Interplay between self-assembly and phase separation in a polymer-complex model

Tianhao Li,[1] W. Benjamin Rogers,[2] and William M. Jacobs[1,*]
[1]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*
[2]*Martin A. Fisher School of Physics, Brandeis University, Waltham, Massachusetts 02453, USA*

(Received 20 July 2023; accepted 14 November 2023; published 7 December 2023)

We present a theoretical model for predicting the phase behavior of polymer solutions in which phase separation competes with oligomerization. Specifically, we consider scenarios in which the assembly of polymer chains into stoichiometric complexes prevents the chains from phase-separating via attractive polymer–polymer interactions. Combining statistical associating fluid theory with a two-state description of self-assembly, we find that this model exhibits rich phase behavior, including reentrance, and we show how system-specific phase diagrams can be derived graphically. Importantly, we discuss why these phase diagrams can resemble—and yet are qualitatively distinct from—phase diagrams of polymer solutions with lower critical solution temperatures.

DOI: 10.1103/PhysRevE.108.064501

I. INTRODUCTION

Predicting the phase behavior of biopolymer systems is challenging due to the diversity of conformational and oligomeric states that biopolymers can populate [1,2]. In this context, oligomers refer to small molecular clusters that form due to attractive interactions among biomacromolecules, including proteins [3,4] and nucleic acids [5,6]. The interplay between oligomerization and phase separation is complex and can give rise to qualitatively different behaviors in different scenarios. In some cases, phase separation correlates with oligomer stability, either because phase separation is enhanced by the presence of oligomers [3,4,7–10] or because phase separation promotes oligomer formation in condensed phases [1,11]. Yet in other cases, phase separation appears to compete with oligomer formation, meaning that the presence of stable oligomers tends to inhibit demixing into condensed biopolymer-rich phases [12].

Here we explore this latter scenario—competition between the self-assembly of oligomers and phase separation—using a thermodynamically consistent mean-field model. We consider polymer solutions in which multiple polymer chains can assemble into a *stoichiometric complex*, which is an oligomer comprising a precise number of chains in a well-defined geometry [13]. A theoretical challenge is that this scenario cannot be described in terms of pairwise interactions alone, since the assembly of discrete stoichiometric complexes implies that spatial correlations among multiple polymer chains cannot be ignored [14,15]. We therefore propose a coarse-graining strategy in which a two-state model of stoichiometric complex self-assembly is coupled to a mean-field model of interacting, yet spatially uncorrelated, polymer chains within the framework of statistical associating fluid theory (SAFT) [16].

This theoretical approach treats self-assembly and phase separation in a self-consistent framework [17,18], allowing us to predict the effects of self-assembly on the equilibrium phase behavior of associating polymer solutions [19]. In the present article, we systematically develop this theoretical approach and study the implications of the resulting model under wide-ranging scenarios. In a closely related article [20], we separately provide validation of our theoretical predictions in a particular DNA-based experimental system, which utilizes the programmability of Watson-Crick base pairing [21] to tune both the thermodynamic stability of a designed stoichiometric complex and the phase behavior of the DNA oligonucleotides. Taken together, the results of these studies demonstrate the validity of the assumptions employed in our theoretical approach and confirm the quantitative accuracy of our model's predictions.

The key theoretical results of our approach are presented in this article as follows. In Secs. II A and II B, we describe the mean-field model and our numerical strategy for calculating its phase behavior in various parameter regimes. Then, in Secs. II C and II D, we investigate two scenarios in which the excluded volume of a stoichiometric complex is either the same as or greater than the total excluded volume of its constituent chains. Both of these scenarios exhibit reentrant phase behavior, in which a constant-concentration polymer solution transitions from one phase to two coexisting phases and back to a single phase as a single control parameter is tuned monotonically. This control parameter represents an experimentally accessible quantity such as, but not limited to, temperature, ionic strength, pH, or a cosolvent concentration. Importantly, although many features of the phase behavior resemble qualitative aspects of phase diagrams with a lower critical solution temperature (LCST), we find that competition between self-assembly and phase separation does not result in a critical point of this type. Instead, we find that phase coexistence can terminate at a first-order transition due to the competition between self-assembly and phase separation. Our approach shows how qualitatively different phase diagrams can arise by altering the dependence of key dimensionless parameters on experimentally controllable variables. A more thorough

*wjacobs@princeton.edu

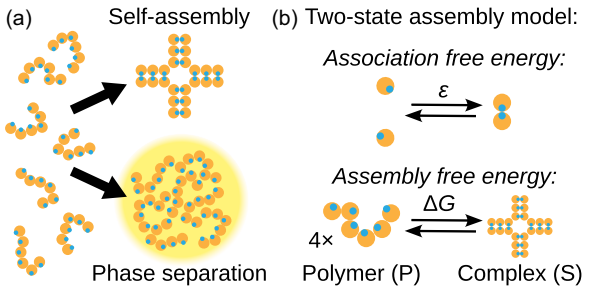

FIG. 1. Schematic of the theoretical model. (a) Biopolymers can either assemble into stoichiometric complexes or form a condensed phase via phase separation. Polymers are represented as freely jointed chains of blobs (yellow circles), each of which contains a single binding site (blue dots). Here we consider stoichiometric complexes consisting of $n = 4$ polymers. (b) The model involves two independent free-energy parameters, ϵ and ΔG, which govern the pairwise association between binding sites on blobs and the assembly of chains into stoichiometric complexes, respectively. Associative interactions between blobs are treated using statistical associating fluid theory (SAFT), while the two-state model of stoichiometric complex assembly is treated as an equilibrium chemical reaction between the polymer (P) and complex (S) states.

discussion of the implications of these results is provided in Sec. III.

II. THEORETICAL MODEL AND RESULTS

A. Two-state self-assembly and phase separation

We first develop a mean-field model of coupled self-assembly and liquid–liquid phase separation Fig. [1(a)]. We consider a polymer solution consisting of a single polymeric species in an implicit solvent. Polymers are coarse grained into chains of N_p "blobs," each of which represents a Kuhn segment [22] of the polymer with excluded volume v_0. This coarse-graining allows us to treat polymers as freely jointed chains, meaning that the relative orientations of adjacent blobs are uncorrelated. The total excluded volume of a free polymer is $v_p = N_p v_0$. To account for "specific" short-ranged attractive interactions, such as hybridization of DNA oligomers [21] or screened electrostatic interactions between charged sidechains of polypeptides [23], we introduce N_p distinct binding sites on each polymer. Each blob can accommodate a single coarse-grained binding site within this framework, since the orientations of monomers within a Kuhn-segment blob are highly correlated and thus not independent of one another [22]. Specific interactions between pairs of binding sites therefore represent coarse-grained, effective interactions between two Kuhn-length groups of monomers.

This modeling approach relies on three key assumptions. First, we assume that each binding site can participate in at most one specific attractive interaction at a time. Second, we assume that the assembly of n polymers into a stoichiometric complex is highly cooperative, such that partially assembled intermediates involving fewer than n polymers are vanishingly rare at equilibrium (i.e., two-state self-assembly). Third, we assume that all binding sites engage in specific interactions within a stoichiometric complex in the assembled state. This final assumption can be relaxed, leaving additional binding sites on assembled complexes available for interactions with binding sites either on other complexes or on free polymers [20], but we do not consider this possibility in the present work.

Taken together, these assumptions imply that self-assembly and phase separation are mutually exclusive [Fig. 1(a)]: Associative interactions between available binding sites are required to drive phase separation, and yet these binding sites are completely saturated by the formation of a complete stoichiometric complex. We note that, in practice, satisfying the second, "two-state" assumption typically requires a stoichiometric mixture of n different polymers with distinct binding sites, such that each polymer species is used exactly once in constructing a unique stoichiometric complex [24,25]. The model that we describe here treats all polymers as being indistinguishable for simplicity, which means that the specific interactions between free polymers represent effective interactions that average over all possible pairs of binding sites.

We implement these assumptions within a mean-field theory by combining SAFT [16] with a variation of the Flory–Huggins model of a polymer solution [22]. In accordance with the second key assumption of two-state self-assembly, chains can either exist in the free polymer state (P) or as part of a stoichiometric complex (S) consisting of n chains. The excluded volume of a complex is v_s, which we relate to the excluded volume of a polymer blob via $N_s \equiv v_s/v_0$ for notational convenience. Importantly, because our partitioning of chains into either P and S states allows us to treat the configuration of an assembled complex independently from that of individual polymers, it is not necessarily the case that $v_s = n v_p$. We shall consider both scenarios, where the excluded volume is either conserved or not conserved, in Secs. II C and II D, respectively.

The free-energy density, f, comprises four terms,

$$f = f^{\mathrm{id}} + f^{\mathrm{ev}} + f^{\mathrm{assoc}} + f^{\mathrm{int}}, \tag{1}$$

accounting for the ideal (id), excluded-volume (ev), associative (assoc), and internal (int) contributions, respectively. Each term has units of energy per volume. The ideal contribution depends on the concentrations (i.e., number densities) of polymers in both the P and S states,

$$\beta f^{\mathrm{id}} v_0 = \frac{\phi_p}{N_p}(\ln \phi_p - 1) + \frac{\phi_s}{N_s}(\ln \phi_s - 1), \tag{2}$$

where $\phi_i \equiv v_i \rho_i$ is the volume fraction and ρ_i is the concentration (or number density) of state $i \in \{\mathrm{P}, \mathrm{S}\}$. To avoid discussing explicit temperature dependence at this point, we use the inverse temperature $\beta \equiv 1/k_B T$, where k_B is the Boltzmann constant and T is the absolute temperature, to express all free energies as dimensionless quantities. We employ a mean-field formula for the excluded volume contribution that accounts for the fact that assembled complexes are essentially rigid bodies, whereas free polymers are freely jointed chains,

$$\beta f^{\mathrm{ev}} v_0 = \frac{\phi_s}{N_s} + \frac{1-\phi_s}{N_s}\ln(1-\phi_s) + \frac{\phi_p}{N_p}[1 - \ln(1-\phi_s)] + \phi_0 \ln\left(\frac{\phi_0}{1-\phi_s}\right), \tag{3}$$

where $\phi_0 \equiv 1 - \phi_s - \phi_p$ is the volume fraction not occupied by either free polymers or stoichiometric complexes. A derivation of this formula is presented in Appendix A, and a comparison to the Flory-Huggins model is provided in Appendix B.

Following the two-state assumption, the associative contribution only accounts for interactions among binding sites on P-state polymers. The free-energy density is given by Wertheim's thermodynamic perturbation theory [14], taking the P/S mixture without attractive interactions [i.e., Eqs. (2) and 3)] as the reference state [15],

$$\beta f^{\text{assoc}} v_0 = \phi_p \left(\ln X_p - \frac{X_p}{2} + \frac{1}{2} \right), \tag{4}$$

where X_p is the fraction of binding sites on P-state polymers that are not associated. X_p is determined by the chemical equilibrium expression

$$X_p + \phi_p e^{-\beta\epsilon} X_p^2 = 1, \tag{5}$$

where ϵ is the association free energy between a pair of binding sites [Fig. 1(b)]. Equation (4) can be derived by minimizing the Helmholtz free energy of a mixture of reactive binding sites; Eq. (5) follows directly from the derivation of f^{assoc}. We refer the reader to Ref. [26] for an elegant derivation and a thorough discussion. Multiplying Eq. (5) by ϕ_p, we can easily identify this chemical equilibrium condition as a mass-action expression relating the volume fraction of blobs that are not associated, $X_p\phi_p$, and the volume fraction of blobs that are associated, $(1 - X_p)\phi_p$, subject to an equilibrium constant $\exp(-\beta\epsilon)$. Association does not occur in the limit $\beta\epsilon \to \infty$. Importantly, Eqs. (4) and (5) assume that the association states of different binding sites are uncorrelated. This requirement is consistent with our coarse-grained treatment of association, in which we assume that each Kuhn-segment blob accommodates a single binding site.

By contrast with the associative interactions, the internal contribution to the free-energy density only depends on the concentration of S-state complexes,

$$\beta f^{\text{int}} v_0 = \frac{\phi_s}{N_s} [\beta\Delta G + n(1 - N_p)], \tag{6}$$

where ΔG represents the free-energy change (excluding translational entropy) associated with the self-assembly reaction $n\text{P} \rightleftharpoons \text{S}$ [Fig. 1(b)]. This reaction is at equilibrium when

$$n\mu_p = \mu_s, \tag{7}$$

where the P- and S-state chemical potentials are $\mu_p \equiv \partial f/\partial\rho_p$ and $\mu_s \equiv \partial f/\partial\rho_s$, respectively. Intuitively, Eq. (6) says that the contribution to the free energy due to polymer–polymer interactions within an S-state complex is proportional to the concentration of S-state complexes, $\rho_s = \phi_s/N_s v_0$. We motivate the particular definition of $\beta\Delta G$ established by Eq. (6) by discussing a limiting case at the end of this section.

This treatment allows us to study the consequences of stoichiometric complex self-assembly within a mean-field theory. In reality, both self-assembly and association among free polymers are driven by the same physical interactions among binding sites [Fig. 1(a)]. These interactions among binding sites can be assumed to be essentially pairwise in nature. However, the probability that a specific pair of binding sites is associated is highly correlated with the binding status of proximal binding sites on the same chains when self-assembly takes place. Such correlations are challenging to capture within mean-field models via thermodynamic perturbation theory. By contrast, including a discrete S state via the two-state approximation and the equilibrium condition Eq. (7) allows us to account for these correlations, while still treating the associative interactions between binding sites on P-state polymers as uncorrelated in Eqs. (4) and (5).

Nonetheless, the common physical origin of polymer association and self-assembly implies that the association free energy, ϵ, and the assembly free energy, ΔG, are coupled parameters that must change in tandem when a control parameter is tuned. To be precise, ΔG depends on both the sum total of the $nN_p/2$ binding interactions with strength ϵ within a stoichiometric complex, as well as entropic considerations related to the geometry and flexibility of the stoichiometric complex. In Secs. II C and II D, we show how system-specific relationships between ϵ and ΔG can be used to derive a phase diagram as a function of a single control parameter.

Before proceeding, it is instructive to consider the behavior of this model in three limiting scenarios. If $\beta\epsilon$ is finite and $\beta\Delta G \to \infty$, then self-assembly cannot occur and the model reduces to that of an associative polymer solution in which the polymers obey the statistics of freely jointed chains [19,22]. This polymer solution has a critical point at a dimensionless associative free energy $(\beta\epsilon)_c$ and polymer volume fraction ϕ_c. This critical point is commonly described as an upper critical solution temperature (UCST), since phase separation only occurs at temperatures below $T_c = \epsilon k_B^{-1} (\beta\epsilon)_c^{-1}$ if ϵ is taken to be a temperature-independent constant [22]. Alternatively, if $\beta\epsilon$ is finite and $\beta\Delta G \to -\infty$, then the solution is a fluid of S-state complexes in implicit solvent. The model reduces to a regular solution [27] with excluded volume interactions only in this case, since we have assumed that there are no associative interactions among completely assembled complexes. Lastly, if $\beta\epsilon \to \infty$ and $\beta\Delta G$ is finite, then there are no associative interactions among free polymers, and the model reduces to a concentration-dependent description of two-state self-assembly. In this case, Eq. (7) simplifies to an expression of chemical equilibrium for a reaction involving ideal gases in the dilute limit, $\phi_p \ll 1$ and $\phi_s \ll 1$:

$$\frac{\phi_s}{\phi_p^n} = e^{-\beta\Delta G}. \tag{8}$$

This limiting scenario justifies our expression for the internal free energy in Eq. (6).

B. Phase-coexistence calculations and master phase diagrams

Our aim is to compute the phase behavior of the mean-field model introduced in the previous section. This model is completely specified by the five independent parameters n, N_p, N_s, $\beta\epsilon$, and $\beta\Delta G$, which are all considered to be constants when carrying out phase-coexistence calculations. We shall assume that the number of chains per stoichiometric complex is $n = 4$ throughout this article, and then consider representative scenarios for various combinations of the remaining four parameters. Although the free-energy density, f, is a function

of both ϕ_p and ϕ_s, the concentrations of free polymers and stoichiometric complexes are related via the chemical equilibrium constraint, Eq. (7). The total polymer concentration, $\rho \equiv \rho_p + n\rho_s$, is therefore the sole independent variable when performing phase-coexistence calculations. For convenience, we present our results in terms of a dimensionless total concentration $\rho N_p v_0$, which is equal to the volume fraction if all polymers are in the P state, since this dimensionless quantity is bounded by 0 and 1. We present the composition of a solution with fixed total concentration in terms of the fraction of polymers in the S state, i.e., the S fraction, which is equal to $n\rho_s/\rho$. The S fraction is also bounded by 0 and 1.

For a given set of parameters, we determine the concentrations and compositions of coexisting phases by computing and analyzing a free-energy landscape as a function of ρ. This calculation is carried out in two steps. In the first step, we determine the equilibrium S fraction and the corresponding free-energy density at a fixed value of ρ by solving Eq. (7). In cases where there are multiple solutions to this chemical equilibrium equation, we choose the one with the lowest free-energy density; this choice is equivalent to performing a global minimization of f as a function of the S fraction at constant ρ. An example calculation showing the equivalence between constrained free-energy minimization and chemical equilibrium is illustrated in Fig. 2(a).

In the second step, we repeat this chemical-equilibrium calculation for many values of the total concentration to construct the equilibrium free-energy landscape $f(\rho)$ numerically. We then use the common-tangent construction [27] to identify one or more coexistence regions, where two total concentrations have the same chemical potential $\mu \equiv \partial f/\partial\rho$ and osmotic pressure $P \equiv f - (\partial f/\partial\rho)\rho$, if any such regions exist. An example landscape with two coexistence regions is illustrated in Fig. 2(b). In practice, we compute the convex hull of a free-energy landscape computed on a grid of discrete total concentration points, $\{(\rho, f(\rho))\}$ [28]. Any point $(\rho, f(\rho))$ that is not part of the convex hull must be within a coexistence region, in accordance with the common-tangent construction; adjacent grid points that are not part of the convex hull must be within the same coexistence region. In this way, we are able to identify the number of coexistence regions and obtain estimates of the total concentrations of the coexisting phases for each region. We then refine these estimates to compute the total concentrations and compositions of each pair of coexisting phases α and β. This final refinement step is achieved by solving the equal-chemical-potential, $\mu^{(\alpha)} = \mu^{(\beta)}$, and equal-osmotic-pressure, $P^{(\alpha)} = P^{(\beta)}$, conditions to numerical precision. The results of these calculations are most easily visualized by examining the grand-potential density, $f - \mu\rho$, with μ chosen to be equal to the chemical potential of the two coexisting phases. The grand-potential density is equal to the osmotic pressure at equilibrium. For example, the case shown in Fig. 2(c) corresponds to the lower-concentration coexistence region found in Fig. 2(b), in which a dilute solution consisting primarily of assembled stoichiometric complexes coexists with a condensed phase of disassembled polymers stabilized by their associative interactions.

To summarize the dependence of the phase behavior on the parameters $\beta\epsilon$ and $\beta\Delta G$, we introduce the concept of a *master phase diagram*. Specifically, for fixed choices of N_p and N_s,

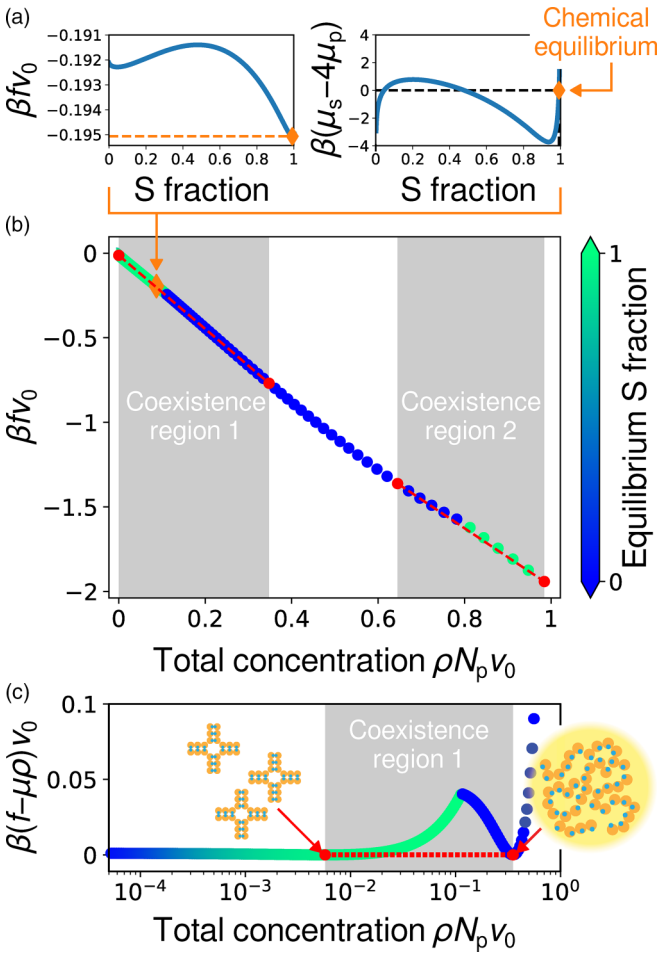

FIG. 2. Schematic of phase-coexistence calculations. (a) For a given chain concentration ρ, we first calculate the equilibrium composition by minimizing the free energy with respect to the stoichiometric complex (S) fraction (left). This calculation satisfies the condition of chemical equilibrium, $\mu_s = n\mu_p$, where $n = 4$ in this example (right). The equilibrium free energy is indicated by the orange diamond ($\rho N_p v_0 = 0.093$), assuming $N_p = 6$, $N_s = 24$, $\beta\epsilon = -4.025$, and $\beta\Delta G = -27$ [corresponding to point d in Fig. 3(a)]. (b) Performing chemical equilibrium calculations at various total concentrations results in a free-energy profile. Points on this profile are colored according to the equilibrium fraction of chains that assemble into stoichiometric complexes. Common-tangent constructions (red dashed lines) are then used to identify coexistence regions (shaded regions). Red points indicate the concentrations and free energies of the coexisting phases. In this example, there are two distinct coexistence regions at different ranges of the total concentration. (c) The common-tangent construction ensures that the coexisting phases have the same chemical potential, $\mu \equiv \partial f/\partial\rho$, and the same grand potential, $f - \mu\rho$. As an illustrative example, we show the grand potential at the coexistence chemical potential of the first coexistence region defined in (b). The equilibrium compositions indicate that a dilute phase of mostly assembled complexes coexists with a condensed phase of mostly free chains.

we determine the number of distinct coexistence regions that exist over the complete range of total concentrations, $0 \leqslant \rho N_p v_0 \leqslant 1$, as a function of $\beta\epsilon$ and $\beta\Delta G$. Example master phase diagrams are shown in Figs. 3(a) and 5(a), and are discussed in the following sections. In this representation,

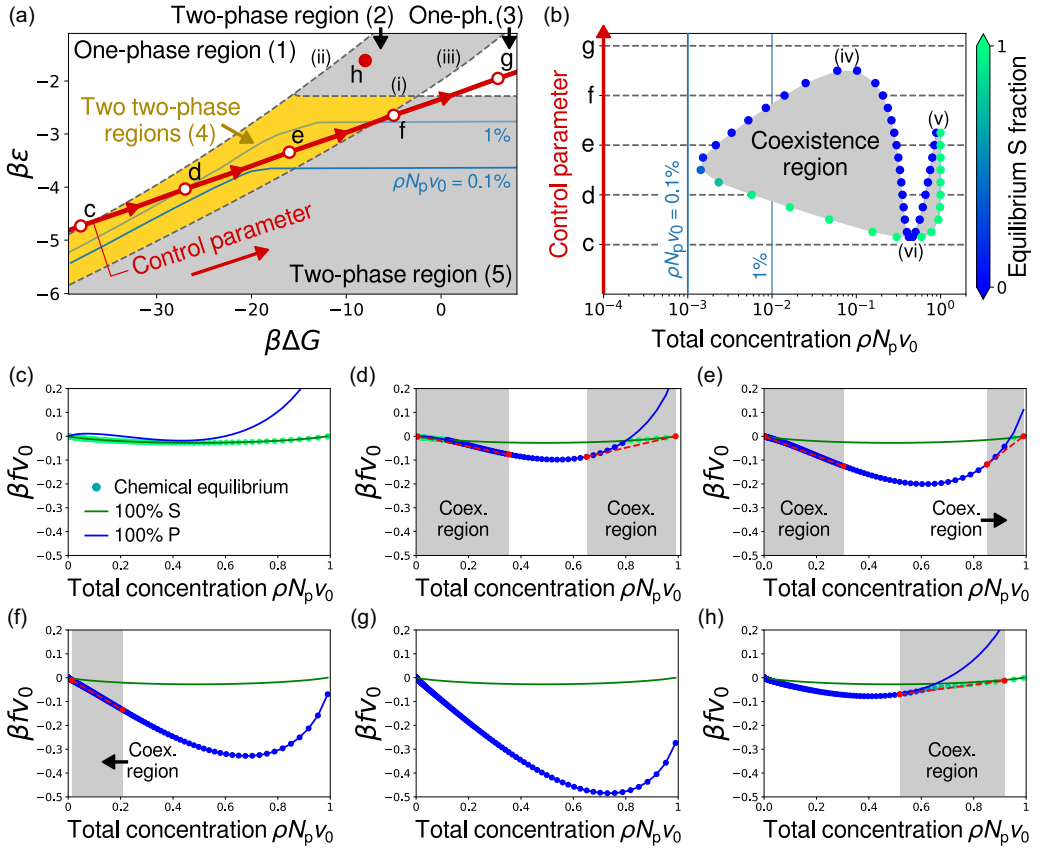

FIG. 3. Phase-behavior predictions when self-assembly conserves the excluded volume of the constituent chains. (a) The master phase diagram predicted by our model exhibits rich phase behavior. The parameter space is shaded according to the number of distinct coexistence regions at a given $\beta\epsilon$ and $\beta\Delta G$. In the unshaded regions, the system is homogeneous at all concentrations. Gray shaded regions indicate that phase separation occurs in a single coexistence region, while the gold region indicates the existence of two distinct coexistence regions within different concentration ranges. These regions are bounded by dashed curves (i), (ii), and (iii), as described in the text. Solid blue curves show where the dilute phase (of the lower-concentration coexistence region, if there are two) has the indicated total concentration. The path taken by a hypothetical control parameter, which relates $\beta\epsilon$ to $\beta\Delta G$, is shown by a solid red line; labeled points correspond to the free-energy profiles in (c)–(h). (b) The phase diagram mapped out by the hypothetical control parameter shown in panel **a**. Points representing the concentrations of coexisting phases are colored according to the equilibrium composition of stoichiometric complexes. The coexistence regions terminate at the points (iv), (v), and (vi), as described in the text. (c)–(g) Representative free-energy profiles (points) along the path taken by the control parameter shown in panel **a**. For comparison, we also show the free-energy profiles of pure chains (100% P, solid blue curves) and pure complexes (100% S, solid green curves). Red dashed lines indicate common-tangent constructions, red points indicate coexisting concentrations, and gray shaded regions indicate coexistence regions. (h) A representative free-energy profile in two-phase region (2) of the master phase diagram, (a).

the boundaries between parameter regimes in the $\beta\epsilon$–$\beta\Delta G$ plane that have differing numbers of coexistence regions indicate lines of either first-order or second-order transitions where coexistence regions terminate. In other words, crossing one of these boundaries on the master phase diagram means that one or more coexistence regions either appear or disappear within some range(s) of total concentration as $\beta\epsilon$ and/or $\beta\Delta G$ are tuned.

Master phase diagrams are useful for understanding how different relationships between the association and assembly free energies can result in qualitatively different phase diagrams when calculated as a function of a single control parameter. For example, consider a biopolymer solution in which temperature controls both $\beta\epsilon$ and $\beta\Delta G$. Depending on the precise chemical details of the biopolymers, changing temperature traces out a curve in the $\beta\epsilon$–$\beta\Delta G$ plane [see, e.g., Figs. 3(a) and 5(a)]. Computing the phase behavior as a function of the total concentration at points along this temperature-parameterized curve therefore allows us to construct a conventional temperature-concentration phase diagram [see, e.g., Figs. 3(b) and 5(b)]. The topology of this conventional phase diagram is determined by the way in which the temperature-parameterized curve intersects the boundaries between parameter regimes with different numbers of coexistence regions in the $\beta\epsilon$–$\beta\Delta G$ plane. Importantly, this approach allows us to disentangle the predictions of the mean-field model, which depend on the free-energy parameters $\beta\epsilon$ and $\beta\Delta G$, from system-specific parametrizations of these quantities.

C. Self-assembly with conserved excluded volume

We first consider the scenario where the self-assembly of n polymers into a stoichiometric complex conserves the

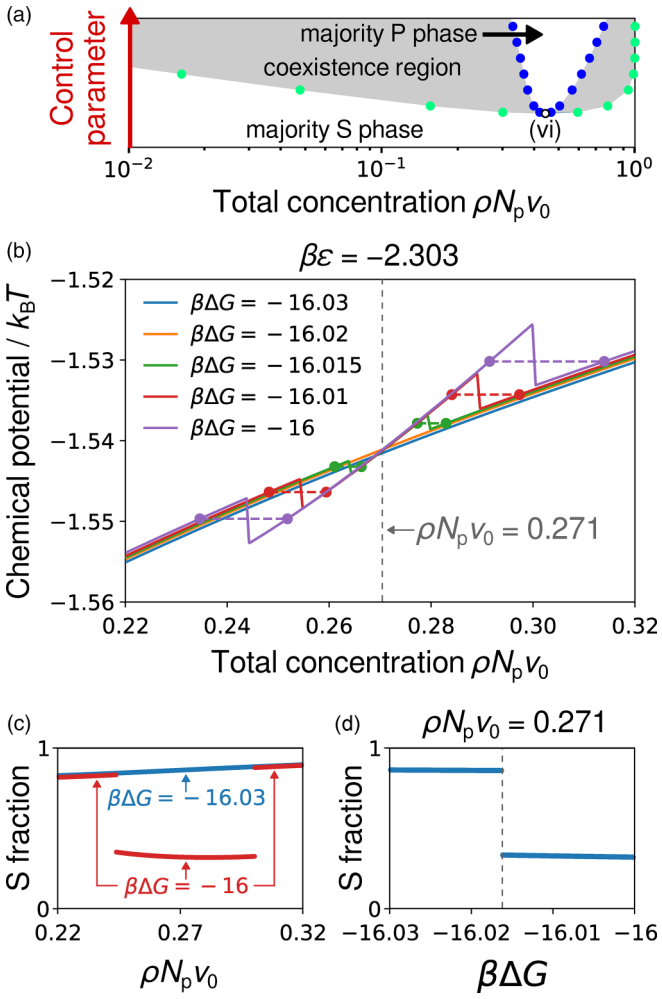

FIG. 4. Coexistence regions terminate at a first-order phase transition in the conserved-excluded-volume scenario. (a) Here we examine the nature of the transition from a one-phase region to two two-phase regions along line (ii) in Fig. 3(a), which corresponds to point (vi) in this zoomed-in portion of the phase-diagram shown in Fig. 3(b). (b) The concentration-dependent chemical potential, μ, at fixed $\beta\epsilon = -2.303$ and variable $\beta\Delta G$. Coexisting phases are indicated by circles and connected by dashed lines. The transition from two coexistence regions ($\beta\Delta G \gtrsim -16.015$) to zero coexistence regions ($\beta\Delta G \lesssim -16.02$) does not involve a critical point, which would imply a concentration where $\partial\mu/\partial\rho = 0$. Instead, the two coexistence regions move towards one another as $\beta\Delta G$ is decreased and vanish at the azeotrope concentration $\rho N_p v_0 \simeq 0.271$ (dashed line), where $\partial\mu/\partial\rho$ jumps discontinuously. (c) When no coexistence regions are present ($\beta\Delta G = -16.03$), the composition is a continuous function of ρ. Yet when two coexistence regions are present ($\beta\Delta G = -16$), the composition changes discontinuously across each coexistence region. (d) The first-order nature of this transition at finite $\beta\Delta G$ (dashed line), corresponding to line (ii) in Fig. 3, is clear from the discontinuity in the composition at the azeotrope concentration. Moving from left to right across this plot is analogous to moving through point (vi) in Fig. 3(b) at constant ρ from below to above.

excluded volume, such that $v_s = nv_p$. Here we compute phase diagrams for an example case with $N_p = 6$ blobs per polymer chain. Conservation of the excluded volume thus implies that $N_s = nN_p = 24$, since we assume that $n = 4$ chains comprise each stoichiometric complex.

The master phase diagram for this scenario is shown in Fig. 3(a). The model predicts five parameter regimes in the $\beta\epsilon$–$\beta\Delta G$ plane: two distinct regimes, (1) and (3), in which only one phase exists at all concentrations; two distinct regimes, (2) and (5), in which a single coexistence region exists over a finite range of concentrations; and one regime, (4), in which two different coexistence regions exist over different concentration ranges. We discuss the physical origin and significance of these various regimes in the remainder of this section.

1. Phase diagram construction as a function of a control parameter

As discussed in Secs. II A and II B, the free-energy parameters $\beta\epsilon$ and $\beta\Delta G$ must be related to one another due to their common dependence on the interactions between pairs of binding sites [Fig. 1(a)]. We therefore construct a conventional concentration-dependent phase diagram by tracing a path through the $\beta\epsilon$–$\beta\Delta G$ plane that depends on a control parameter. For example, if the control parameter is temperature, then the resulting conventional phase diagram depicts the phase behavior in the temperature–concentration plane. We show a possible parametrization of $\beta\epsilon$ and $\beta\Delta G$ in Fig. 3(a), which results in the conventional phase diagram shown in Fig. 3(b). This example curve is chosen because it passes through four of the five regimes on the master phase diagram. Importantly, this example also satisfies the relation $\partial\beta\epsilon/\partial\beta\Delta G > 0$, since both free energies should typically be expected to increase or decrease in tandem. Specific points along the parameterized curve are labeled c–g in Fig. 3(a) and are indicated on the conventional phase diagram in Fig. 3(b).

The appearance and disappearance of various coexistence regions in Fig. 3(b) is determined by the intersections between the control-parameter curve and the boundaries of the distinct regimes in Fig. 3(a). Starting from low values of the control parameter, the polymer solution is initially in a one-phase region at all concentrations at point (c). Upon crossing line (ii) in Fig. 3(a), two distinct coexistence regions appear simultaneously at point (vi) in Fig. 3(b). These coexistence regions imply that a polymer solution at a total concentration within a shaded region will phase separate at equilibrium into two coexisting phases with different concentrations and compositions, as indicated by the equilibrium S fraction shown on the coexistence-region boundaries in Fig. 3(b). Thus, while the number of coexistence regions is the same at points (d) and (e), phase coexistence in the lower-concentration coexistence region is qualitatively different at these two points. Specifically, the dilute phase at point (d) is dominated by assembled stoichiometric complexes, while the dilute phase at point (e) is primarily composed of free polymers. The control-parameter curve then crosses line (iii) in Fig. 3(a), which indicates the termination of the high-concentration coexistence region at point (v) in Fig. 3(b), leaving only the lower-concentration coexistence region at point (f). Finally, the control-parameter curve crosses line (i) in Fig. 3(a), corresponding to point (iv) in Fig. 3(b), so the polymer solution exists in a single phase at all concentrations once again at point (g).

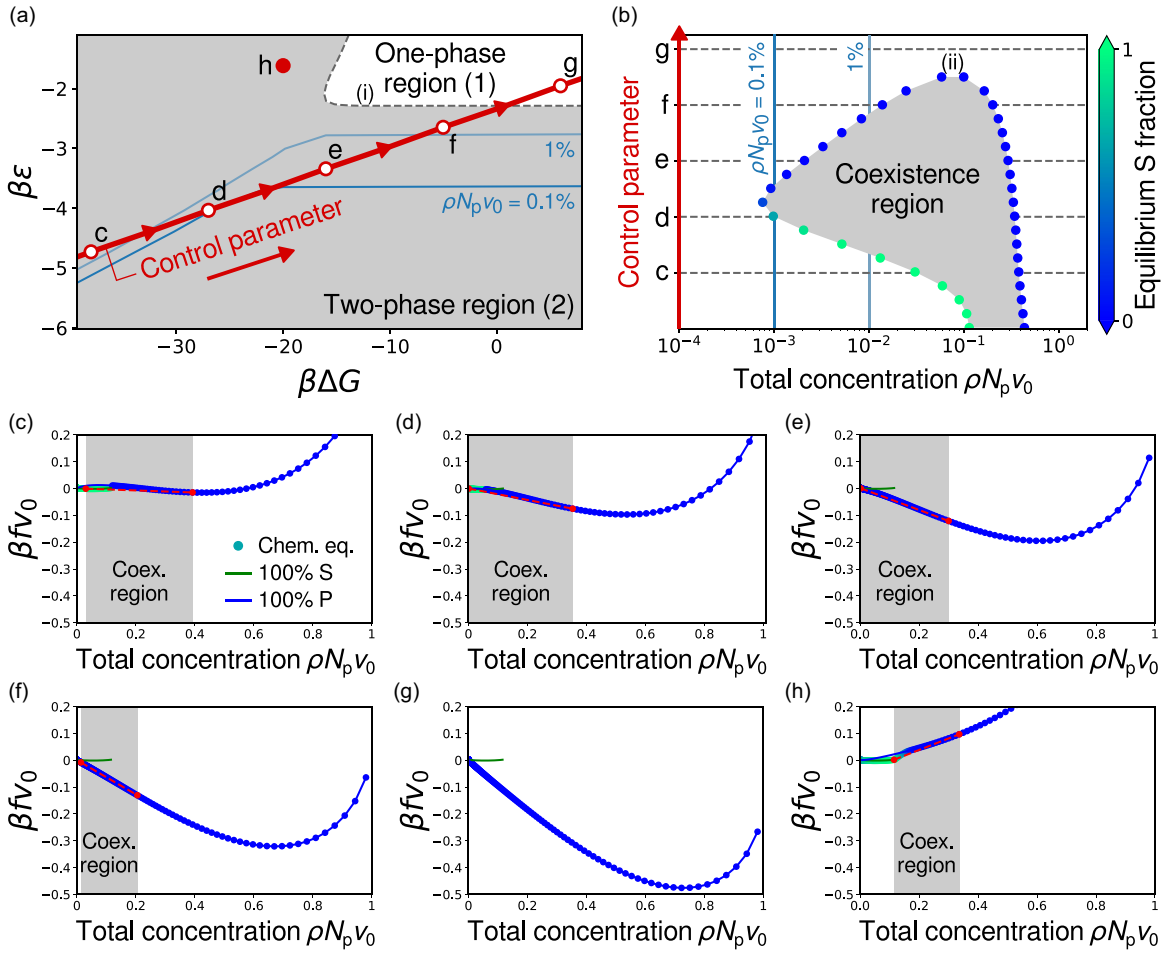

FIG. 5. Phase-behavior predictions when self-assembly does not conserve the excluded volume of the constituent chains. (a) A single one-phase region (unshaded), where the system is homogeneous at all concentrations, lies to the right of dashed curve (i) and above dashed curve (ii), as described in the text. A two-phase region exists within some concentration range everywhere else within the master phase diagram. See Fig. 3(a) for explanations of the control parameter path and total concentration curves. (b) When the excluded volume is not conserved by the two-state model of self-assembly, the lone coexistence region terminates at a single point (iii), as described in the text. This phase diagram is constructed analogously to that shown in Fig. 3(b). (c)–(h) Representative free-energy profiles along the path taken by the control parameter shown in (a), and at the additional labeled point (h) in panel (a). See the description of Figs. 3(c)–3(h) for details. Because of the excluded volume difference, the 100% S curves terminate at the total concentration where the complexes are close-packed, $\rho N_s v_0/n = 1$, where $n = 4$ in this example. Consequently, condensed phases with $\rho N_p v_0 > n N_p/N_s$ are always polymer-dominated.

Both the higher-concentration and the lower-concentration coexistence regions in Fig. 3(b) exhibit reentrance, since a polymer solution at constant total concentration can pass from a single phase to two phases and back to a single phase as the control parameter is varied monotonically from point (c) to point (g). However, reentrance only occurs over a finite range of total concentrations. To determine whether reentrance will occur at a specific total concentration, we also plot curves of constant dilute-phase concentration on the master phase diagram in Fig. 3(a). These curves indicate the conditions in the $\beta\epsilon$–$\beta\Delta G$ plane at which the dilute phase has the indicated total concentration, and thus where the binodal intersects a line of fixed total concentration in a control parameter versus total concentration phase diagram [e.g., Fig. 3(b)]. Because the example control-parameter curve shown in Fig. 3(a) crosses the curve of a 1% dilute-phase concentration twice but does not cross the curve of a 0.1% dilute-phase concentration, we can easily predict that reentrance will occur for $\rho N_p v_0 = 1\%$ but not at $\rho N_p v_0 = 0.1\%$ in Fig. 3(b). These examples demonstrate how the essential features of the conventional phase diagram can be inferred graphically by comparing a control-parameter curve, which is specific to a particular system, with the master phase diagram, which depends only on the parameters N_p, N_s, and n in this model.

2. *Free-energy landscapes*

Representative free-energy landscapes at various points in the $\beta\epsilon$–$\beta\Delta G$ plane, corresponding to each labeled point on the control-parameter curve as well as the additional labeled point (h), are shown in Figs. 3(c)–3(h). Equilibrium landscapes are indicated by points whose coloring reflects the composition, as in Fig. 3(b). Coexistence regions and tie-line constructions are indicated for each equilibrium landscape. For comparison, we also show free-energy landscapes computed assuming a *nonequilibrium* composition of either 100% S or 100% P.

We find that we can predict the qualitative phase behavior at each point along the control-parameter curve in Fig. 3(b) by examining the relationship between these nonequilibrium landscapes, despite the fact that the composition is never exactly 100% S or 100% P due to the entropy of mixing. In Fig. 3(c), the equilibrium landscape is close to the 100% S curve, which is always below the 100% P curve. Thus, the polymer solution primarily consists of assembled stoichiometric complexes in a single phase at all concentrations. In Fig. 3(d), the lower-concentration coexistence region is approximately given by a common-tangent construction between the 100% S curve (on the left) and the 100% P curve (on the right). However, the lower-concentration coexistence region in Fig. 3(e) and the sole coexistence region in Fig. 3(f) are approximately given by a common-tangent construction using the 100% P curve alone, since this curve is nonconvex due to the associative contribution to the free energy, Eq. (4). This nonconvexity in the 100% P curve disappears at high $\beta\epsilon$, resulting in no phase separation in Fig. 3(g). By contrast, the higher-concentration coexistence region in Figs. 3(d) and 3(e) and the sole coexistence region in Fig. 3(h) are approximately given by a common-tangent construction between the 100% P curve (on the left) and the 100% S curve (on the right).

3. *Parameter-regime boundaries on the master phase diagram*

The nature of the boundaries between regimes in the master phase diagram can be understood by examining the behavior of the system at points where the coexistence regions terminate in Fig. 3(b). Point (iv) in Fig. 3(b) is a critical point, since the derivatives $\partial\mu/\partial\rho$ and $\partial^2\mu/\partial\rho^2$ vanish. As noted at the end of Sec. II A, this critical point is typically referred to as an UCST. Line (i) in Fig. 3(a) is therefore a line of critical points that bound the lower-concentration coexistence region. Since the two phases that merge at point (iv) are composed mostly of free polymers, line (i) is nearly independent of $\beta\Delta G$. The value of $\beta\epsilon$ along this line is thus approximately equal to the critical association free energy of an associating polymer solution in the absence of self-assembly (i.e., $\beta\Delta G \to \infty$).

By contrast, point (v) in Fig. 3(b) is a first-order transition, where the composition changes discontinuously from mostly assembled stoichiometric complexes to mostly free polymers upon a change in $\beta\epsilon$ and/or $\beta\Delta G$. This transition occurs when the free energy of a 100% P solution drops below that of a 100% S solution at $\rho N_\mathrm{p} v_0 = 1$, which we can see by comparing the free-energy landscapes shown in Figs. 3(e) and 3(f). We can therefore predict the line of first-order transitions indicated by line (iii) in Fig. 3(a) by equating the free energies of 100% P and 100% S solutions at $\rho N_\mathrm{p} v_0 = 1$. To the left of line (iii), both coexisting phases have total concentrations less than one, as evidenced by the example coexistence region shown in Fig. 3(h). To the right of line (iii), the higher-concentration coexistence region does not exist.

4. *Physical origin of reentrance*

We now focus on the physical origin of the reentrant behavior that occurs at relatively low values of the total concentration in Fig. 3(b). In essence, reentrance manifests because a solution of free polymers phase separates at high values of the control parameter, but a single-phase solution of predominantly assembled stoichiometric complexes becomes stable at low values of the control parameter. Reentrance therefore results from mutually exclusive competition between polymer phase separation and stoichiometric complex self-assembly. More precisely, reentrance occurs when the control-parameter curve crosses a line of constant dilute-phase concentration (e.g., $\rho N_\mathrm{p} v_0 = 1\%$) *twice* in Fig. 3(a). These constant dilute-phase concentration curves are bounded by lines (i) and (ii), so this form of reentrance can only appear in a limited portion of the $\beta\epsilon$–$\beta\Delta G$ plane. While the upper bound of the lower-concentration coexistence region with respect to the control parameter is set by point (iv) in Fig. 3(b), the lower bound on this lower-concentration coexistence region is determined by point (vi). This lower bound corresponds to the intersection between line (ii) and the control-parameter curve in Fig. 3(a).

Yet unlike point (iv) in Fig. 3(b), point (vi) is *not* a critical point. Instead, point (vi) indicates a first-order transition where the composition jumps from a high value to a low value of the equilibrium S fraction as the control parameter is increased. The distinction between the behavior at point (vi) and a LCST can be understood by examining the chemical potential in the vicinity of the transition [Fig. 4(a)]. In Fig. 4(b), we show the chemical potential as a function of the total concentration at fixed $\beta\epsilon$. As we tune $\beta\Delta G$ near the transition point, we see that the two coexistence regions converge at a crossover concentration [vertical dashed line in Fig. 4(b)]. However, the derivative of the chemical potential at the crossover concentration does not vanish, as would be required for a second-order transition. By contrast, the equilibrium S fraction at the crossover concentration changes discontinuously, as shown in Figs. 4(c) and 4(d). Since the composition is determined from a first derivative of the free energy, this behavior is consistent with a first-order transition.

We propose that this behavior can be understood intuitively by drawing an analogy to a positive azeotrope [29]. In a binary mixture of two fluids, a positive azeotrope occurs at a specific composition at which the two fluids share the same boiling point; thus, raising the temperature of the mixture at the azeotrope composition causes the liquid mixture to transform directly into a mixture of gases with the same composition. Qualitatively, the free-energy landscape of a fluid mixture with an azeotrope is analogous to that of the nonequilibrium 100% S and 100% P curves shown in Figs. 3(c) and 3(d), whereby the liquidlike 100% S curve is stable below a transition point ($\beta\Delta G$, $\beta\epsilon$) and the vaporlike 100% P curve is stable above. We therefore propose that the crossover concentration shown in Fig. 4(b) is analogous to an azeotrope composition, even though there is technically only one equilibrium free-energy landscape as a function of the total concentration in our model due to the chemical equilibrium constraint, Eq. (7). Finally, we note that line (i) terminates at line (ii) in Fig. 3(a) for a similar reason: For sufficiently negative $\beta\Delta G$, the 100% S free-energy landscape curve becomes more stable than the 100% P curve at all concentrations. Consequently, the nonconvex portion of the free-energy landscape, which describes a polymer solution composed of mostly free polymers, becomes metastable to the left of line (ii) in Fig. 3(a).

5. Dependence of the master phase diagram on N_p and n

Up to this point, we have discussed the master phase diagram assuming the specific parameter choices $N_p = 6$ and $n = 4$. These calculations can easily be repeated using alternative parameters. In this way, we find that the topology of the master phase diagram remains unchanged as long as $n > 1$ and the excluded volume is conserved, such that $N_s = nN_p$. However, changing the parameters N_p and n alters the positions of the phase boundaries in the $\beta\epsilon$–$\beta\Delta G$ plane, as we discuss next.

The line of critical points (i) is nearly independent of the degree of complexation n because these critical points only depend on the properties of free polymers, except when in close proximity to the intersection of lines (i) and (ii). Line (i) shifts upward in the $\beta\epsilon$–$\beta\Delta G$ plane as the degree of polymerization N_p increases, since the critical per-binding-site association free energy becomes weaker as the number of binding sites per polymer increases. This behavior is analogous to that of the Flory-Huggins homopolymer model, in which the critical interaction strength weakens as the degree of polymerization increases [22].

By contrast, the two first-order phase boundaries, lines (ii) and (iii), depend on both n and N_p. The slope of line (iii) is exactly proportional to nN_p; this can be shown by equating the free energies of the 100% P and 100% S solutions at $\rho N_p v_0 = 1$. The slope of line (ii), as well as the difference in $\beta\Delta G$ between lines (ii) and (iii) at constant $\beta\epsilon$, also scales roughly linearly with nN_p. This dependence can be understood by considering the transition between regions (1) and (4) in Fig. 3(a) in the strong association limit, $\beta\epsilon \to -\infty$ [20]. In this limit, the relative stabilities of the 100% P and 100% S solutions are primarily determined by the balance between the association free energy in the P state, in which nearly all binding sites are associated since $X_p \to 0$, and the internal contribution to the free energy in the S state, which is nearly proportional to $\beta\Delta G$. Consequently, $\beta\Delta G$ must scale with the total number of associative interactions in which the complex-forming polymers could potentially engage, which is equal to nN_p. Thus, nN_p sets the scale for both lines (ii) and (iii). Increasing nN_p tends to stretch the master phase diagram in the horizontal direction.

D. Self-assembly with nonconserved excluded volume

Next, we consider the alternative case in which self-assembly does not conserve the excluded volume. This scenario can be motivated by considering the second-virial coefficient between two assembled stoichiometric complexes in solution, which is given by $B_{ss} = N_s v_0/2$ in our model (see Appendix B). This second-virial coefficient represents the total volume (divided by two) that the center of mass of one stoichiometric complex cannot access due to the presence of the other complex. If the polymers self-assemble into a structure that has a lower density than a polymer melt, then it is possible to have $v_s = 2B_{ss} > nv_p$. For example, this scenario is relevant in the case of nucleic-acid complexes, because secondary-structure formation increases the persistence length of duplexes relative to free strands [30,31]; nucleic-acid complexes can therefore have a relatively low density of polymer within a relatively large pervaded volume. We therefore consider an example scenario in which $N_s = 200 > nN_p = 24$ in this section.

1. Similarities and differences with the conserved-excluded-volume scenario

The master phase diagram for this scenario is shown in Fig. 5(a). The control-parameter curve, which is the same as that in Fig. 3(a), maps out the conventional concentration-dependent phase diagram shown in Fig. 5(b). In contrast with the conserved-excluded-volume scenario, this master phase diagram only exhibits two parameter regimes: one regime, (1), in which the solution exists in a single phase at all concentrations; and one regime, (2), with a single coexistence region. Yet, despite these drastic changes in the topology of the master phase diagram relative to the conserved-excluded-volume scenario, the coexistence region in Fig. 5(b) strongly resembles the lower-concentration coexistence regime in Fig. 3(b) for values of the control parameter above point (c).

These differences can be understood by noting that the increased excluded volume of the stoichiometric complex means that chains must exist in the free-polymer state at high total concentrations. More precisely, since the volume fraction ϕ_s cannot exceed one, the maximum total concentration at which stoichiometric complexes can exist is given by $\rho N_p v_0 = nN_p/N_s$. This upper limit on the concentration of stoichiometric complexes is reflected in the free-energy landscapes shown in Figs. 5(c)–5(h), where the 100% S landscapes do not extend over the full range of total concentrations. Consequently, there is no analog of a higher-concentration coexistence region, as observed in the conserved-excluded-volume case, when we consider the nonconserved scenario. However, the phase behavior at low total concentrations, $\rho N_p v_0 \ll nN_p/N_s$, is qualitatively the same in both excluded-volume scenarios. We also note that the boundary between the two regimes, line (i) in Fig. 5(a), is a single, continuous line of critical points. This critical line is the same as line (i) in the conserved-excluded-volume case at large $\beta\Delta G$, but turns into a continuous transition between assembled stoichiometric complexes and free polymers when it bends upward near point (h) in Fig. 5(a).

2. Physical origin of re-entrance

Reentrance occurs in the nonconserved-excluded-volume scenario due to the same competition between self-assembly and polymer phase separation discussed in the context of the conserved-excluded-volume scenario. Minor quantitative differences, such as the maximum total concentration at which reentrance occurs, appear between the two scenarios due to the differences in the excluded-volume and internal contributions to the free energy, f^{ev} and f^{int}, in Eqs. (3) and (6). However, the absence of a one-phase parameter regime at low $\beta\Delta G$ in the nonconserved case means that the coexistence region does not terminate at low values of the control parameter in Fig. 5(b). Instead, the maximum concentration at which re-entrance can occur tends to $\phi_s = 1$ in the limit $\beta\Delta G \to -\infty$. In this limit, phase coexistence occurs between a lower-concentration phase composed entirely of assembled stoichiometric complexes and a higher-concentration phase of free polymers, regardless of the value of $\beta\epsilon$.

III. DISCUSSION

We have demonstrated how competition between self-assembly and liquid–liquid phase separation can emerge from the same set of molecular interactions among biopolymers in solution. Our theoretical approach captures the many-body interactions that are inherent to this competition by coarse-graining the conformational states of polymers into either free chains and assembled stoichiometric complexes. The resulting mean-field model predicts that reentrance can emerge from the competition between self-assembly and phase separation, both of which involve only attractive associative interactions. Our results provide insight into the key features that control the phase behavior of these systems, providing a unified view of different phase diagrams that can result given a parametric dependence on temperature, ionic strength, pH, or other experimentally controllable variables.

Importantly, we find that reentrance in this model is not a consequence of a LCST, despite qualitative similarities between the phase diagrams predicted by our model and those involving UCST+LCST miscibility loops [22,32]. In the conserved-excluded-volume scenario, the lower termination of the coexistence regions occurs at a first-order transition, which is qualitatively different from an LCST. By contrast, when self-assembly does not conserve excluded volume, the coexistence region does not terminate at all in the limit of stable stoichiometric complexes. Nonetheless, the physical origin of reentrance is the same in both cases, and the phase behavior at low polymer concentrations is essentially unchanged between these two scenarios. We therefore believe that the qualitative predictions of our model—particularly the absence of an LCST—are robust despite the approximations invoked in our mean-field treatment. We note that analogous behavior has also been observed in conceptually similar mean-field models of monomeric as opposed to polymeric solutions, in which phase separation either competes with self-assembly [17] or a unimolecular chemical reaction [33]. In the case of unimolecular chemical reactions [33], however, the low-temperature terminus of the coexistence region [i.e., point (vi) in Fig. 3(b)] always occurs at $\phi = 1$.

A key approximation in this article is the use of a two-state model of self-assembly. In principle, the accuracy of our model could be improved by accounting for additional intermediate states between completely dissociated polymers and completely assembled stoichiometric complexes. Our model also assumes that the excluded volume of the stoichiometric complexes, v_s, is a constant. A more versatile model could account for variations in v_s as a function of concentration, which would provide a better description of complexes that deform at high concentrations in the nonconserved-excluded-volume scenario. However, we do not believe that these modifications would affect our qualitative conclusions regarding the nature of the various phase transitions. In particular, the first-order transitions that we observe emerge, fundamentally, from the cooperative nature of the self-assembly process, which implies the existence of a free-energy barrier between the free-polymer (P) and assembled stoichiometric-complex (S) states. We therefore believe that the phase behavior will remain qualitatively unchanged as long as this essential feature is preserved. For the same reason, we believe that this model can apply to a wide variety of self-assembling polymer solutions in which self-assembly is cooperative and mutually exclusive with phase separation. Our approach would be less appropriate for modeling the phase behavior of polymer solutions in which less cooperative intramolecular bonding competes with associative polymer phase separation, as is the case in many experiments on nucleic-acid repeat sequences [34].

In conclusion, we have introduced and analyzed a broadly applicable mean-field model of biopolymer assembly and phase separation. Our master-phase-diagram approach makes it possible to derive system-specific phase diagrams in a unified and intuitive fashion. It is also straightforward to extend our theoretical framework to model more complex scenarios involving hierarchical assembly and phase separation, such as polymer solutions in which self-assembled complexes can themselves undergo phase separation via associative interactions [20]. Most importantly, the simple manner in which our model can be parameterized makes this framework useful for generating experimentally testable predictions for complex biopolymer solutions.

ACKNOWLEDGMENTS

This paper was supported by a grant from the National Science Foundation (No. DMR-2143670) to W.M.J., as well as support from the Human Frontier Science Program (No. RGP0029) and the Smith Family Foundation to W.B.R. Source code is provided at Ref. [35].

APPENDIX A: EXCLUDED VOLUME FORMULAE

To derive an expression for the excluded-volume contribution to the mixing free energy, we consider a cubic lattice with N lattice sites. Each lattice site has the same volume, v_0, as a monomer (i.e., a blob) of the polymer, P. Each polymer chain occupies N_p lattice sites. Each stoichiometric complex, referred to as an S molecule in this section, occupies an excluded volume of $v_s = N_s v_0$. We compute the free energy of mixing, in the absence of associative interactions or chemical reactions, in a two-step manner (Fig. 6). An empty lattice, with a free energy equal to zero, is chosen as the reference state. At each step, the free-energy change is purely entropic

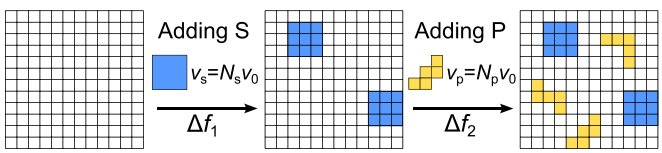

FIG. 6. Two-step calculation of the excluded-volume contribution to the free energy. Starting from the reference state of an empty lattice (state 0, left), stoichiometric complexes (S) are added in step one. Stoichiometric complexes occupy N_s *contiguous* lattice sites. The free-energy change associated with this process is Δf_1, resulting in the intermediate configuration (state 1, middle). In step two, polymers (P) are added to reach the final configuration (state 2, right). Polymers occupy N_p *noncontiguous* lattice sites. The free-energy change associated with this second process is Δf_2.

due to the increase in the number of microstates, $\Omega_{i-1} \to \Omega_i$,

$$\beta N \Delta f_i v_0 = -\frac{\Delta S_i}{k_B} = -\ln\left(\frac{\Omega_i}{\Omega_{i-1}}\right), \tag{A1}$$

where ΔS_i is the entropy change associated with step i from state $i-1$ to state i.

First, we consider the free-energy change due to the addition of S molecules occupying a volume fraction ϕ_s. Importantly, each S molecule is considered to be a compact, rigid unit, and must therefore occupy N_s *contiguous* lattice sites. We therefore compute the approximate entropy change associated with step 1 (Fig. 6) by coarse graining the lattice into super sites of volume $v_0 N_s$, and then applying the standard regular-solution formula [27]:

$$\beta \Delta f_1 v_0 = \frac{\phi_s}{N_s}\ln\phi_s + \frac{1-\phi_s}{N_s}\ln(1-\phi_s). \tag{A2}$$

Second, we add P molecules occupying a volume fraction ϕ_p into the system. Following the Flory–Huggins derivation [22], we ignore chain connectivity when assigning polymer blobs to *noncontiguous* unoccupied lattice sites. This leads to the mean-field result

$$\beta \Delta f_2 v_0 = \frac{\phi_p}{N_p}\ln\left(\frac{\phi_p}{1-\phi_s}\right) + \phi_0 \ln\left(\frac{\phi_0}{1-\phi_s}\right), \tag{A3}$$

where $\phi_0 \equiv 1-\phi_s-\phi_p$ represents the volume fraction not occupied by either free polymers or stoichiometric complexes. Summing the free-energy changes of the two steps, we find that the total free-energy difference is

$$\begin{aligned}\beta(f^{\rm id}+f^{\rm ev})v_0 = {}& \frac{\phi_s}{N_s}\ln\phi_s + \frac{1-\phi_s}{N_s}\ln(1-\phi_s) \\ & + \frac{\phi_p}{N_p}\ln\left(\frac{\phi_p}{1-\phi_s}\right) + \phi_0\ln\left(\frac{\phi_0}{1-\phi_s}\right),\end{aligned} \tag{A4}$$

which accounts for both the ideal-gas and the excluded-volume contributions to the free energy. Taking partial derivatives of Eq. (A4) with respect to ρ_p and ρ_s, and subtracting the ideal chemical potentials $\beta^{-1}\ln\phi_p$ and $\beta^{-1}\ln\phi_s$, gives the excluded-volume contributions to the excess chemical potentials:

$$\begin{aligned}\beta\mu_s^{\rm ev} = {}& -\ln(1-\phi_s) + N_s\ln\left(\frac{1-\phi_s}{\phi_0}\right) \\ & - N_s\left(1-\frac{1}{N_p}\right)\frac{\phi_p}{1-\phi_s},\end{aligned} \tag{A5}$$

$$\beta\mu_p^{\rm ev} = 1 - N_p + (N_p-1)\ln(1-\phi_s) - N_p\ln\phi_0. \tag{A6}$$

Both $\beta\mu_s^{\rm ev}$ and $\beta\mu_p^{\rm ev}$ diverge as $\phi_0 \to 0$. Since $\phi_s+\phi_p \leqslant 1$, this means that both $\beta\mu_s^{\rm ev}$ and $\beta\mu_p^{\rm ev}$ diverge if $\phi_s \to 1$ or $\phi_p \to 1$.

APPENDIX B: COMPARISON WITH THE FLORY–HUGGINS POLYMER MODEL

The excluded-volume contributions embodied in Eqs. (A4)–(A6) differ from those implied by the standard Flory–Huggins free-energy density [22],

$$\beta f^{\rm FH} v_0 = \frac{\phi_s}{N_s}\ln\phi_s + \frac{\phi_p}{N_p}\ln\phi_p + \phi_0\ln\phi_0, \tag{B1}$$

which treats both the S and P species as mean-field polymers. The excluded-volume contribution to the chemical potential of species $i \in \{s, p\}$ is

$$\beta\mu_i^{\rm FH,ev} = 1 - N_i - N_i\ln\phi_0 \tag{B2}$$

in the Flory–Huggins model. Focusing on the case of the S molecule, Eq. (B2) highlights the problem with applying the Flory–Huggins model to our system: The chemical potential of the S molecule depends on N_s regardless of the polymer concentration. As a result, the Flory–Huggins equation of state of a pure solution of S molecules depends on the ratio of the excluded volume of a rigid S molecule to the excluded volume of a polymer blob, even when no polymer is present in the system. This is an undesirable feature, which necessitates the use of the formulas derived in Appendix A.

It is also instructive to compare the second-virial coefficients, B_{ss}, B_{sp}, and B_{pp}, of the two models. To this end, we expand the pressure, $P = -f + \mu_s\rho_s + \mu_p\rho_p$, to second order in the concentrations ρ_s and ρ_p, such that $P = \sum_{i\in\{s,p\}}\rho_i + \sum_{i,j\in\{s,p\}} B_{ij}\rho_i\rho_j + O(\rho^3)$. For the Flory–Huggins model, this yields the coefficients

$$B_{ss}^{\rm FH} = \frac{N_s^2 v_0}{2}, \quad B_{sp}^{\rm FH} = \frac{N_s N_p v_0}{2}, \quad B_{pp}^{\rm FH} = \frac{N_p^2 v_0}{2}, \tag{B3}$$

whereas the coefficients derived from our excluded-volume model, Eqs. (A6) and (A5), are

$$B_{ss}^{\rm ev} = \frac{N_s v_0}{2}, \quad B_{sp}^{\rm ev} = \frac{N_s v_0}{2}, \quad B_{pp}^{\rm ev} = \frac{N_p^2 v_0}{2}. \tag{B4}$$

The B_{pp} coefficients are the same in both expressions as expected, since the polymers are treated equivalently in the two models. However, the second-virial coefficients involving the S molecules differ. $B_{ss}^{\rm ev}$ is a more physically reasonable result than $B_{ss}^{\rm FH}$, since the S molecules are assumed to be compact, rigid units. In fact, $B_{ss}^{\rm ev}$ is the correct result for monodisperse hard particles (such as hard spheres [36]) that exclude a volume of $N_s v_0$. The cross term, $B_{sp}^{\rm ev}$, is less obvious, but we can build intuition by considering the insertion of polymers into a low-density gas of S molecules. For simplicity, let us assume that the S molecules are hard spheres with diameter $D \simeq N_s^{1/3} v_0^{1/3}$. The average end-to-end distance of the freely jointed polymer chains is $d \simeq v_0^{1/3} N_p^{1/2}$. A simple approximation of the cross term is therefore $B_{sp}^{\rm ev} \approx (D+d)^3/2 = N_s(1+r)^3 v_0/2$, where $r = N_p^{1/2}/N_s^{1/3}$. Our excluded-volume model of S/P mixtures is therefore most accurate when the S molecules are large compared to the individual polymer chains (such as in the case considered in Sec. II D), while the standard Flory-Huggins model is more appropriate when the size of an S molecule is comparable to that of an individual polymer blob. We therefore conclude that the excluded-volume formulas Eqs. (A6) and (A5) are the more appropriate choice for the model and parameter regimes considered in this article.

[1] Y. Shin and C. P. Brangwynne, Liquid phase condensation in cell physiology and disease, Science **357**, eaaf4382 (2017).

[2] R. V. Pappu, S. R. Cohen, F. Dar, M. Farag, and M. Kar, Phase transitions of associative biomacromolecules, Chem. Rev. **123**, 8945 (2023).

[3] P. Li, S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani *et al.*, Phase transitions in the assembly of multivalent signalling proteins, Nature (London) **483**, 336 (2012).

[4] C. Lan, J. Kim, S. Ulferts, F. Aprile-Garcia, A. Anandamurugan, R. Grosse, R. Sawarkar, A. Reinhardt, and T. Hugel, Quantitative real-time in-cell imaging reveals heterogeneous clusters of proteins prior to condensation, Nat. Commun. **14**, 4831 (2023).

[5] A. Jain and R. D. Vale, RNA phase transitions in repeat expansion disorders, Nature (London) **546**, 243 (2017).

[6] H. T. Nguyen, N. Hori, and D. Thirumalai, Condensates in RNA repeat sequences are heterogeneously organized and exhibit reptation dynamics, Nat. Chem. **14**, 775 (2022).

[7] M. R. Marzahn, S. Marada, J. Lee, A. Nourse, S. Kenrick, H. Zhao, G. Ben-Nissan, R.-M. Kolaitis, J. L. Peters, S. Pounds *et al.*, Higher-order oligomerization promotes localization of SPOP to liquid nuclear speckles, EMBO J. **35**, 1254 (2016).

[8] A. E. Conicella, G. L. Dignon, G. H. Zerze, H. B. Schmidt, A. M. DOrdine, Y. C. Kim, R. Rohatgi, Y. M. Ayala, J. Mittal, and N. L. Fawzi, TDP-43 α-helical structure tunes liquid–liquid phase separation and function, Proc. Natl. Acad. Sci. USA **117**, 5883 (2020).

[9] M. Kar, F. Dar, T. Welsh, L. Vogel, R. Kuhnemuth, A. Majumdar, G. Krainer, T. Franzmann, S. Alberti, C. Seidel *et al.*, Phase separating RNA binding proteins form heterogeneous distributions of clusters in subsaturated solutions, Proc. Natl. Acad. Sci. USA **119**, e2202222119 (2022).

[10] G. He, T. GrandPre, H. Wilson, Y. Zhang, M. C. Jonikas, N. S. Wingreen, and Q. Wang, Phase-separating pyrenoid proteins form complexes in the dilute phase, Commun. Biol. **6**, 19 (2023).

[11] C. Weber, T. Michaels, and L. Mahadevan, Spatial control of irreversible protein aggregation, Elife **8**, e42315 (2019).

[12] I. Seim, A. E. Posey, W. T. Snead, B. M. Stormo, D. Klotsa, R. V. Pappu, and A. S. Gladfelter, Dilute phase oligomerization can oppose phase separation and modulate material properties of a ribonucleoprotein condensate, Proc. Natl. Acad. Sci. USA **119**, e2120799119 (2022).

[13] J. A. Marsh and S. A. Teichmann, Structure, dynamics, assembly, and evolution of protein complexes, Annu. Rev. Biochem. **84**, 551 (2015).

[14] M. S. Wertheim, Fluids with highly directional attractive forces. I. Statistical thermodynamics, J. Stat. Phys. **35**, 19 (1984).

[15] W. M. Jacobs, D. W. Oxtoby, and D. Frenkel, Phase separation in solutions with specific and nonspecific interactions, J. Chem. Phys. **140**, 204109 (2014).

[16] W. Chapman, K. Gubbins, G. Jackson, and M. Radosz, SAFT: Equation-of-state solution model for associating fluids, Fluid Phase Equilib. **52**, 31 (1989).

[17] A. Reinhardt, A. J. Williamson, J. P. Doye, J. Carrete, L. M. Varela, and A. A. Louis, Re-entrant phase behavior for systems with competition between phase separation and self-assembly, J. Chem. Phys. **134**, 104905 (2011).

[18] J. Bauermann, S. Laha, P. M. McCall, J. Frank, and C. A. Weber, Chemical kinetics and mass action in coexisting phases, J. Am. Chem. Soc. **144**, 19294 (2022).

[19] A. N. Semenov and M. Rubinstein, Thermoreversible gelation in solutions of associative polymers. 1. Statics, Macromolecules **31**, 1373 (1998).

[20] O. Hegde, T. Li, A. Sharma, M. Borja, W. M. Jacobs, and W. B. Rogers, Competition between self-assembly and phase separation induces re-entrant condensation of DNA liquids, arXiv:2301.06134.

[21] J. SantaLucia, Jr. and D. Hicks, The thermodynamics of DNA structural motifs, Annu. Rev. Biophys. Biomol. Struct. **33**, 415 (2004).

[22] M. Rubinstein and R. H. Colby, *Polymer Physics* (Oxford University Press, New York, 2003).

[23] C. P. Brangwynne, P. Tompa, and R. V. Pappu, Polymer physics of intracellular phase transitions, Nat. Phys. **11**, 899 (2015).

[24] S. Biffi, R. Cerbino, F. Bomboi, E. M. Paraboschi, R. Asselta, F. Sciortino, and T. Bellini, Phase behavior and critical activated dynamics of limited-valence DNA nanostars, Proc. Natl. Acad. Sci. USA **110**, 15633 (2013).

[25] S. Biffi, R. Cerbino, G. Nava, F. Bomboi, F. Sciortino, and T. Bellini, Equilibrium gels of low-valence DNA nanostars: A colloidal model for strong glass formers, Soft Matter **11**, 3132 (2015).

[26] M. L. Michelsen and E. M. Hendriks, Physical properties from association models, Fluid Phase Equilib. **180**, 165 (2001).

[27] D. A. Porter, K. E. Easterling, and M. Y. Sherif, *Phase Transformations in Metals and Alloys* (CRC Press, Boca Raton, 2021).

[28] J. Wolff, C. M. Marques, and F. Thalmann, Thermodynamic approach to phase coexistence in ternary phospholipid-cholesterol mixtures, Phys. Rev. Lett. **106**, 128104 (2011).

[29] J. Shephard, S. Callear, S. Imberti, J. Evans, and C. Salzmann, Microstructures of negative and positive azeotropes, Phys. Chem. Chem. Phys. **18**, 19227 (2016).

[30] L. Rovigatti, F. Bomboi, and F. Sciortino, Accurate phase diagram of tetravalent DNA nanostars, J. Chem. Phys. **140**, 154903 (2014).

[31] N. Conrad, G. Chang, D. K. Fygenson, and O. A. Saleh, Emulsion imaging of a DNA nanostar condensate phase diagram reveals valence and electrostatic effects, J. Chem. Phys. **157**, 234203 (2022).

[32] L. Rovigatti, J. M. Tavares, and F. Sciortino, Self-assembly in chains, rings, and branches: A single component system with two critical points, Phys. Rev. Lett. **111**, 168302 (2013).

[33] G. Bartolucci, O. Adame-Arana, X. Zhao, and C. A. Weber, Controlling composition of coexisting phases via molecular transitions, Biophys. J. **120**, 4682 (2021).

[34] O. Kimchi, E. M. King, and M. P. Brenner, Uncovering the mechanism for aggregation in repeat expanded RNA reveals a reentrant transition, Nat. Commun. **14**, 332 (2023).

[35] `https://github.com/wmjac/polymer-phase-sep-self-assembly`.

[36] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids: With Applications to Soft Matter* (Academic Press, Oxford, 2013).