathrm{restr}}) \approx 0$ across the scanned densities.

This kernel selectivity persists under the size extension. While the detailed shape of the spanning probability curve is modulated by L (as discussed above), the qualitative ordering of kernels is unchanged: the emergence of system-scale restricted-water connectivity requires an interaction neighbourhood that is sufficiently isotropic and spatially extended so that restriction domains can overlap and transmit connectivity at the system scale.

This highlights that the phenomenon is not governed solely by amphiphile density, but by how solvent sites ''sense'' the presence of nearby amphiphiles through the kernel. Equivalently, it is the effective dimensionality and isotropy of the interaction shell that controls whether local restriction patches can overlap into a connected structure. In the present minimal model, this offers a natural interpretation in terms of solvent structuring: only when the perturbation induced by amphiphiles is spatially isotropic enough does it propagate connectivity through the solvent micro-environment in a way that can become system-spanning.

While the present study focuses on a representative threshold θ and interaction range R, the qualitative mechanism underlying the restricted-water connectivity crossover is expected to be robust to moderate parameter variations. Changing θ primarily shifts the density window where restriction becomes significant, whereas increasing R tends to enhance isotropy and promote overlap of restriction domains. A systematic parameter sweep would be an interesting direction for future work but is not expected to alter the central topology-driven interpretation.

4.4 Decoupling solvent amount from solvent connectivity

An important and somewhat counterintuitive result is the decoupling between the amount of restricted water and its connectivity. The extensive restriction measure $\langle F \rangle$ increases strongly and (within the explored range) monotonically with amphiphile density for all kernels and for both system sizes. However, restricted-water spanning is highly structured and kernel dependent: it is absent for some kernels and, for isotropic kernels, it turns on over a limited density range and exhibits substantial nontrivial r-dependence.

The $L = 64$ extension strengthens this conclusion by separating two classes of observables according to their sensitivity. Quantities such as $\langle F \rangle$ (an extensive count) and $\langle f \rangle$ (an averaged solute-connectivity proxy) are comparatively stable under size increase, showing the same global trends with improved resolution. By contrast, spanning probabilities are intrinsically sensitive to large-scale morphology and thus are expected to show stronger finite-size modulation. This contrast supports an interpretational point that is central to the manuscript:

Local restriction can grow steadily with density while global connectivity reorganizes sharply and nontrivially.

Consequently, hydrophobic aggregation in this model cannot be inferred solely from local solvent ordering or confinement metrics. Instead, global topology must be considered. In particular, a large fraction of restricted water does not guarantee that solvent-mediated correlations are system-spanning. The hydrophobic effect, in this sense, appears as a collective phenomenon that depends on how local restrictions are spatially arranged and connected.

4.5 Relation to micellization and the critical micelle concentration

The connectivity crossover in restricted water naturally suggests a reinterpretation of the critical micelle concentration (CMC). In real amphiphilic systems, micellization is often described as the point beyond which additional amphiphiles preferentially join existing aggregates rather than remaining dispersed. Within the present model, an analogous picture emerges in terms of the solvent restriction field: there exists a density regime where restricted-water domains overlap sufficiently to realize system-scale connectivity, and beyond which further increases in amphiphile density are accommodated primarily by solute reorganization rather than by a qualitatively stronger solvent-connectivity effect.

In the $L = 48$ dataset, this appears as a restricted-water spanning window followed by a loss of spanning at higher r. In the $L = 64$ extension, the onset of restricted-water spanning remains localized, but the post-onset regime can remain topologically active (intermittently spanning) over a wider range. Both outcomes are consistent with the same conceptual translation to a CMC-like interpretation: there is a characteristic density range where the solvent micro-environment reorganizes most strongly at the system scale, after which further densification does not necessarily amplify solvent-mediated correlations in a simple monotonic fashion.

Crucially, the model does not predict a disappearance of free solvent above this regime. Both system sizes indicate that the free-water subset remains largely spanning across the scanned densities, and that the connectivity of the bulk solvent backbone is not destroyed. Rather, the model predicts that the solvent reorganizes such that the subset defined as "restricted" undergoes a structured connectivity crossover. This aligns qualitatively with the empirical idea that above the CMC, aggregates grow or multiply while the surrounding solvent remains continuous. From an experimental perspective, a solvent-topological crossover of this type could manifest through changes in hydration-shell connectivity, scattering signatures sensitive to solvent structuring, or kinetic markers in surfactant micellization and nanoparticle self-assembly. While the present model is deliberately minimal, these connections suggest concrete routes through which solvent-connectivity diagnostics might be probed in real soft-matter systems.

4.6 Absence of a strong first-order signature

The barrier-like statistic ΔB constructed from the distribution of the amphiphile order parameter does not show a pronounced growth or divergence across the explored density range. Together with the smooth behaviour of $\langle f \rangle$, this suggests that, within the resolution of the present simulations and order-parameter definition, the system does not exhibit a strong first-order-like transition in amphiphile ordering.

The size-extension results are consistent with this assessment. At $L = 64$, the additional resolution in r does not reveal a systematic growth of ΔB into a large-barrier regime; instead, the barrier-like curves remain relatively flat with scatter. Likewise, the susceptibility-like proxy χ_f (where available) shows localized fluctuation enhancements but not a universal, large-amplitude divergence. This combination of observations supports a conservative interpretation: the ordering proxy f evolves through a finite-size connectivity crossover rather than a strongly metastable, first-order-like jump under the present model definition.

Instead, the dominant "transition-like" structure resides in solvent topology. The most prominent sharp feature in the results is the onset (and structured evolution) of restricted-water spanning for isotropic kernels, which is a global-connectivity phenomenon. This supports the view that the relevant reconfiguration in this model is not a classical phase transition, but a topological crossover. Such crossovers can have sharp physical consequences (*e.g.*, abrupt changes in response or aggregation pathways) without requiring a discontinuity in conventional thermodynamic observables.

4.7 Interpretational limits of the model

It is important to emphasize the intentional simplifications of the present approach. The model represents only apolar segments of amphiphiles and does not explicitly encode polar head groups. As a consequence, the post-micellization regime—where real amphiphiles shield their hydrophobic cores with polar interfaces—is not represented. Accordingly, interpretations of high-r behaviour should be restricted to solvent-topological implications rather than detailed micelle morphology. Accordingly, the model should be interpreted primarily as a pre-micellar/near-micellar description rather than a representation of fully developed micellar structures.

This caveat becomes particularly relevant when interpreting the high-density behaviour of restricted-water spanning. Because the restricted-water subset is defined by a local count threshold, it does not distinguish between qualitatively different microscopic realizations of "restriction" (*e.g.*, narrow interfacial shells *versus* thick trapped domains). At high r, morphological rearrangements can produce different global-connectivity patterns depending on system size and finite-size geometry, as reflected in the differences between the $L = 48$ and $L = 64$ spanning curves. Thus, the most reliable conclusions should emphasize the existence and kernel selectivity of the connectivity crossover, rather than the precise fate of spanning at the largest scanned densities.

Within these limits, the results are best understood as describing a pre-micellar and near-micellar regime in which solvent reorganization plays a dominant role and in which hydrophobic aggregation can be meaningfully interpreted as emerging from solvent connectivity rather than direct solute attraction.

4.8 Broader implications

The present findings resonate with a growing body of work that interprets hydrophobic effects, and more generally self-assembly

in soft matter, as solvent-driven phenomena. By explicitly identifying a connectivity crossover of a restricted-solvent subset as a key feature, this work provides a concrete, topology-based mechanism linking local solvent perturbations to global aggregation behaviour.

More broadly, the results suggest that solvent topology—rather than only local energetics—may serve as a unifying descriptor for aggregation phenomena across different length scales and interaction geometries. The size-extension analysis reinforces this point: extensive local measures (such as $\langle F \rangle$) can behave smoothly and robustly with density, while topology-based measures (such as spanning) can reorganize sharply and remain sensitive to system-scale morphology. In this sense, hydrophobic aggregation emerges not as a direct force between apolar units, but as an entropic and topological consequence of how the solvent maintains global connectivity under increasing local constraints.

5. Conclusions

We presented a minimal two-dimensional lattice model in which hydrophobic (apolar) solute segments influence the solvent solely through a local neighborhood restriction rule, and configurations are sampled with a Boltzmann-like weight penalizing the total number of restricted-water sites at fixed composition. Within this framework, the most salient transition-like signature is not a dramatic discontinuity in a solute aggregation observable, but a topological crossover in the solvent when it is decomposed into free and restricted subsets.

A central methodological point of the study is that this conclusion is supported at two levels of resolution. The baseline dataset (L = 48) provides a clear qualitative mapping of regimes, while a systematic size extension (L = 64) with a denser sampling in solute fraction r tests the robustness of the connectivity phenomenology under changes in system size. This two-layer structure is essential because spanning/percolation is a global, binary observable in finite systems and is therefore particularly sensitive to morphology and finite-size effects.

Across the explored parameter range and for both system sizes, free water remains almost always spanning, indicating that the model does not predict a breakdown of the bulk solvent backbone under these conditions. Small deviations of $P_{\mathrm{span}}(W_{\mathrm{free}})$ from unity appear only at the largest r values and do not organize into a sharp connectivity threshold, supporting the interpretation that the bulk solvent remains globally connected even as its free volume fraction decreases with increasing solute loading.

In contrast, restricted water exhibits a strongly kernel-dependent connectivity response. For sufficiently isotropic and extended kernels (Chebyshev_R2 and Euclidean_R2 in the investigated set), the restricted-water subset undergoes a pronounced connectivity onset over an intermediate range of r, whereas more anisotropic or smaller-support kernels (Manhattan_R1, Cross_R1) fail to generate restricted-water spanning altogether. In the L = 48 baseline, the restricted-water spanning probability displays a transient percolation window: restricted water becomes system-spanning over an intermediate range of solute fractions and loses spanning again at higher r. The L = 64 extension refines this statement by showing that, once system size is increased and r is sampled more densely, the onset of restricted-water spanning remains localized and kernel selective, but the detailed high-r behavior becomes smoother and more fluctuation dominated, with spanning persisting intermittently within the explored range. This sensitivity is consistent with the nature of spanning as a finite-size topological diagnostic and indicates that the most robust claim is a restricted-water connectivity crossover whose apparent ''transient'' character at high r can be modulated by finite-size morphology and scan resolution.

Importantly, the connectivity behavior of restricted water does not simply mirror the monotonic growth in the amount of restricted water. The average restriction measure $\langle F \rangle$ grows strongly with r for all kernels and remains stable in its qualitative trend under the size extension, while restricted-water spanning can be absent, sharply activated, or strongly structured depending on kernel geometry. This establishes a clear separation between how much restricted water exists and whether it forms a system-spanning network, and it supports the central conceptual framing of the manuscript: the relevant organizing mechanism is not a purely local increase of restriction but the global spatial arrangement and overlap of restricted-water domains.

These findings support a restrained but firm interpretation: in this model, hydrophobic aggregation is naturally framed as a consequence of solvent reorganization and connectivity constraints rather than as evidence for an intrinsic attractive interaction between apolar units. The solvent can remain globally connected as ''free'' water while still undergoing a qualitatively distinct reconfiguration in the topology of the restricted subset, and this reconfiguration can be concentrated into a finite concentration interval. At the same time, conventional solute ordering diagnostics (the smooth evolution of $\langle f \rangle$ and the absence of a systematic growth of the barrier-like statistic ΔB) suggest that the model does not exhibit a strong first-order-like transition in amphiphile ordering under the present definition of the order parameter, reinforcing the interpretation that the most prominent ''transition-like'' feature is topological and solvent-centered.

Finally, the model's interpretational scope should be kept explicit. Because only apolar segments are represented and polar head-group physics is absent, the high-r regime should not be overread as a faithful description of post-micellization morphology. The model is best viewed as isolating a pre-micellar/near-micellar mechanism: a solvent-topology crossover capable of producing sharp connectivity signatures without requiring a strong first-order-like signal in conventional solute order parameters.

Future work can test and sharpen the present picture along several concrete directions: (i) extending the degrees of freedom to include explicit polar segments and amphiphile architecture; (ii) performing broader finite-size scaling of spanning

and connectivity observables to distinguish crossover sharpening from genuine asymptotic behavior; and (iii) exploring alternative restriction criteria that more directly encode solvent structural constraints (*e.g.*, thresholds tied to local solvent coordination or multi-scale neighborhood definitions). Together, such extensions would clarify the extent to which solvent-topology signatures provide a general, transferable descriptor of hydrophobic aggregation across models and length scales.

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

All scripts necessary to reproduce the simulations and analysis reported in this work, together with the full parameter sets used in the study, are openly available on Zenodo: https://doi.org/10.5281/zenodo.18776053.

Acknowledgements

Financial support from 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 M. Busato, G. Mannucci, L. A. Rocchi, M. E. Di Pietro, A. Capocefalo, E. Zorzi, P. Postorino, D. Veclani, F. Castiglione, A. Martinelli, A. Martinelli, P. Postorino and P. D'Angelo, *ACS Sustainable Chem. Eng.*, 2023, **11**, 8988–8999.
2 M. Kowacz and P. Warszynski, *Colloids Surf., A*, 2018, **557**, 94–105.
3 G. Raffaini, A. Mazzaglia and F. Ganazzoli, *Beilstein J. Org. Chem.*, 2015, **11**, 2459–2473.
4 R. Gu, J. Lamas, S. K. Rastogi, X. Li, W. Brittain and S. Zauscher, *Colloids Surf., B*, 2015, **135**, 126–132.
5 B. Raman, C. Sundari and D. Balasubramanian, *Indian J. Biochem. Biophys.*, 1992, **29**, 143–147.
6 H. Frank and M. Evans, *J. Chem. Phys.*, 1945, **13**, 507–532.
7 C. Tanford, *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, Wiley, 2nd edn, 1980, p. 233.
8 K. Lum, D. Chandler and J. Weeks, *J. Phys. Chem. B*, 1999, **103**, 4570–4577.
9 D. Chandler, *Nature*, 2005, **437**, 640–647.
10 P. G. Debenedetti, *Metastable liquids: concepts and principles*, Princeton university press, 2020.
11 T. F. Miller, III, E. Vanden-Eijnden and D. Chandler, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 14559–14564.
12 A. J. Patel, P. Varilly, S. N. Jamadagni, H. Acharya, S. Garde and D. Chandler, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 17678–17683.
13 V. Domínguez-Arca, *Soft Matter*, 2026, **22**, 109–118.
14 M. I. Chaudhari, S. B. Rempe, D. Asthagiri, L. Tan and L. R. Pratt, *J. Phys. Chem. B*, 2016, **120**, 1864–1870.
15 A. Gao, L. Tan, M. I. Chaudhari, D. Asthagiri, L. R. Pratt, S. B. Rempe and J. D. Weeks, *J. Phys. Chem. B*, 2018, **122**, 6272–6276.
16 M. Kinoshita, *J. Chem. Phys.*, 2008, **128**, 024507.
17 T. Urbic and K. A. Dill, *Phys. Rev. E*, 2017, **96**, 032101.
18 F. Stillinger, *Science*, 1980, **209**, 451–457.
19 M. V. Athawale, G. Goel, T. Ghosh, T. M. Truskett and S. Garde, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 733–738.
20 J. N. Israelachvili, *Intermolecular and Surface Forces*, Academic Press, 1991.