

RESEARCH ARTICLE | APRIL 17 2023

Phase separation and aggregation in multiblock chains

Athanassios Z. Panagiotopoulos  



J. Chem. Phys. 158, 154901 (2023)

<https://doi.org/10.1063/5.0146673>



15 April 2024 16:24:44



Lock-in Amplifier



Boxcar Averager



Zurich
Instruments

Boost Your Optics and
Photonics Measurements

Find out more

Phase separation and aggregation in multiblock chains

Cite as: J. Chem. Phys. 158, 154901 (2023); doi: 10.1063/5.0146673

Submitted: 15 February 2023 • Accepted: 30 March 2023 •

Published Online: 17 April 2023



Athanassios Z. Panagiotopoulos^{a)} 

AFFILIATIONS

Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, USA

^{a)} Author to whom correspondence should be addressed: azp@princeton.edu

ABSTRACT

This paper focuses on phase and aggregation behavior for linear chains composed of blocks of hydrophilic and hydrophobic segments. Phase and conformational transitions of patterned chains are relevant for understanding liquid–liquid separation of biomolecular condensates, which play a prominent role in cellular biophysics and for surfactant and polymer applications. Previous studies of simple models for multiblock chains have shown that, depending on the sequence pattern and chain length, such systems can fall into one of two categories: displaying either phase separation or aggregation into finite-size clusters. The key new result of this paper is that both formation of finite-size aggregates and phase separation can be observed for certain chain architectures at appropriate conditions of temperature and concentration. For such systems, a bulk dense liquid condenses from a dilute phase that already contains multi-chain finite-size aggregates. The computational approach used in this study involves several distinct steps using histogram-reweighting grand canonical Monte Carlo simulations, which are described in some level of detail.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0146673>

I. INTRODUCTION

Liquid–liquid phase separation in biological systems is considered a key mechanism for subcellular organization, with important implications for the functioning of living cells.^{1–3} More broadly, the formation of two liquid phases by polymers in both aqueous⁴ and non-aqueous⁵ solvents plays an important role in separation and purification technologies and provides a challenging domain for thermodynamic and statistical mechanical modeling methods. Similarly, the formation of finite-size aggregates (micelles) from both low-molecular-weight surfactants⁶ and macromolecular components⁷ is important for detergency, catalysis, and nanotechnology applications.

While sophisticated models have been proposed for describing biomolecular condensation,^{8,9} significant insights have also been obtained from simple models containing a small number of residue types. Such models have also proven useful in modeling non-biological two-phase systems and in understanding micellization in surfactant and block copolymer solutions. For example, a linear chain model with just hydrophobic and hydrophilic beads has been deployed to illustrate the sensitivity of phase behavior to sequence and generate many complex condensate morphologies,^{10,11} using

large-scale molecular dynamics simulations. A significant limitation of these simulations is that the precise character of the observed transitions (e.g., first-order vs continuous) cannot be determined. To overcome these issues and to investigate the interplay between aggregation and phase separation, it is important to have access to accurate estimates of the free energy of a system as a function of thermodynamic state variables over a broad range of concentrations and temperatures. For this purpose, a particularly useful approach involves utilizing a lattice chain model originally proposed by Larson *et al.*¹² The model was designed to represent chains with hydrophilic (head, “H”) and hydrophobic (tail, “T”) segments and has provided insights on micellization and phase transitions of diblock and triblock surfactant chains^{13,14} as well as multiblock chains consisting of identical repeat units,¹⁵ and it has been used to model liquid–liquid phase separation in disordered proteins.¹⁶

Prior work on aggregation and phase behavior of lattice chains with hydrophobic and hydrophilic beads has produced some general trends on the effects of sequence and chain length on the observed properties. The “all-T” version of the model is identical to the well-known lattice homopolymer system underlying Flory–Huggins theory.¹⁷ It has a first-order transition between dense and dilute fluid phases, for which the critical temperature increases and critical

volume fraction decreases as the chain length is increased. Short diblock chains such as H_2T_4 have been shown to form finite-size aggregates in a continuous transition, whereas longer ones, such as H_2T_8 , have a first-order dilute-dense fluid phase transition.¹⁴ For triblock chains, hydrophilic groups at the end favor micellization, whereas hydrophobic end groups favor phase separation; once again, longer sequences of T blocks lead to phase separation. For example the sequence $T_2H_8T_2$ micellizes, but $T_4H_8T_4$ phase separates. Multiblock chains follow the same pattern of chain length favoring phase separation: e.g., $(H_3T_3)_3$ forms finite-size aggregates, while $(H_3T_3)_4$ phase separates.¹⁵ In the most recent study of this model,¹⁶ a large number of sequences with lengths between 20 and 100 beads were examined and classified as phase separating or aggregating. In that study, sequences with long blocks of T beads were determined to phase separate while more evenly dispersed sequences phase separate. As chain length is increased, a large majority of long sequences was determined to phase separate, a fact of possible biological relevance for disordered proteins.

Remarkably, all prior studies mentioned earlier were not able to observe both formation of finite-size aggregates and phase separation for any specific chain architecture. This poses the interesting question of the interconnection between the different transition types. Specifically, it would be important to determine how one type of behavior changes to another for systems “on the cusp” between aggregation and phase separation. In this paper, we aim to address this question by revisiting some of these previously studied chain architectures reported to be at the boundary between the two types of behavior using new simulations combined with careful analysis of the results.

Specifically, we use grand canonical Monte Carlo simulations combined with histogram reweighting for the determination of phase coexistence curves, critical points, and aggregation boundaries. This approach is now quite mature, having served the community well for about twenty-five years. However, there are several subtle implementation issues for systems that may have multiple types of transitions. For this reason, this paper contains a pedagogical component with a relatively detailed description of the computational approach, in addition to the presentation of the new findings. The codes as well as example input and output files are available online, as detailed in the Data Availability Statement.

II. MODEL AND METHODS

A. Model

The model used in this study is shown schematically in Fig. 1. It consists of linear chains on a simple cubic lattice, entailing two types of beads, namely, hydrophobic T (red color) and hydrophilic H (blue color). Excluded volume interactions are active among beads so that only a single one can occupy any given lattice site. Chain connectivity is along the main directions of the lattice, with bonds between bead centers of length 1 in lattice units, but also diagonal in-plane bonds of length $\sqrt{2}$ and diagonal off-plane bonds of length $\sqrt{3}$. In other words, bond directions can be $[0\ 0\ 1]$, $[0\ 1\ 1]$, $[1\ 1\ 1]$, and their allowable rotations and reflections, resulting in 26 total possible connectivity vectors. These same vectors define nearest-neighbor attractive T–T interactions for non-bonded beads, of magnitude -1 ,

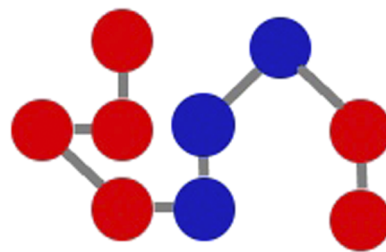


FIG. 1. A configuration for the triblock $T_4H_3T_2$ chain with energy $E = -2$. For simplicity, a two-dimensional snapshot is shown, even though all calculations were performed in three-dimensional space.

in reduced units that also set the temperature scale. All other interactions, namely, between H–T and H–H beads, as well as interactions involving the monomeric solvent beads occupying space not containing chains, are set to zero. In some early work on this model,^{13–15} the nearest-neighbor interaction energy was set to -2 for consistency with the original publications.^{12,18} This necessitates a scaling by a factor of 2 of temperatures, energies, and chemical potentials in this paper relative to the results reported in Ref. 16.

The model describes liquid–liquid separation between solvent-rich and solvent-lean phases in a two-component system of solvent and chains on a fully occupied lattice. A completely equivalent description of the model would be that it describes vapor–liquid equilibria of the one-component system of chains in a partly occupied lattice.

B. Grand canonical Monte Carlo

Monte Carlo sampling in the grand canonical ensemble was used to obtain all aspects of thermodynamic behavior for the lattice chain model described in Subsection II A. For a one-component system, the control variables of this ensemble are the chemical potential μ , volume V , and temperature T . For the two-component “filled lattice” equivalent description of the model, the first control variable is the chemical potential difference between chains and a number of monomers equal to the chain length, but we adopt the simpler one-component terminology. Cubic boxes of edge length L were simulated, so that the total number of sites was $V = L^3$. At least two different box sizes were used for each system studied, in order to quantify finite-size effects on the results.

The elementary Monte Carlo moves used in the simulations were chain insertion or removal attempts (typically 60% of total moves), partial chain regrowth (39.8% of total moves), and cluster moves (0.2% of total moves). The regrowth moves involve removing half of an existing chain and regrowing it again in an entirely new conformation. To facilitate insertion/removal and partial chain regrowth moves, an athermal Rosenbluth algorithm¹⁹ was implemented, as detailed in Ref. 20. Cluster moves involved identifying a cluster of interacting chains and displacing it by a single lattice unit in a randomly chosen spatial direction.¹³ To satisfy detailed balance, cluster moves are rejected if they result in the creation of a new cluster. These moves help with equilibration at low densities when aggregates are present, but become inefficient and computationally expensive at higher densities for which system-spanning large clusters frequently exist. Typical runs consisted of 2×10^9 Monte Carlo

moves and required a few hours of wall-clock time on a single core of a 3.6 GHz Intel Core i9 processor. The source codes, example input and output files, and a list of the runs performed for each system are available online, as explained in the Data Availability Statement at the end of this paper. Statistical uncertainties were obtained from the standard deviation of results from four independent runs at identical conditions.

Histogram reweighting was used to allow for full utilization of information from the grand canonical simulations, thus minimizing the number of separate simulation runs required. Specifically, we used the Ferrenberg–Swendsen algorithm²¹ to combine information from separate runs that have a reasonable degree of overlap in the number of particles N and energy E covered in each. The combined histograms were then used to obtain distributions of particle numbers and energy, $P(N, E)$, as a continuous function of μ and T . This capability is key to the determination of phase diagrams and critical points, as explained in the subsection that follows. Histogram reweighting methods used here are described in more detail in a review article.²² Computational requirements for histogram reweighting increase as $N^{3/2}$, so it is beneficial to operate at small system sizes, especially when exploring the behavior of a new chain sequence for which the type of transition and location of a possible critical point are not known. This scaling results from the fact that relative fluctuations in density, which are responsible for histogram overlap, are proportional to $N^{-0.5}$, so more runs are needed when bigger boxes are used. Also, sampling needs to be at least proportional to N , to allow for a similar number of moves per particle in the simulations.

C. Determination of phase diagrams and critical points

For a macroscopic system at equilibrium, one expects an abrupt jump in the number N of particles in the grand canonical simulation box at subcritical temperatures, occurring at a precise value of the chemical potential μ . However, in grand canonical simulations of finite systems, these transitions are subject to significant hysteresis because of the free energy barriers for the formation of a new phase. Hysteresis disappears at sufficiently high temperatures where barriers are small and runs can sample a range of densities covering both coexisting phases. For example, this is the case for the runs underlying the probability functions $P(N)$ shown in Fig. 2 (top). Such runs can be combined through histogram reweighting with subcritical runs of the separate phases to construct coexistence curves. This must be done in a self-consistent manner: the temperature–chemical potential curve at coexistence for a given system must be consistent with the underlying runs used to construct it. In other words, a dense phase simulation must have been performed at a chemical potential close to, but a little greater than, the chemical potential at coexistence at that temperature. A dilute-phase simulation must have been performed at a chemical potential a little lower than the chemical potential at coexistence at the temperature of interest. In practice, only a small number of new runs are required to achieve self-consistency.

Coexistence is determined by the condition of area equality for the low- N and high- N regions of the $P(N)$ distributions. This construction is meaningful only at lower temperatures than the ones

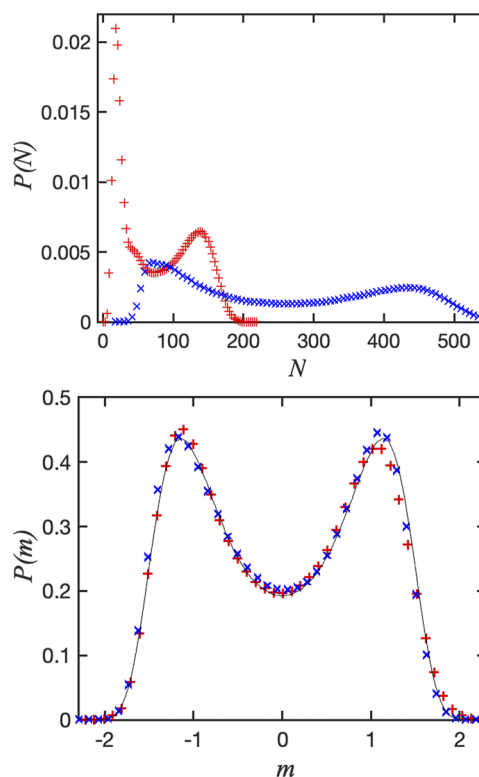


FIG. 2. (Top) Probability functions $P(N)$ for $T_4H_8T_4$ at conditions corresponding to the critical point. Red '+'s are for $L = 20$ and blue 'x's for $L = 30$. (Bottom) Transformed probability functions $P(m)$ for the same system and conditions. The thin black line is the universal three-dimensional Ising order parameter distribution from Ref. 23.

shown in Fig. 2 (top): in the figure, the two peaks potentially corresponding to coexisting phases have a great deal of overlap. The matching to the universal Ising distribution actually suggests that the runs are at the critical point, so an assignment of the peaks to coexisting phases is not appropriate for these conditions. At lower temperatures, however, there is a broad range of intermediate N values for which the probability $P(N)$ is low, resulting in a clear separation between high-density and low-density peaks. It should be emphasized that these deeply subcritical distributions cannot usually be sampled directly, as a single grand canonical run can only sample one phase because of the barriers to nucleation of the other one. Only the combination of these runs with near-critical ones that sample both sides makes it possible to determine coexistence (binodal) curves.

To determine critical points and provide unequivocal confirmation that an apparent phase transition is truly a first-order one between macroscopic phases, the mixed-field finite-size scaling method of Bruce and Wilding^{24,25} is used. The method transforms the two-dimensional histograms $P(N, E)$ into one-dimensional probability distributions $P(m)$, where the order parameter for the transition is now $m = N - sE$. The critical chemical potential μ_c , critical temperature T_c , and “field mixing” parameter s are determined for each chain architecture and system size L from the combined

histograms obtained at appropriate state conditions near the critical point. A typical two-dimensional frequency distribution $P(N, E)$ for a system at coexistence is shown in Fig. 5 of Ref. 22—the full two-dimensional distribution needs to be used for obtaining the critical parameters. For simplicity, we show one-dimensional probability distributions of particle numbers $P(N)$ in Fig. 2 (top), using two different system sizes for chains of type $T_4H_8T_4$. As can be seen in the figure, the $P(N)$ distributions are asymmetric, with the one for larger system size extending over a broader range of particle numbers. The transformed distributions $P(m)$ are shown in Fig. 2 (bottom) and compared to the universal Ising three-dimensional curve. While the fit is not perfect, especially for the smaller system size, it is clear that for this particular chain architecture and conditions, the order parameter probability distributions match the universal curve quite well.

D. Determination of critical micellar concentrations

Following earlier work,¹³ we use the equation of state data obtained through histogram reweighting to pinpoint the location of the critical micellar concentration (cmc), which marks the first appearance of multi-chain aggregates in systems of sufficiently large size. The grand canonical partition function Ξ is obtained directly from histogram reweighting for the range of particle numbers N covered by the combined histograms, within an unknown additive constant. Given that $\ln \Xi = \beta PV$, where $\beta = 1/k_B T$, a plot of this quantity vs N has a linear portion of unit slope at sufficiently low densities at which systems are close to an ideal gas and chain monomers dominate the dilute phase, as shown in Fig. 3. The unknown run-specific vertical shift of $\ln \Xi$ is obtained by matching this low-density portion to a line of unit slope going through the origin, but this shift does not influence the cmc estimation. At higher densities and sufficiently low temperatures, chain interactions start being active and the slope of the βPV vs N curve becomes markedly lower, signifying that the independent kinetic entities in the system are no longer monomers, but rather aggregates of multiple chains. The point at which this curve has a maximum second derivative (rate of change of the slope) is designated as the cmc. Remarkably,

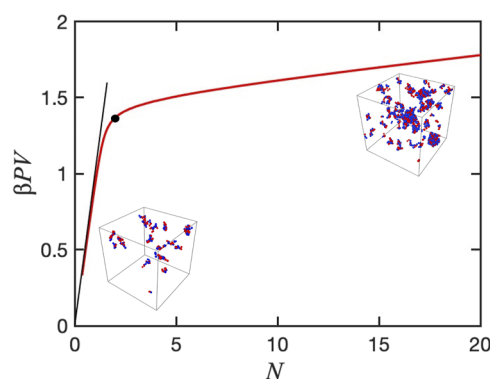


FIG. 3. (Main plot) The combination βPV as a function of N for $T_4H_8T_4$ at $T = 3.2$ for $L = 20$. A black dot marks the position of the maximum second derivative and, thus, the assigned cmc. The thin black line is drawn with a unit slope. (Insets) Snapshots of configurations in a box with $L = 50$ at conditions below and above the cmc.

the location of this transition obtained from the maximum in the second derivative is insensitive to simulation box size. In relatively small boxes, such as the one with $L = 20$ for which Fig. 3 is drawn, there are not enough chains in the system at the cmc to form aggregates of the size that would appear in larger boxes. However, the insets to Fig. 3 show representative snapshots of this system in boxes with $L = 50$ and clearly show the presence of multichain aggregates at concentrations above the cmc.

It should be pointed out that the sharp change in slope of the βPV vs N curves is not unique to systems forming aggregates—a first-order phase transition is also associated with a sharp decrease in slope, with the final value of the second segment depending on system size. Thus, a confirmation of the character of the transition (micellization into finite-size aggregates, or formation of a bulk condensed phase) needs to rely on cluster size distribution analyses at conditions beyond the computed cmc as well as on the determination of system-size effects on the equation of state.

III. RESULTS AND DISCUSSION

As suggested earlier, we focus on chain architectures that were identified in previous work^{13,15} to be at the boundary of phase separation (formation of macroscopic fluid phases) and micellization (formation of finite-size aggregates). The key questions we would like to address are the interplay between the two transition types and the possible identification of chain architectures for which both transitions are present at different thermodynamic conditions.

A. Systems exhibiting only phase separation

Diblock chains with long sections of solvophobic T beads and short sections of solvophilic H beads undergo phase separation: e.g., the system H_2T_8 was previously identified¹³ as being on the boundary between phase separation and micellization: the related sequence H_2T_4 was observed to form aggregates and H_2T_8 was characterized as phase separating. However, coexistence data were not reported in the prior study. The phase behavior for this system from the present work is shown in Fig. 4 for two different system sizes, in terms of

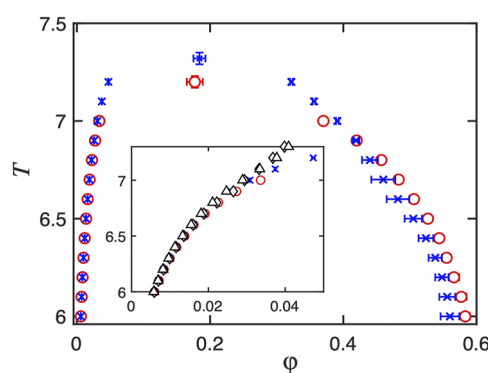


FIG. 4. Phase behavior for H_2T_8 . The main plot shows volume fractions at coexistence and critical points for $L = 30$ (red circles) and $L = 20$ (blue $\times$'s). The inset magnifies the low-density region and also shows apparent cmc's for $L = 30$ (triangles) and $L = 20$ (diamonds). Statistical uncertainties in this and subsequent figures are visible only when greater than symbol size.

temperature T vs volume fraction $\phi = nN/V$, where n is the number of beads per chain ($n = 10$ for H_2T_8). The main plot shows binodal curves and critical points determined as detailed in Sec. II C. These binodal curves and critical points are seen to be relatively insensitive to simulation box size. It is expected that the critical temperature would be a little lower for larger systems that allow density fluctuations of longer wavelengths within the simulation box and this is, indeed, what is observed. A minor system-size dependence is also seen for the subcritical dense-phase volume fraction. The relative invariance of the phase envelope to simulation box size confirms that this chain architecture has a normal bulk liquid phase separated from the dilute phase by a first-order phase transition.

The inset to Fig. 4 shows the low-density coexistence region and apparent cmc's, that is, loci of the maximum second derivative in the βPV vs N curves. The apparent cmc's are insensitive to box length L and occur at densities very near the corresponding dilute phase boundary. Additional simulations in large boxes ($L = 60$) in the small "gap" region between the apparent cmc's and the saturation (binodal) curve showed no micellar aggregates. Instead, the probability distributions decayed monotonically from monomers to dimers, trimers, etc., as expected in any system in which attractive interactions are present. For brevity, these distributions are not shown here but are available online as indicated in the Data Availability Statement. Thus, H_2T_8 chains show no sign of the formation of stable finite-size micellar aggregates, just as is the case for pure T (homopolymer) chains. The presence of the two-bead hydrophilic end segment results in significant reductions, of around 40%, for both the critical temperature and critical volume fraction relative to the T_8 homopolymer values reported in Ref. 20. A reduction of the critical volume fraction for sequences near the boundary between phase separation and aggregation was also observed in Ref. 16 for general chain sequences and a variety of chain lengths.

A similar "exclusively phase separation" behavior is obtained for a multiblock chain $(H_3T_3)_4$, previously identified as being on the boundary between phase separation and aggregation.¹⁵ The phase diagram for this system from this paper is shown in Fig. 5. As was the case for the previous H_2T_8 system, simulations in large boxes ($L = 80$) in the "gap" region between the apparent cmc's and the saturation (binodal) curve showed no micellar aggregates. The cluster probability distribution function decayed monotonically from monomers to oligomers, with no separate micellar peak, but in this

case had a long tail, with some larger aggregates of dozens of chains at temperatures near the critical point for the first-order transition. The critical temperature for the $(H_3T_3)_4$ system is about half that of H_2T_8 , despite the greater chain length and the presence of 50% more attractive T centers on each chain for the former system.

Even though a limited number of chain architectures were examined to provide examples of systems that display only phase separation and show no signs of aggregation in the dilute phase prior to saturation, it is anticipated that the vast majority of systems identified in prior studies as phase separating fall in this category. This includes chain architectures without extended hydrophilic blocks, as well as long chains of almost any sequence, except the most blocky ones. Formation of finite-size aggregates is the exception, rather than rule, occurring only when there is sufficient "protection" of solvophobic segments from further aggregation by sufficiently numerous solvophilic groups on the exterior of any multichain aggregates that form in the dilute phase.

B. Systems with both phase separation and aggregation

As stated earlier, prior studies had failed to identify chain architectures displaying both phase separation and aggregation. The reason for this failure turns out to be the fact that chain architectures and lengths for which this dual behavior is accessible from simulations are quite limited. This "dual" behavior only occurs over a relatively narrow range of thermodynamic conditions. In this paper, we were able to identify such systems by examining specific chain architectures in large boxes at chemical potentials corresponding to conditions between the putative cmc (determined from the inflection points of the βPV curves) and the binodal conditions [determined from the equal-area construction for $P(N)$]. Specifically, such behavior was identified for two triblock chain architectures and a multiblock one, as follows:

The first system determined to have dual behavior is the triblock chain architecture $T_4H_8T_4$, which had been previously identified as phase separating, but being close to the boundary for aggregation behavior.¹⁴ The phase behavior computed is shown in Fig. 6. As seen in the inset to the figure, there appears to be a gap between the concentration at which the cmc is assigned from the maximum in the second derivative of the βPV vs N curves and the coexistence (binodal) curve on the dilute phase side. As also seen

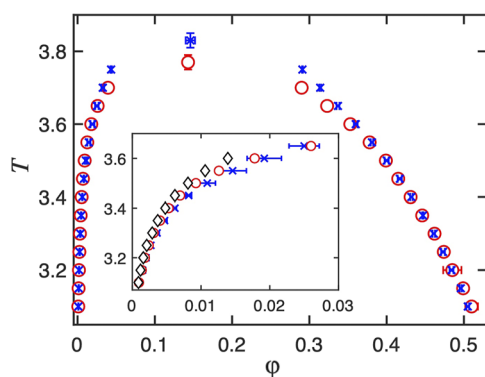


FIG. 5. Phase behavior for $(H_3T_3)_4$. Notation is the same as for Fig. 4.

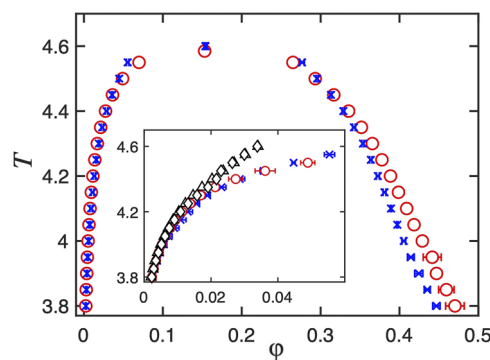


FIG. 6. Phase behavior for $T_4H_8T_4$. Notation is the same as for Fig. 4.

earlier, the cmc location is insensitive to system size. Relatively small boxes of $L = 20$ or $L = 30$ are optimal for cmc and coexistence curve calculations, as they afford good overlap between runs at different conditions through histogram reweighting. However, in these smaller boxes, at the coexistence density, there are relatively few chains present in the simulation box at any given time, not enough to form micellar aggregates. For this reason, it is necessary to perform simulations in large boxes to find out if aggregates form in the vapor phase. This needs to be done at chemical potentials between that corresponding to the cmc and phase coexistence. Probability distributions $P(M)$, where M is the size of aggregates, are determined through a cluster algorithm as detailed in Ref. 13. This function is shown in Fig. 7 for a box of $L = 60$ at concentrations within the dilute phase, above the cmc but fairly close to coexistence. A well-separated micellar peak can be seen in the data. At these conditions, about 15% of the chains in the dilute phase participate in micellar aggregates, which have an average size $\langle M \rangle = 34$.

The second system found to display both aggregation and phase separation is also a triblock chain, namely, $T_2H_8T_2$. This system was identified in Ref. 14 to be aggregating. Physically, we expect that the higher proportion of H blocks will result in a more protected hydrophobic core of the aggregates relative to $T_4H_8T_4$, promoting stable aggregates. The phase envelope and cmc obtained in this paper are shown in Fig. 8—these cmc's are in good agreement with results from Ref. 14. As for the previous triblock system, we find both aggregation and phase separation, but now with a significantly broader gap between the cmc and the phase boundary. There is also a more pronounced system-size dependence of the phase envelope. The critical temperature is about half that of the previous case, as expected from the smaller number of T beads in the attractive end segments. Simulations of the dilute phase at near-coexistence conditions in large boxes show micellar aggregates for this system, even at temperatures above the critical point for the first-order transition. For example, at the supercritical temperature of $T = 2.60$, a run with $\langle \phi \rangle = 0.09$, which is 50% higher than the cmc, had a well-separated micellar peak with $\langle M \rangle = 30$ and about 1/3 of the chains in the system being part of aggregates. Thus, for strongly micellizing systems such as $T_2H_8T_2$, there is a sizable temperature range over which aggregation is observed, but no first-order transition. This is

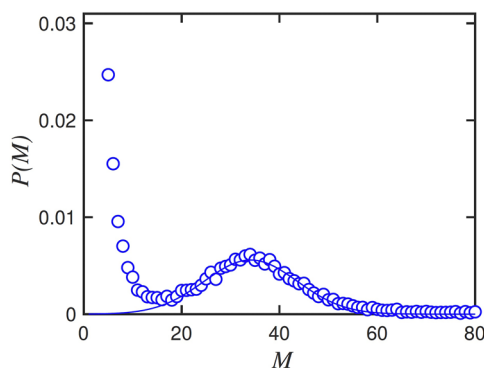


FIG. 7. Aggregation number probability function $P(M)$ for $T_4H_8T_4$, in a box of $L = 60$, at $T = 4.4$, $\langle \phi \rangle = 0.025$. The line is a Gaussian fit to the data points between $M = 20$ and $M = 70$.

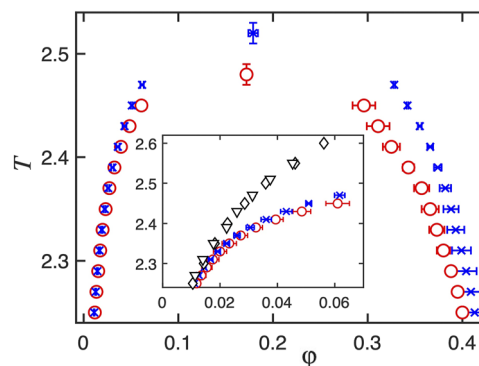


FIG. 8. Phase behavior for $T_2H_8T_2$. Notation is the same as for Fig. 4.

significant for interpreting the behavior of systems for which only aggregation is seen, discussed in Subsection III C.

Triblock chains with solvophobic outside blocks such as the two example architectures discussed here can have both of these attractive blocks be part of a single aggregate's solvophobic core, forming loops. Another option would be for the solvophobic blocks to bridge two different aggregates, which causes further aggregation. In experiments, such systems can lead to physical gelation as well as phase separation into chain-rich and chain-lean phases.²⁶

Another example of a system showing both aggregation and fluid-phase separation is the multiblock system $(H_3T_3)_3$. This system had been previously identified¹⁵ as aggregating. We have discussed earlier in this paper that the related $(H_3T_3)_4$ system, with one extra repeat unit, is phase-separating with no aggregates in the dilute phase prior to condensation. Simulations at multiple system sizes for $(H_3T_3)_3$ reveal a behavior qualitatively similar to that of the triblock chains discussed in this subsection. Data for this system are available as explained in the Data Availability Statement. There is a broad region of volume fractions (or chemical potentials) in which aggregates form in an unsaturated dilute phase prior to reaching the binodal limit. This system, however, suffers from slow equilibration because of the multiple connection possibilities of the aggregates and the relatively low temperature at which the phase transition is observed.

At this point, it is appropriate to comment on the relatively strong system-size dependence of the saturated liquid density seen for phase-separating architectures. One would normally expect small finite-size effects away from critical points. The larger differences seen here result from the microstructured character of the liquid phases, even when the overall phase behavior is normal. At lower temperatures, where the structuring of the liquid phases is more pronounced, larger systems can pack more efficiently, and this results in a faster change in the density of the liquid phases at lower temperatures, as observed in Figs. 4–6 and 8.

C. Systems exhibiting only aggregation

Systems with relatively short hydrophobic blocks have been previously determined¹³ to aggregate into finite-size clusters, rather than phase separate into a bulk liquid phase. An example of this

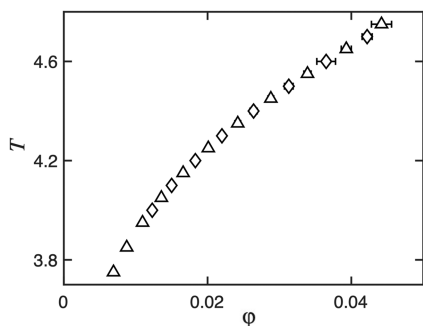


FIG. 9. Critical micellar concentration for H_2T_4 with $L = 30$ (triangles) and $L = 20$ (diamonds).

is H_2T_4 , which was identified as aggregating, whereas H_2T_8 is confirmed to be phase separating only, as discussed in Subsection III A.

The cmc curves obtained in this paper for H_2T_4 at two different system sizes are shown in Fig. 9. These cmc values agree well with those reported in Ref. 13. Because H_2T_4 has a relatively short hydrophilic segment that is not effective in protecting micelles from further aggregation, this system forms relatively large, disordered micellar clusters at concentrations above the cmc. For example, at $T = 4$, a run with $\langle\phi\rangle = 0.015$, which is just 20% above the cmc, in a box with $L = 60$, showed a broad micellar peak with $\langle M \rangle = 85$. At higher concentrations, aggregates grow and become roughly cylindrical in shape.

Probability distributions $P(N)$ vs N for two temperatures ($T = 4.3$ and 4) and two system sizes ($L = 20$ and 30) are shown in Fig. 10 for volume fractions $\langle\phi\rangle = 0.075$ (red curves) and $\langle\phi\rangle = 0.21$ (blue curves). These conditions are significantly above the cmc at either one of the two temperatures. The left-most peak, which is visible at $\langle\phi\rangle = 0.075$ for both temperatures and system sizes, corresponds to states with only oligomers (no micelles) present within the

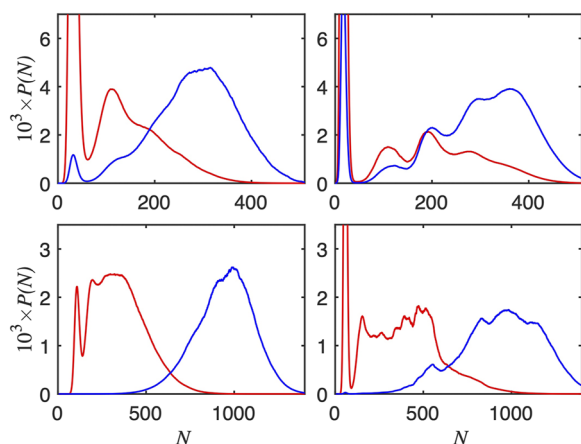


FIG. 10. Number probability functions $P(N)$ for H_2T_4 . Top row subplots are for $L = 20$, and bottom row for $L = 30$. Left column subplots are at $T = 4.3$, and right column subplots are at $T = 4.0$. The mean volume fractions are $\langle\phi\rangle = 0.075$ for the red curves and $\langle\phi\rangle = 0.21$ for the blue curves in all four subplots.

simulation box. The rest of the distributions corresponds to states with micellar aggregates of a broad range of sizes. These probability distributions of particle numbers are entirely inconsistent in shape with Ising-type order parameter distributions (Fig. 2 bottom), so there is no critical point for a first-order phase transition associated with H_2T_4 at any of the conditions studied.

The conclusion that there is no first-order phase transition for this system is also confirmed by forcing computation of a “pseudo-phase diagram” through assigning the low-density peak to a putative dilute phase and the rest of the $P(N)$ distribution to a putative dense phase. The result of this calculation is shown in Fig. 11. The apparent phase boundaries on the dilute-phase side are system-size invariant, but the dense-phase “coexisting” volume fractions are very strongly system-size dependent. The physical basis for this behavior can be seen from the representative configurations in the two different boxes shown in Fig. 11. The system-spanning irregular cylindrical micelles that form within the dense-phase boxes have roughly the same diameter, but occupy a bigger fraction of space in the smaller box. However, the chemical potential (and corresponding volume fraction) at which the aggregates form is approximately independent of system size, so the dilute-phase boundary is invariant with respect to system size. The large statistical uncertainties seen for the dense-phase volume fractions are a consequence of slow dynamics for this system due to the presence of system-spanning aggregates.

The H_2T_4 system is representative of many chain architectures assigned previously to aggregation behavior. From continuity arguments with respect to systems such as $T_2H_8T_2$, for which we have confirmed both aggregation and phase separation, we stipulate that the most likely fate of the first-order fluid-phase transition and associated Ising-type critical point is that the transition moves to lower temperatures than the ones at which micellization is readily observed. Any bulk liquid will have to be generated from attractive aggregate–aggregate interactions mediated by chain segments extending into the solvent. However, these chains have a strong preference for the hydrophobic cores of the aggregates, so they are unlikely to be found in such extended configurations. This leads to large barriers to the formation of a bulk liquid phase. Clearly, the dynamics of these systems also slow down considerably as

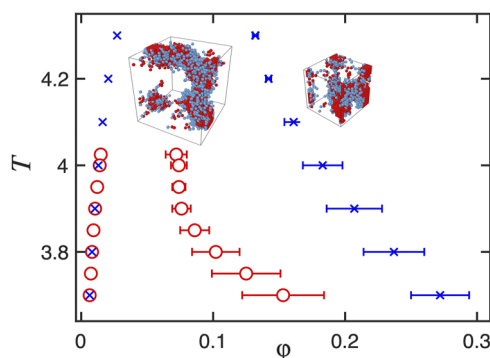


FIG. 11. Pseudo-phase diagram for H_2T_4 . The main plot shows volume fractions obtained from the area equality condition of the number probability functions for $L = 30$ (red circles) and $L = 20$ (blue x's). The inset shows representative configurations of the dense “phase” for the two system sizes.

temperature is reduced. As a result, we are unable to observe a true first-order transition for many aggregating chain architectures.

IV. CONCLUSIONS

In this work, we have re-examined the previously proposed distinction between macroscopic phase separation and formation of finite-size aggregates in systems of patterned linear chains consisting of hydrophobic and hydrophilic segments. By employing a careful analysis using grand canonical Monte Carlo simulations of a simple lattice mode, we were able to identify for the first time examples of both phase separation and aggregation behavior occurring within the same system at different thermodynamic conditions. Triblock chains with attractive outside segments are particularly prone to exhibiting this dual behavior, in agreement with experiments.²⁶ Short multiblock chains with a repeating hydrophobic-hydrophilic pattern are also seen to exhibit this behavior. In future work, we plan to examine the effects of weak sequence randomness on these observations.

In chain systems with attractive interactions, phase separation into macroscopic dilute and dense fluid phases is a universal feature, taking place for the majority of sequences for long chains.¹⁶ The exception to this behavior is chains with significant blocks of solvophilic beads, or very short chains, for which finite-size aggregates can be sterically prevented from attracting other aggregates and forming a macroscopic liquid, as confirmed in this paper. Aggregation behavior, which takes place at relatively low concentrations, was shown here to be possible at temperatures higher than the critical point for the fluid first-order phase transition.

We hypothesize based on our results that at sufficiently low temperatures, all systems, even those found to exhibit “aggregation only” behavior, have a first-order phase transition between a dilute fluid phase and a dense liquid. The micellization line, which marks the formation of a separate aggregate peak in the probability function $P(M)$ vs size M , and the binodal for the dilute phase, which marks the formation of the infinite-extent liquid, may be quite far apart for some chain architectures. Physically, this is due to the weakness of attractive interactions between aggregates because of the presence of solvophilic segments on their surface. However, it is expected that dynamical slow-down could prevent the actual observation of a first-order transition at temperatures for which a given system can reach equilibrium, in computer simulations or the laboratory. Multiblock chains were determined to be particularly prone to such non-equilibrium structure formation.

ACKNOWLEDGMENTS

This work was supported by the Princeton Center for Complex Materials (PCCM), a U.S. National Science Foundation Materials Research Science and Engineering Center (Award No. DMR-2011750).

AUTHOR DECLARATIONS

Conflict of Interest

The author has no conflicts to disclose.

Author Contributions

Athanassios Z. Panagiotopoulos: Conceptualization (equal); Investigation (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal).

DATA AVAILABILITY

Computer codes used in this work, example input and output files, information on the runs performed, and numerical data for the phase boundaries and cmc's are freely available for download from the Princeton Data-Space repository, at <https://doi.org/10.34770/ykvp-8b36>.

REFERENCES

- C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, “Germline P granules are liquid droplets that localize by controlled dissolution/condensation,” *Science* **324**, 1729 (2009).
- S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, “Biomolecular condensates: Organizers of cellular biochemistry,” *Nat. Rev. Mol. Cell Biol.* **18**, 285 (2017).
- Y. Shin and C. P. Brangwynne, “Liquid phase condensation in cell physiology and disease,” *Science* **357**, eaaf4382 (2017).
- R. Hatti-Kaul, “Aqueous two-phase systems,” *Mol. Biotechnol.* **19**, 269 (2001).
- T. Dobashi, M. Nakata, and M. Kaneko, “Coexistence curve of polystyrene in methylcyclohexane. II. Comparison of coexistence curves observed and calculated from classical free energy,” *J. Chem. Phys.* **72**, 6692 (1980).
- M. J. Rosen and J. T. Kunjappu, *Surfactants and Interfacial Phenomena*, 4th ed. (Wiley, Hoboken, NJ, 2012).
- G. Riess, “Micellization of block copolymers,” *Prog. Polym. Sci.* **28**, 1107 (2003).
- G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, “Sequence determinants of protein phase behavior from a coarse-grained model,” *PLoS Comput. Biol.* **14**, e1005941 (2018).
- T. M. Perdikari, N. Jovic, G. L. Dignon, Y. C. Kim, N. L. Fawzi, and J. Mittal, “A predictive coarse-grained model for position-specific effects of post-translational modifications,” *Biophys. J.* **120**, 1187 (2021).
- A. Statt, H. Casademunt, C. P. Brangwynne, and A. Z. Panagiotopoulos, “Model for disordered proteins with strongly sequence-dependent liquid phase behavior,” *J. Chem. Phys.* **152**, 075101 (2020).
- A. Statt, D. C. Kleeblatt, and W. F. Reinhart, “Unsupervised learning of sequence-specific aggregation behavior for a model copolymer,” *Soft Matter* **17**, 7697 (2021).
- R. G. Larson, L. E. Scriven, and H. T. Davis, “Monte Carlo simulation of model amphiphile-oil-water systems,” *J. Chem. Phys.* **83**, 2411 (1985).
- M. A. Floriano, E. Caponetti, and A. Z. Panagiotopoulos, “Micellization in model surfactant systems,” *Langmuir* **15**, 3143 (1999).
- A. Z. Panagiotopoulos, M. A. Floriano, and S. K. Kumar, “Micellization and phase separation of diblock and triblock model surfactants,” *Langmuir* **18**, 2940 (2002).
- M. E. Gindy, R. K. Prud'homme, and A. Z. Panagiotopoulos, “Phase behavior and structure formation in linear multiblock copolymer solutions by Monte Carlo simulation,” *J. Chem. Phys.* **128**, 164906 (2008).
- U. Rana, C. P. Brangwynne, and A. Z. Panagiotopoulos, “Phase separation vs aggregation behavior for model disordered proteins,” *J. Chem. Phys.* **155**, 125101 (2021).
- P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, NY, 1953).
- R. G. Larson, “Monte-Carlo lattice simulation of amphiphilic systems in two and three dimensions,” *J. Chem. Phys.* **89**, 1642 (1988).

- ¹⁹M. N. Rosenbluth and A. W. Rosenbluth, "Monte Carlo calculation of the average extension of molecular chains," *J. Chem. Phys.* **23**, 356 (1955).
- ²⁰A. Z. Panagiotopoulos, V. Wong, and M. A. Floriano, "Phase equilibria of lattice polymers from histogram reweighting Monte Carlo simulations," *Macromolecules* **31**, 912 (1998).
- ²¹A. M. Ferrenberg and R. H. Swendsen, "Optimized Monte Carlo data analysis," *Phys. Rev. Lett.* **63**, 1195 (1989).
- ²²A. Z. Panagiotopoulos, "Monte Carlo methods for phase equilibria of fluids," *J. Phys.: Condens. Matter* **12**, R25 (2000).
- ²³M. M. Tsy-pin and H. W. J. Blöte, "Probability distribution of the order parameter for the three-dimensional Ising-model universality class: A high-precision Monte Carlo study," *Phys. Rev. E* **62**, 73 (2000).
- ²⁴A. D. Bruce and N. B. Wilding, "Scaling fields and universality of the liquid-gas critical point," *Phys. Rev. Lett.* **68**, 193 (1992).
- ²⁵N. B. Wilding and A. D. Bruce, "Density fluctuations and field mixing in the critical fluid," *J. Phys.: Condens. Matter* **4**, 3087 (1992).
- ²⁶D. A. Vega, J. M. Sebastian, Y.-L. Loo, and R. A. Register, "Phase behavior and viscoelastic properties of entangled block copolymer gels," *J. Polym. Sci., Part B: Polym. Phys.* **39**, 2183 (2001).