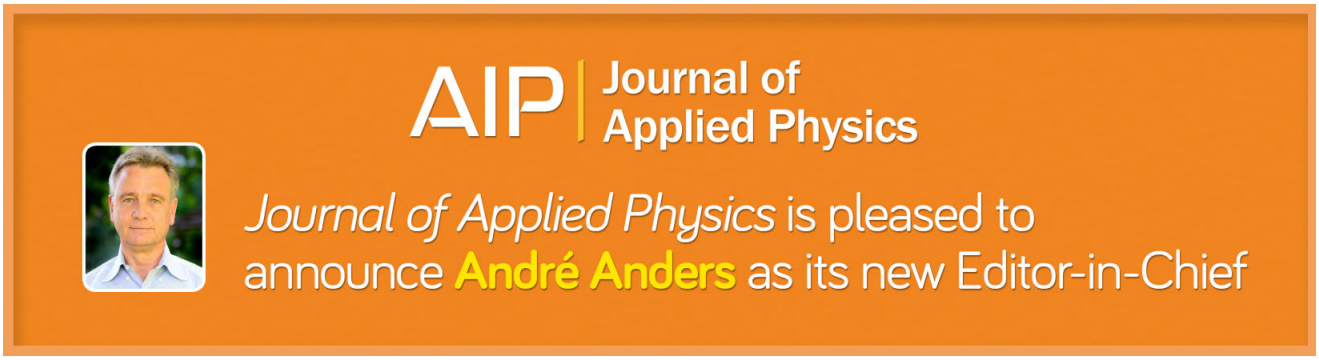
Phase separation in solutions with specific and nonspecific interactions

William M. Jacobs, David W. Oxtoby, and Daan Frenkel

Citation: The Journal of Chemical Physics **140**, 204109 (2014); doi: 10.1063/1.4878836
View online: http://dx.doi.org/10.1063/1.4878836
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/140/20?ver=pdfcov
Published by the AIP Publishing



Phase separation in solutions with specific and nonspecific interactions

William M. Jacobs,[1] David W. Oxtoby,[2] and Daan Frenkel[1]
[1]*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*
[2]*Pomona College, 550 North College Avenue, Claremont, California 91711, USA*

(Received 25 March 2014; accepted 8 May 2014; published online 28 May 2014)

Protein solutions, which tend to be thermodynamically stable under physiological conditions, can demix into protein-enriched and protein-depleted phases when stressed. Using a lattice-gas model of proteins with both isotropic and specific, directional interactions, we calculate the critical conditions for phase separation for model proteins with up to four patches via Monte Carlo simulations and statistical associating fluid theory. Given a fixed specific interaction strength, the critical value of the isotropic energy, which accounts for dispersion forces and nonspecific interactions, measures the stability of the solution with respect to nonspecific interactions. Phase separation is suppressed by the formation of protein complexes, which effectively passivate the strongly associating sites on the monomers. Nevertheless, we find that protein models with three or more patches can form extended aggregates that phase separate despite the assembly of passivated complexes, even in the absence of nonspecific interactions. We present a unified view of the critical behavior of model fluids with anisotropic interactions, and we discuss the implications of these results for the thermodynamic stability of protein solutions. © *2014 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4878836]

I. INTRODUCTION

Naturally occurring protein solutions rely on specific protein-protein interactions for biological functionality. In order to maintain solubility and to prevent potentially detrimental aggregation, the structures of globular proteins are well optimized against nonfunctional interactions.[1] Yet under deviations from physiological solution conditions, it is possible to observe phase separation driven primarily by nonspecific interactions.[2–5] Recent experimental results indicate that phase separation does occur *in vivo*[6–9] and, in some cases, is an essential mechanism for protein functionality.[10–12] It is thus natural to ask, how sensitive are protein solutions to nonspecific interactions?

The nonspecific forces that mediate protein interactions are on average repulsive due to electrostatic repulsion and the predominant hydrophilicity of solvent-facing residues on the surface of globular proteins. Nevertheless, experimental measurements of the Hamaker constant of many globular proteins suggest that dispersion forces can give rise to short-range attractive interactions with sufficient strength to bring about phase separation.[13,14] Nonspecific hydrophobic interactions, particularly among proteins in non-native configurations, can also contribute to an attractive free energy at small intermolecular separation. Even if the majority of nonspecific interactions are repulsive, strong attractions among a dangerous subset of components in a multicomponent environment may lead to demixing phase transitions.[15–17] In this work, we restrict our attention to the simplest nontrivial case of a single-component solution in which nonspecific interactions are assumed to be orientationally averaged and always attractive.

In a one-component colloidal solution with isotropic interactions, the equilibrium thermodynamics are completely described by two fields: the reduced isotropic energy, $\beta\epsilon$, and the reduced chemical potential, $\beta\mu$, where $\beta \equiv (k_B T)^{-1}$ is the inverse temperature. The isotropic energy models the effect on nonspecific interactions, which may be tuned through the control of temperature or the solution salt concentration.[14,18] Liquid-vapor phase separation occurs on a coexistence curve, which terminates at a critical point located at $(\beta\epsilon_c, \beta\mu_c)$. With isotropic interactions weaker than $|\beta\epsilon_c|$, the colloidal fluid is supercritical and thus stable with respect to phase separation regardless of the chemical potential.

To model an associating fluid, we consider anisotropic interactions mediated by specific binding sites superimposed on a colloidal model of a globular protein. We must therefore add a third independent thermodynamic field describing the strength of the associative interactions, $\beta\epsilon_A$. The liquid-vapor coexistence curve generalizes to a surface in the three-dimensional field space, and the stable, homogeneous fluid is bounded by a continuous line of critical points. By calculating this critical line in the $\beta\epsilon$–$\beta\epsilon_A$ plane, we can determine how the placement of associative sites and the strength of specific interactions affect the sensitivity of a model protein solution to nonspecific interactions.

In this paper, we explore the generic features of the phase behavior of model proteins with both specific and nonspecific interactions. In Sec. II, we introduce a family of lattice models that comprise a representative set of colloidal particles with anisotropic interactions. Then in Sections III and IV, we describe an analytical theory incorporating a thermodynamic perturbation theory of associating fluids and a Monte Carlo simulation method based on finite-size scaling theory for locating critical points. The phase diagrams for the family of lattice models are presented and analyzed in Sec. V. Finally, we discuss the implications of the results of these simple models for real protein solutions in Sec. VI.

II. LATTICE MODELS

We model a solution of globular proteins as a lattice gas on a three-dimensional cubic lattice. Occupied lattice sites, which can accommodate at most one particle, represent proteins, while vacancies on the lattice represent an implicit solvent. Each lattice particle also possesses an independent orientation vector that indicates one of the 24 possible orientations of a cubic particle on the lattice. The m association sites of each lattice protein model considered in this work are fixed relative to this orientation vector as shown in Figure 1. Due to the high symmetry of the lattice, there is a small number of unique particle geometries for modeling anisotropic globular proteins with between one and four associative patches.

We consider attractive, short-range potentials by only counting interactions between occupied nearest-neighbor lattice sites. All nearest-neighbor particle pairs, regardless of the orientations of the particles, interact with a reduced isotropic energy of magnitude $\beta\epsilon$. Associative interactions, which depend on the orientations of the particles at nearest-neighbor lattice sites, contribute an additional associative energy of magnitude $\beta\epsilon_{\mathrm{A}}$. In this coarse-grained model, both $-\beta\epsilon$ and $-\beta\epsilon_{\mathrm{A}}$ represent effective free energies of interaction.

In the simplest realization of these anisotropic lattice particles, two particles are said to associate whenever their patches are adjacent. Variations on these models can be constructed by placing additional requirements of the particle orientations. In models (b) and (e), we implement a dihedral constraint by requiring that the polarities of two adjacent patches align; this constraint reduces the number of associating configurations for a pair of patches by a factor of four. We also consider a patch-antipatch constraint, in which both patches on a two-patch particle are uniquely labeled, in model (d). The patch-antipatch condition only allows for association between differently labeled patches.

The placement of associating patches, as well as their polarity in the case of dihedral constraints, controls the topologies of specifically associated clusters that can be formed. For instance, in the two-patch models, models (c) and (d) can form both linear and ring polymers in the absence of further constraints, while the largest oligomer that the dihedrally constrained model (e) can form is a tetramer. Among the three-patch models, model (g), with mutually adjacent patches, is capable of associating all patches simultaneously in an octamer, while model (h) can only form branched polymers that have at least three unassociated patches. This sensitivity to the patch orientations suggests that, within the symmetry group imposed by the cubic lattice, these models are a realistic coarse-graining of globular proteins, which are known to form oligomeric complexes with specific geometries.

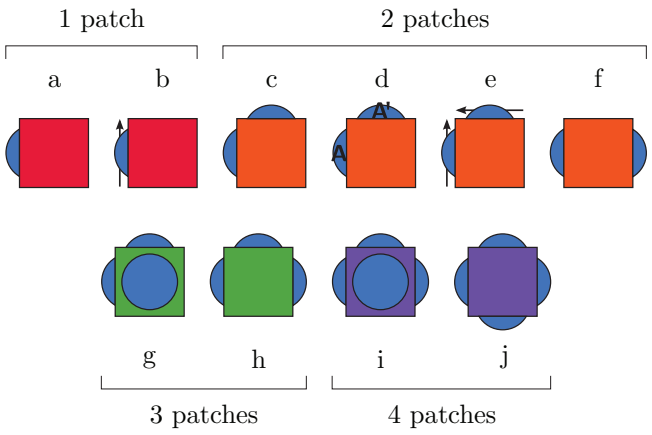


FIG. 1. The three-dimensional patchy protein models studied in this work, with patch locations indicated by blue spherical caps. This set of models includes all unique particle geometries on a cubic lattice with up to four patches. In addition, we have considered models with further requirements for specific interactions: models (b) and (e) implement a dihedral constraint, in which only one of four possible relative nearest-neighbor orientations results in specific association; while model (d) requires complementary patch-antipatch interactions, for which only **A**-**A′** contacts result in specific association.

III. STATISTICAL ASSOCIATING FLUID THEORY

We construct an analytic approximation for the free energy of our protein models by developing a variant of statistical associating fluid theory (SAFT)[19,20] adapted for lattice calculations. In its original form, SAFT utilizes Wertheim's thermodynamic perturbation theory[21,22] to compute the first-order correction to the Helmholtz free energy of a reference fluid due to intermolecular association. At this level of approximation, the number of associative patches and the steric constraint that each patch associate with only one partner are explicitly considered, while the relative orientations of the patches on each monomer are not. As a result, we do not expect the results of the SAFT calculations to distinguish between multiply connected oligomers, such as a tetramer with four specific interactions, and linear or tree-like polymeric configurations.

We take as our reference state an ideal lattice gas of N indistinguishable monomers occupying L^3 lattice sites. The stationary point on the free-energy surface describing associative interactions imposes a chemical equilibrium between bound and unbound patches[23] that depends on the monomer number density, $\rho \equiv NL^{-3}$, and the free energy of association, $-\beta\epsilon_{\mathrm{A}}$:

$$\frac{p_{\mathrm{bound}}}{p_{\mathrm{unbound}}^2} = \frac{m\rho}{M}(e^{\beta\epsilon_{\mathrm{A}}} - 1), \tag{1}$$

where M is the coordination number of the lattice and $p_{\mathrm{bound}} + p_{\mathrm{unbound}} = 1$. For simplicity, we model the isotropic interactions at the mean-field level, so that the free-energy difference due to nonspecific interactions depends only on the number density, ρ. We then write the total Helmholtz free energy per monomer, a, as the sum of the ideal, associative and mean-field contributions, $a = a_{\mathrm{id}} + a_{\mathrm{assoc}} + a_{\mathrm{mf}}$:

$$\beta a_{\mathrm{id}} \equiv \frac{1-\rho}{\rho}\ln(1-\rho) + \ln\rho, \tag{2a}$$

$$\beta a_{\mathrm{assoc}} \equiv m\left(\ln p_{\mathrm{unbound}} - \frac{p_{\mathrm{unbound}}}{2} + \frac{1}{2}\right), \tag{2b}$$

$$\beta a_{\mathrm{mf}} \equiv -\frac{1}{2}M\beta\epsilon\rho. \tag{2c}$$

For a fixed value of $\beta\epsilon_{\mathrm{A}}$, Eqs. (2a)–(2c) have a critical point at the smallest value of $\beta\epsilon$ for which the density fluctuations

diverge. We take the derivative of the pressure with respect to volume and solve for the double root, i.e., the value of $\beta\epsilon$ at which there exists only one solution for ρ:

$$-2\left(\frac{\partial a}{\partial \rho}\right)_\beta - \rho\left(\frac{\partial^2 a}{\partial \rho^2}\right)_\beta = 0. \tag{3}$$

As a result of our mean-field approximation of the isotropic lattice-gas, we do not expect the critical points calculated from Eqs. (1)–(3) to be quantitatively accurate. We do, however, expect that the critical lines derived from these equations will predict, at least qualitatively, the effects of specific interactions on the phase behavior of associating fluids.

This theory can also be modified to account for patch-antipatch interactions. In this case, each patch may only associate with a complementary patch, and thus the number of patches m in Eq. (1) must be modified; for the patch-antipatch model with two patches considered in this work, the m in Eq. (1) must be reduced from 2 to 1. The number of patches appearing in Eq. (2b) remains unchanged.

IV. MONTE CARLO CALCULATIONS

We verify the qualitative predictions of SAFT, and investigate the influence of explicit patch orientations, by calculating the $\beta\epsilon$-$\beta\epsilon_A$ critical lines with grand-canonical Monte Carlo simulations. Starting from the known critical point of the three-dimensional lattice gas, $\beta\epsilon_0 \simeq 0.887$ and $\beta\mu_0 = -3\beta\epsilon_0$, we trace the critical line for $\beta\epsilon_A > 0$ by iteratively sampling and updating the thermodynamic fields using finite-size scaling techniques. We choose an $L \times L \times L$ lattice with periodic boundary conditions and $L = 12$ so that all simulations, including those with stable multiply connected oligomers, are carried out in the scaling regime. In Secs. IV A and IV B, we describe our finite-size scaling approach and detail our Monte Carlo algorithm for sampling in the grand-canonical ensemble.

A. Universality and finite-size scaling theory

We follow the mixed-field finite-size scaling theory developed by Wilding and Bruce to fit our grand-canonical simulation data to the universal distribution of the three-dimensional Ising model.[24,25] Because the particle-hole symmetry of the lattice gas is destroyed by the addition of anisotropic interactions, the critical ordering operator $\mathcal{M}$, which describes the direction of diverging fluctuations among all simulation observables, is proportional to a linear combination of the number density, ρ, and the mean internal energy per lattice site, u.[26] Since the interactions remain short-ranged, however, we expect that all critical phenomena observed in these simulations belong to the universality class of the undecorated lattice gas.

Finite-size scaling theory predicts that, at the critical point defined by a unique set of thermodynamic fields, fluctuations in the critical ordering operator $\mathcal{M}$ follow a universal distribution $p_{\mathcal{M}}$, subject to a model-dependent choice of $\mathcal{M}$ and a system-size-dependent proportionality constant Λ.[27] The unknown quantities in our lattice protein models are thus the critical thermodynamic fields, a mixing parameter s that determines the contribution of the internal energy to the ordering operator, and the scaling constant Λ. Only Λ depends strongly on the system size; all other quantities are independent of the system size up to a small correction that decays as a power law in L.

Finding the optimal values for all of the unknown quantities thus reduces the calculation of the critical point at fixed $\beta\epsilon_A$ to a multidimensional curve-fitting problem. We construct a histogram of the calculated grand-canonical probability density distribution, p_{gc}, by projecting p_{gc} onto the basis vector $\hat{\mathcal{M}} \equiv \hat{\rho} - s\hat{u}$ and computing $p_{gc,k}^{(\mathcal{M})} \equiv \Lambda p_{gc}(\rho - su = \delta\mathcal{M}_k)$,

$$p_{gc,k}^{(\mathcal{M})} \equiv \Lambda \sum_v w_v \mathbf{1}\{\delta\mathcal{M}_k \le [(\rho, u)_v \cdot \hat{\mathcal{M}}] < \delta\mathcal{M}_{k+1}\}, \tag{4}$$

where the index v runs over all statistically independent observations $(\rho, u)_v$. Deviations from the ensemble-averaged critical densities are expressed as $\delta\mathcal{M} \equiv \Lambda'(\mathcal{M} - \mathcal{M}_c)$, with $\Lambda' \equiv (1 - sr)\Lambda$; the model-dependent constant r is a second field-mixing parameter that is assumed to be negligibly small. The indicator function $\mathbf{1}\{\cdot\}$ evaluates to unity if its argument is true and to zero otherwise, and the weight w_v is the relative Boltzmann factor introduced by the reweighting of simulation data.[28,29] The histogram bin size is chosen such that $(\delta\mathcal{M}_{k+1} - \delta\mathcal{M}_k) = L^{-3}$.

We treat each histogram bin independently and define the χ^2 function

$$\chi^2 \equiv \sum_k \frac{\left[p_{gc,k}^{(\mathcal{M})}(\beta\epsilon, \beta\epsilon_A, \beta\mu) - p_{\mathcal{M}}(\delta\mathcal{M}_k/\Lambda')\right]^2}{\sigma_k^2}, \tag{5}$$

where the index k runs over all bins with nonzero grand-canonical probability density. The mean value of the critical ordering operator $\mathcal{M}_c$ and the thermodynamic fields $(\beta\epsilon, \beta\epsilon_A, \beta\mu)$ are calculated self-consistently by reweighting the grand-canonical probability density distribution and minimizing χ^2. We estimate the statistical error, σ_k^2, on each histogram bin to first order in the reweighted fields,[30]

$$\sigma_k^2 \equiv \frac{\left(\sum_v w_v^2 \mathbf{1}_{k,v}\right) - \left(\sum_v w_v \mathbf{1}_{k,v}\right)^2 / n_{\text{samples}}}{\sum_v w_v}, \tag{6}$$

where n_{samples} is the total number of statistically independent samples and $\mathbf{1}_{k,v}$ denotes the same indicator function written out explicitly in Eq. (4).

Given simulation data taken from the immediate vicinity of any point on the critical line, we can solve for the critical point at nearby values of $\beta\epsilon_A$ by minimizing Eq. (5). (One exception, in which a critical line is found to terminate in a critical endpoint, is described in Sec. V C 3.) An example of a fitted histogram obtained by this procedure is shown in Figure 2. We find that we can obtain reliable fits, with the reduced χ^2 per occupied histogram bin approximately equal to unity, for steps of $\Delta\beta\epsilon_A \simeq 0.1$; extrapolation to larger differences in the associative free energy fails due to errors accumulated by considering only a finite number of samples. We therefore apply this minimization procedure iteratively by alternately sampling from the grand-canonical distribution at

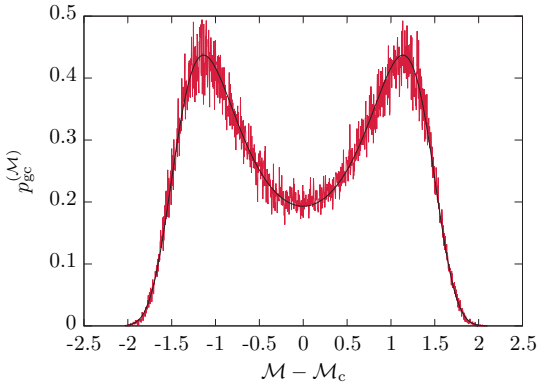


FIG. 2. The universal distribution of the ordering operator, $\mathcal{M}$, of the three-dimensional Ising model. Critical points are calculated by fitting the grand-canonical probability density, $p_{\rm gc}^{(\mathcal{M})}$, (red curve) to the universal distribution, $p_{\mathcal{M}}$, (black curve) for an optimal choice of the ordering operator, $\mathcal{M} \equiv \rho - su$, and its mean value, $\mathcal{M}_{\rm c}$.

the calculated thermodynamic fields and then extrapolating in order to trace out the critical line.

B. Grand-canonical simulations

We perform grand-canonical Monte Carlo simulations by attempting both monomer and cluster insertions and deletions at constant chemical potential. Monomer moves are carried out according to the standard Metropolis acceptance criterion, with the orientations of inserted monomers chosen randomly.[31] Cluster moves are necessary to remove multiply connected oligomers at high values of $\beta\epsilon_{\rm A}$, where individual monomer deletions are unlikely to be accepted. Insertions and deletions of uniquely identifiable oligomers are accepted with probabilities

$$p_{\substack{\text{insert}\\\text{delete}}} = \min\left[1, \left(\frac{N_{\rm clust}}{L^3}\right)^{\mp 1} e^{-\beta\Delta U + \beta\Delta\mu_{\rm clust}}\right], \qquad (7)$$

where $N_{\rm clust}$ is the number of such clusters in existence on the lattice either prior to the move, in the case of a deletion, or after its completion, in the case of an insertion. The change in the total energy of the lattice is represented by ΔU. In Eq. (7), the chemical potential of the cluster, $\mu_{\rm clust}$, includes the entropy lost through association, such that $\beta\mu_{\rm clust} \equiv n_{\rm clust}(\beta\mu - \ln q) + \ln q_{\rm clust}$, where $n_{\rm clust}$ is the number of monomers comprising the cluster. The monomer and cluster partition functions q and $q_{\rm clust}$ are equal to the number of unique particle orientations on the lattice divided by the symmetry number of the monomer or cluster, respectively. The location and orientation of cluster insertions are chosen randomly, and cluster insertions that would overlap occupied lattice sites are rejected automatically.

In order to further accelerate the sampling of statistically independent configurations, we implement one-dimensional multicanonical biasing[32,33] in the direction of the critical ordering operator. We use the known distribution of this one-dimensional projection, $p_{\mathcal{M}}$, to obtain the dimensionless biasing potential $\beta u_{\rm bias}(\delta\mathcal{M}) \equiv \ln[p_{\mathcal{M}}(\delta\mathcal{M}/\Lambda')]$ for the appropriate definition of $\mathcal{M}$. Simulations are then conducted in a modified ensemble in which the Boltzmann weights are proportional to $\exp(N\beta\mu - \beta U - \beta u_{\rm bias})$, and the calculated probability distributions are subsequently reweighted to obtain $p_{\rm gc}$. Since the critical scaling parameters are all calculated prior to running a simulation, implementing this procedure requires little additional work and results in significantly accelerated recrossings of the free-energy barrier between the incipient phases. The statistical independence of samples is ensured by first calculating the $\mathcal{M}$-$\mathcal{M}$ autocorrelation function in a short simulation and then recording configurations of the lattice at intervals greater than the decorrelation time.

V. PHASE DIAGRAMS

Phase diagrams showing the critical lines and critical densities for all models introduced in Figure 1, calculated with both SAFT and numerical simulation, are plotted in Figure 3. In the full space of independent thermodynamic fields $(\beta\epsilon, \beta\epsilon_{\rm A}, \beta\mu)$, a critical line terminates a coexistence surface, which signifies a discontinuous transition in the monomer number densities of two stable, macroscopic phases. The critical lines in the $\beta\epsilon$-$\beta\epsilon_{\rm A}$ plane, shown in Figures 3(a) and 3(b), therefore define the boundary of the homogeneous, supercritical solution. Below and to the left of each critical line, the solution is stable with respect to phase separation at all monomer number densities. Above and to the right, the solution phase separates into phases either enriched or depleted in protein for a finite range of the total monomer number density.

In the quarter-plane shown in Figures 3(a) and 3(b), both the nonspecific and specific free energies of interaction are attractive. As a result, we expect that the critical lines bound a region in the $\beta\epsilon$-$\beta\epsilon_{\rm A}$ parameter space that is unstable with respect to either the condensation of monomers or the condensation, collapse or percolation of specifically associated oligomers. A mean-field regime, in which an orientationally averaged attraction drives a liquid-vapor phase transition, exists at small $\beta\epsilon_{\rm A}$ and is described in Sec. V A. In the presence of stronger specific interactions, the number and placement of the associating patches dramatically affect the phase diagram of a protein solution. The critical nonspecific energy is always decreased relative to the undecorated lattice gas, reflecting the enhanced tendency of larger oligomers to aggregate nonspecifically.[34] With two or fewer associating patches, significant nonspecific interactions are required to observe phase separation, while all models with three or more patches phase separate in the absence of nonspecific interactions. Sections V B–V D discuss how all of these cases can be understood as various limits of specifically and nonspecifically induced phase separation within the Ising universality class.

A. Mean-field regime

When $\beta\epsilon_{\rm A} = 0$, all models shown in Figure 1 reduce to an undecorated lattice gas. Every critical line therefore passes through the lattice-gas critical point in Figure 3(a) and through the mean-field isotropic critical point, $4/M$, in Figure 3(b). As a consequence of the particle-hole symme-

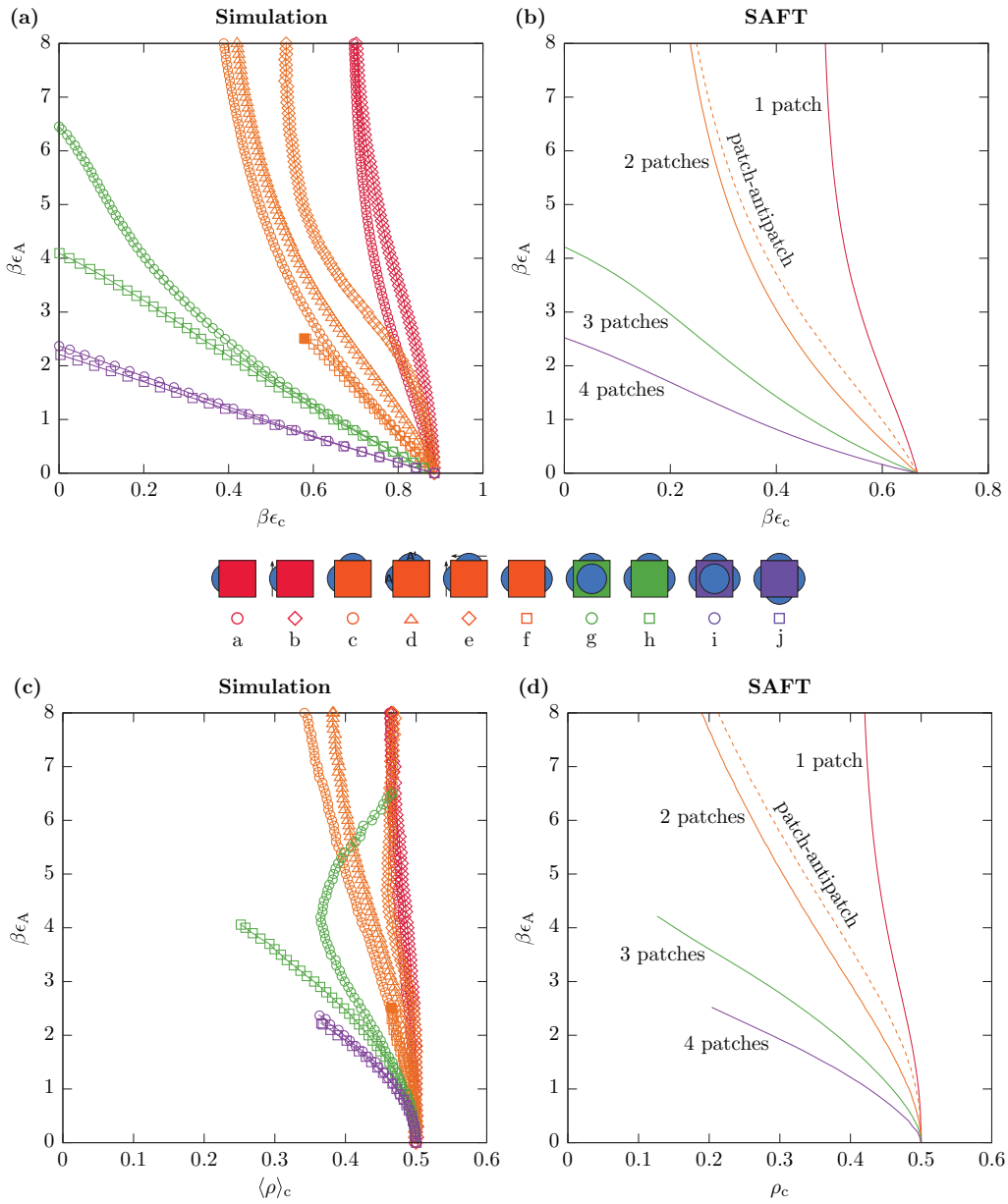


FIG. 3. Critical lines in the $\beta\epsilon$-$\beta\epsilon_A$ plane calculated with (a) Monte Carlo simulations and (b) statistical associating fluid theory (SAFT). All critical lines intersect the lattice-gas critical point when $\beta\epsilon_A = 0$. For each model, the region in parameter space below and to the left of the critical line is stable with respect to phase separation. Significant deviations from the SAFT predictions are observed in models that assemble into multiply connected oligomers, e.g., model (g). The filled point of model (f) indicates a critical endpoint, which is discussed in Sec. V C 3. The critical densities are calculated with (c) Monte Carlo simulations and (d) SAFT. Since particle-hole symmetry is destroyed when $\beta\epsilon_A > 0$, less efficiently packing oligomers come to dominate the incipient condensed phase, and the critical densities decrease with increasing $\beta\epsilon_A$. The critical density of model (g) is nonmonotonic, however, due to the assembly of passivated complexes at large $\beta\epsilon_A$. In (a), statistical errors are smaller than the symbol sizes, while in (c), the errors on $\langle\rho\rangle_c$ are approximately equal to the symbol sizes.

try of the undecorated lattice gas, the $\beta\epsilon_A = 0$ critical points have an associated critical density of 1/2, as shown in both Figures 3(c) and 3(d).

A small perturbation in $\beta\epsilon_A$ leads to a linear response in the critical line. In this regime, the orientationally averaged contribution of specific association provides a mean-field modification to the nonspecific interaction energy among monomers. In models without dihedral or patch-antipatch constraints, this mean-field behavior implies that for $\beta\epsilon_A \ll \beta\epsilon_0$, the critical lines deviate from the undecorated lattice-gas critical point as $(\beta\epsilon_c - \beta\epsilon_0) = -(m^2/M^2)\beta\epsilon_A$. This result can be derived by considering the virial expansion of the lattice fluid linearized in $\beta\epsilon_A$ or by taking the mean-field limit within the SAFT framework.

A second-order perturbation in $\beta\epsilon_A$ is required to see a change in the critical density. As a result, the ensemble-averaged critical density, $\langle\rho\rangle_c$, in Figure 3(c) and the critical density, ρ_c, calculated from the SAFT equations in Figure 3(d) initially decrease quadratically as $\beta\epsilon_A$ is increased. With particle-hole symmetry destroyed by the introduction of anisotropic patches, specifically associated oligomers pack less efficiently in the condensed phase. The associating homogeneous fluid therefore phase separates at a lower density than the reference monomer fluid with isotropic

interactions alone. This decrease in the critical density has also been observed in off-lattice simulations, for instance in the formation of so-called network fluids.[35,36]

B. Passivation by complex formation

The assembly of stable protein complexes prevents further aggregation due to specific association. In a complex, the monomers are associated internally through specific interactions in such a way that all "bonds" are satisfied. With no patches exposed on the outer surface, complexes are effectively passivated and phase separate only occurs through nonspecific condensation. Although the equilibrium constant associated with the assembly of the complex from a reservoir of monomers can be easily calculated, predicting the fraction of monomers in a particular type of complex at the onset of phase separation requires knowledge of the critical chemical potential. For this reason, the effect of complexation on the stability of the lattice protein solution must be determined by self-consistently calculating $\beta\mu_c$ with numerical simulations.

The topologies of the particular complexes that can be formed are dictated by the arrangement of the specific association sites on the monomers. In the case of a single patch, association leads to the formation of passivated dimers. As the specific interaction energy is increased, the probability of a monomer being associated in a dimer at the critical point increases monotonically until phase separation is driven entirely by nonspecific interactions in a dimer fluid. This continuous transition is shown in Figure 4, where the fraction of associated patches is $\langle p_{\text{bound}}\rangle = \langle 2n_{\text{p-p}}/N\rangle$ and $n_{\text{p-p}}$ is the number of patch-patch nearest-neighbor pairs on the lattice. At the lattice-gas critical point, the bound fraction is equal to $2/3q$, where q is the number of distinguishable monomer orientations: in the case of model (a), $q = M$, while for model (b), $q = 4M$. In close agreement with our simulation results, the SAFT equations predict that, for $\beta\epsilon_A \gg 1$, the asymptotic value of the critical nonspecific energy is $\beta\epsilon_c = (3/2 + \sqrt{2})/4\beta\epsilon_0$. The predicted critical density tends to $\rho_c = \sqrt{2} - 1 \simeq 0.414$ and is independent of the lattice coordination number.

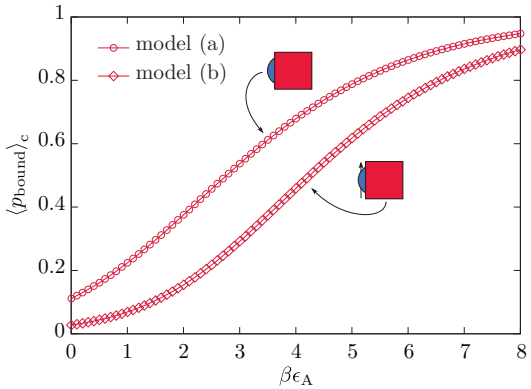


FIG. 4. The probability of a single-patch monomer being associated in a dimer at the onset of nonspecifically driven phase separation. The bound fractions are determined from the results of Monte Carlo simulations conducted along the critical line.

With more than two patches, all lattice protein complexes are multiply connected, in the sense that every pair of monomers in the complex is either directly or indirectly connected by more than one path of specific bonds. These "loops" are not included in the derivation of Eqs. (2a)–(2c), and thus we do not expect the SAFT equations to yield reliable predictions when multiply connected complexes dominate the grand partition function. In models with two adjacently placed patches (models (c), (d), and (e)), the smallest complex is a tetramer, while for the three-patch model with mutually adjacent patches (model (g)) the smallest complex is a cubic octamer. In both cases, passivated complexes composed of more than the minimum number of monomers, such as ring polymers in the two-patch case, are possible although generally not as significantly populated at the critical point.

In the limit $\beta\epsilon_A \to \infty$, all monomers are assumed to be specifically associated in oligomers. We can thus treat a solution consisting solely of passivated complexes as a renormalized lattice gas. In cases where there is competition between the assembly of complexes and the formation of large specifically associated polymeric aggregates, a dihedral constraint can be added to stabilize a particular complex. Applying this constraint to the one-patch model reduces the entropy of the dimer relative to the monomer fluid but has no effect in the limit of strong association. Yet when a dihedral constraint is used to force the assembly of a tetramer from two-patch monomers, the absence of larger oligomers qualitatively changes the phase behavior in this limit. The critical line of model (e) therefore differs markedly from those of the unconstrained two-patch models (c) and (d).

C. Condensation and collapse of specifically associated oligomers

Protein models with two or more patches are capable of forming specifically associated oligomers with linear, ring, or branched topologies. Restricting our attention to models that do not enforce the assembly of specific complexes through dihedral constraints, we can expect to see a much wider variety of phase transitions outside of the mean-field regime. Not only can nonspecific interactions drive phase separation through the condensation of oligomers, but the oligomers themselves can collapse due to internal nonspecific contacts. In general, phase separation tends to demix the homogeneous solution into phases with differing oligomeric size distributions and scaling behaviors.

1. Scaling of specifically associated chain oligomers

We focus first on two-patch models, which cannot percolate at finite density. Along the critical line with $\beta\epsilon_A \gtrsim 2$, we find that the majority of monomers are associated in some form of oligomer for both the case of the adjacently arranged and oppositely arranged two-patch models. The most probable oligomer sizes are generally quite small, although the distribution of associated oligomers broadens as the specific interaction energy is increased. While we cannot determine the mixing parameter s to great accuracy from our limited simulation data and the discrete nature of the fitting process,

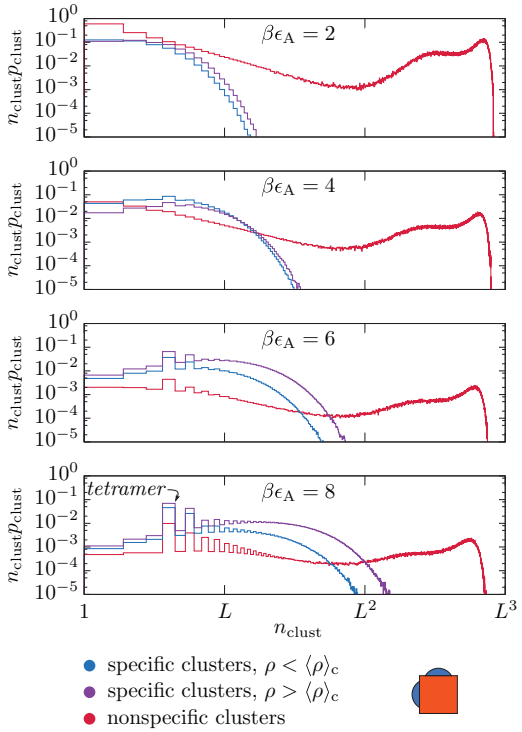


FIG. 5. Distributions of clusters consisting of n_{clust} monomers for model (c). The cluster distributions are calculated with Monte Carlo simulations at the critical point for various specific association energies, $\beta\epsilon_A$. These log-log plots show clusters defined by specific contacts in the incipient low and high-density phases as well as clusters bound by all nonspecific nearest-neighbor interactions. The large fraction of monomers comprising small ring polymers can be seen from the distribution of even n_{clust} calculated at $\beta\epsilon_A = 8$.

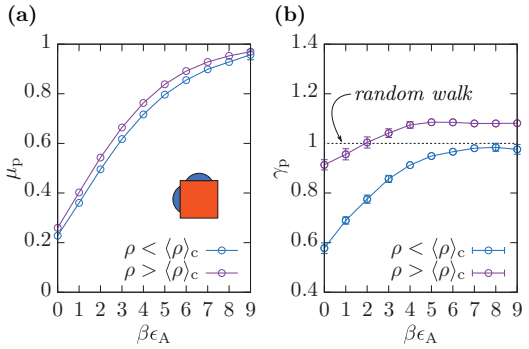


FIG. 6. [(a) and (b)] Connective constants, μ_p, and effective scaling exponents, γ_p, for specifically associated linear polymers formed by monomers of model (c). Although the most probable oligomer sizes in Figure 5 are small, we nevertheless obtain good statistics for calculating μ_p and γ_p from the tails of the cluster distributions. The scaling properties are reported separately for the incipient low-density phase, $\rho < \langle\rho\rangle_c$, and the incipient high-density phase, $\rho > \langle\rho\rangle_c$. Points at $\beta\epsilon_A = 9$ were calculated with $L = 14$ in order to avoid finite-size effects. Unless otherwise indicated by error bars, statistical errors are smaller than the symbol sizes.

we do observe s increasing monotonically with $\beta\epsilon_A$ as the antisymmetric correction to the critical ordering distribution grows.

In order to describe the nature of phase separation along the critical line in greater detail, we have calculated the size distributions of specifically associated clusters in both incipient phases. This can be accomplished because, in a finite-size simulation, it is possible to distinguish the two incipient phases at the pseudo-critical point since they are separated by a nonvanishing free-energy barrier.[37] The distribution of cluster sizes for model (c), calculated with Monte Carlo simulations along the critical line at increasing values of $\beta\epsilon_A$, are plotted in Figure 5. We distinguish between clusters consisting of n_{clust} monomers bound by specific associations alone and those bound by any nearest-neighbor contact. The specific distributions are further separated into clusters observed in the incipient vapor phase, $\rho < \langle\rho\rangle_c$, and the incipient condensed phase, $\rho > \langle\rho\rangle_c$. The Ising-like form of the nonspecific interactions remains essentially unchanged at large n_{clust} for the entire range of $\beta\epsilon_A$. Comparing the nonspecific and specific distributions clearly indicates that the incipient high-density phase is composed of many specifically associated oligomers that are aggregated through nonspecific contacts.

This two-patch model is capable of forming both linear polymers, in which the self-avoiding monomers trace out a zig-zag pattern on the cubic lattice, and ring polymers, for which an even number of monomers makes an equal number of specific contacts in a closed loop. According to general scaling arguments applied to lattice polymers, we expect that the cluster distribution of odd n_{clust}, for which no closed loops can be formed, obeys the scaling law $n_{clust}p(n_{clust}) \propto \mu_p^{n_{clust}} n_{clust}^{\gamma_p}$ for some values of the connectivity constant μ_p and effective scaling exponent γ_p.[38] Within the grand-canonical ensemble, our definition of μ_p includes the monomer fugacity and the Boltzmann factor associated with the $n_{clust} - 1$ intrapolymer specific contacts. Fits to this scaling law are shown in Figure 6.

In Figure 6(a), the connective constants in both the low and high-density phases approach unity with increasing $\beta\epsilon_A$. If μ_p were to equal unity, then the cluster size distribution would no longer be normalizable, and we would expect to see percolating linear polymers. The SAFT equations also suggest this behavior in the limit $\beta\epsilon_A \to \infty$, where the critical density is predicted to go to zero and the critical nonspecific energy to approach $1/M$.[39] The calculated values of μ_p show that, at finite densities and reasonable values of the specific association energy, the relevant critical line bounding the homogeneous solution does not arise from specifically driven polymerization but rather indicates the onset of nonspecifically induced condensation.

Nonspecific interactions also encourage the collapse of monomers within a specifically associated oligomer. For linear polymers, this collapse occurs at the Flory Θ-point, beyond which the polymers assume a compact configuration with a typical radius of $(n_{clust})^{1/3}$.[40,41] This compact phase, which in our circumstances appears as the incipient high-density phase, is characterized by a nonuniversal exponent γ_p, while the coexisting polymer melt at the Θ-point is known to have mean-field exponents.[42] As the specific association energy increases, Figure 6(b) shows that γ_p in the high-density phase converges to a constant value of 1.08, while the low-density phase appears to approach the mean-field exponent $\gamma_p = 1$.

The critical line for the adjacently arranged two-patch model thus describes phase separation in a polydisperse polymer fluid that is analogous to the monomer liquid-vapor

transition: above the critical nonspecific energy, a high-density compact phase coexists with a low-density ideal phase of the same symmetry. In the supercritical homogeneous fluid, where the nonspecific interactions are weaker, we expect the specifically associated oligomers to behave like self-avoiding random walks. This analysis is complicated by the presence of small passivated oligomers, which appear in the cluster distributions for even n_{clust} at large $\beta\epsilon_{\text{A}}$. For this reason, it is appropriate to consider phase separation along this critical line to be the result of two contributions: one from condensation, which results from the nonspecific aggregation of protein complexes and collapsed oligomers, and another from demixing, in which the supercritical mixture of specifically associated oligomers segregates into two phases with distinct oligomer size distributions.

Qualitatively similar phase behavior occurs along the critical lines for models with more than two patches when the strength of the specific interactions is approximately 1 k_BT. Specific association in these cases can lead to the formation of branched polymers, which also undergo collapse, demixing, and condensation at the phase boundary. Because branched polymers can percolate at finite $\beta\epsilon_{\text{A}}$, we shall inspect the critical phenomena associated with these models in more detail in Sec. V D.

2. Enhanced stability of patch-antipatch solutions

Thus far, we have only described the phase behavior of lattice protein models in which all associating patches are identical. The addition of a patch-antipatch constraint, which restricts specific association to interactions between complementary patches, does not alter the possible topologies of specifically associated oligomers but does affect the equilibria among them. We therefore expect to find phase diagrams that are isomorphic to those of monomers with identical patches, although shifted with respect to $\beta\epsilon_{\text{A}}$.

In the case of the two-patch patch-antipatch model, model (d), the symmetry number of the monomers is halved, and all specifically associated oligomers possess an inherent polarization. The entropic penalty for oligomer formation is consequently greater, and, at constant $\beta\epsilon_{\text{A}}$, the smaller oligomers of the patch-antipatch model require an increased nonspecific energy for phase separation relative to the patch-patch model. As a result, the stability of the homogeneous solution with respect to nonspecific aggregation is increased for a wide range of the specific association energy. This result is also obtained from the SAFT equations, which predict equivalence between the patch-patch and patch-antipatch models in the limit $\beta\epsilon_{\text{A}} \to \infty$.

3. Critical endpoint and a "lattice nematic" phase

The critical line for monomers with oppositely placed patches, model (f), terminates in a critical endpoint in our Monte Carlo simulations. At $\beta\epsilon_{\text{A}} \simeq 2.41$, a third distinct phase,[43] characterized by a monomer density greater than $\langle\rho\rangle_{\text{c}}$ and a more negative average energy per monomer, appears in the critical ordering-operator distribution, as shown in Figure 7. The existence of a critical endpoint implies that three phases coexist along a continuous line in parameter space; this is possible in a single-component fluid due to the inclusion of $\beta\epsilon_{\text{A}}$ as a third independent thermodynamic field.

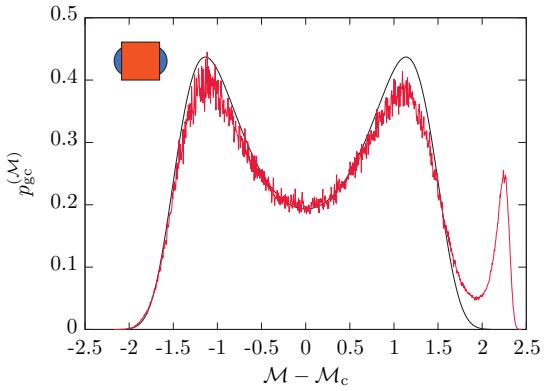


FIG. 7. The distribution of the critical ordering operator of the two-patch model with opposite patches, model (f), in the vicinity of a critical endpoint. The critical line terminates in a critical endpoint at $\beta\epsilon_{\text{A}} \simeq 2.41$ when a lower-energy, higher-density phase, appearing here as the rightmost peak, becomes stable.

We label this third phase a "lattice nematic" phase since it consists of lattice-spanning, specifically associated columnar polymers that are condensed due to lateral nonspecific interactions. The precise determination of the thermodynamic fields at the critical endpoint would require equalizing the total probability density, and thus eliminating the pressure differences, in all three phases, although this could not be accomplished with simulation data obtained with the $L = 12$ lattice. Most likely, this phase is not stable off-lattice, where the persistence length of the linear polymers would be finite.

D. Specifically driven phase separation

In contrast with the previously discussed protein models, monomers with more than two associating patches are capable of phase separating in the absence of nonspecific interactions. As shown in Figure 3, the critical lines for all protein models with three or more patches intersect $\beta\epsilon = 0$ at finite values of the specific association energy, regardless of the arrangement of the associating patches on the monomers. These critical points therefore indicate the uncontrolled aggregation of protein monomers that is driven entirely by specific protein-protein contacts.

The existence of a critical point at $\beta\epsilon = 0$ corresponds to the appearance of a giant component in a polydisperse mixture of specifically associated oligomers. If $\beta\epsilon_{\text{A}}$ is increased beyond this critical point, the monomer fluid may phase separate to form an infinite cluster bound solely by specific contacts. The critical point associated with this phase transition is thus the unique "percolation point," the minimum value of $\beta\epsilon_{\text{A}}$ at which we observe an analogue of bond percolation that is consistent with the fixed patch orientations on the model monomers. Field-mixing at the percolation point is

particularly significant in both three-patch models, where the density and internal energy terms make roughly equal contributions to the critical ordering operator. In all protein models with three or more patches, we find that the critical lines go continuously from the lattice-gas critical point to the percolation point, implying that these qualitatively distinct phenomena are simply two limiting cases of the critical behavior of fluids with both specific and nonspecific interactions.

1. Planar three-patch model

Arranging three associating patches within a plane, as in model (h), prohibits the formation of completely passivated oligomers, since at least three associating patches must always be exposed to the solvent. Beyond the mean-field regime of monomer condensation, nonspecific interactions drive the condensation of a bimodal mixture of small oligomers and collapsed branched polymers. This behavior can be seen in the $\beta\epsilon_A = 2$ panel of Figure 8(b) by comparing the distributions of nonspecific clusters and specifically associated clusters in the incipient high-density phase at large n_{clust}. The coexisting incipient dilute phase consists of a polydisperse mixture of specifically associated polymers that is dominated by small oligomers.

As the specific interaction energy approaches the percolation point, $\beta\epsilon_A \simeq 4.09$, we find that phase separation is driven primarily by aggregation through specific contacts. The fact that this value is almost quantitatively predicted by the SAFT equations suggests that loops of specific contacts do not contribute significantly to the grand partition function at the percolation point, since such multiply connected structures are neglected in the theory. The SAFT equations also provide a qualitative prediction of the significant reduction in the critical density at the percolation point. This low critical density indicates inefficient packing in the high-density incipient phase and provides further evidence that the giant component is dominated by weakly embedded structures with few self contacts.[38]

2. Complex-assembling three-patch model

Arranging the three associating patches on mutually adjacent faces, as in model (g), permits the formation of passivated complexes, in which all patches are internally associated. These complexes, the smallest of which is an octamer, achieve stability by maximizing the number of specific contacts per monomer. Since it is not *a priori* obvious whether monomers prefer to be associated in small complexes or in larger branched polymers along the critical line, it is necessary to determine the fraction of monomers associated in passivated complexes by self-consistently solving for the critical chemical potential without making any simplifying assumptions, as in Ref. 44. In this way, we can assess the extent to which complex assembly affects the onset of phase separation.

We expect to see qualitatively different phase behavior in fluids with dominant fluctuations driven either by specific or by nonspecific interactions. From simple renormalization arguments, we know that the nonspecific condensation of an octamer fluid has a critical point at $\beta\epsilon = \beta\epsilon_0/4$ and a critical density of 1/2. We expect specifically driven phase separation, on the other hand, to result in the formation of a multiply connected branched polymer that forms a giant component at $\beta\epsilon = 0$. Simulations show that the critical line bounding the homogeneous solution has contributions from both of these limiting behaviors. Given sufficiently strong associative interactions, however, the entropically favored branched polymers dominate, and we observe phase separation in spite of the assembly of passivated complexes.

As shown in Figure 8(a), a peak in the low and high-density specific cluster distributions appears at $n_{clust} = 8$ near $\beta\epsilon_A = 4$, indicating a significant concentration of completely assembled octamers at the critical point. As the associative

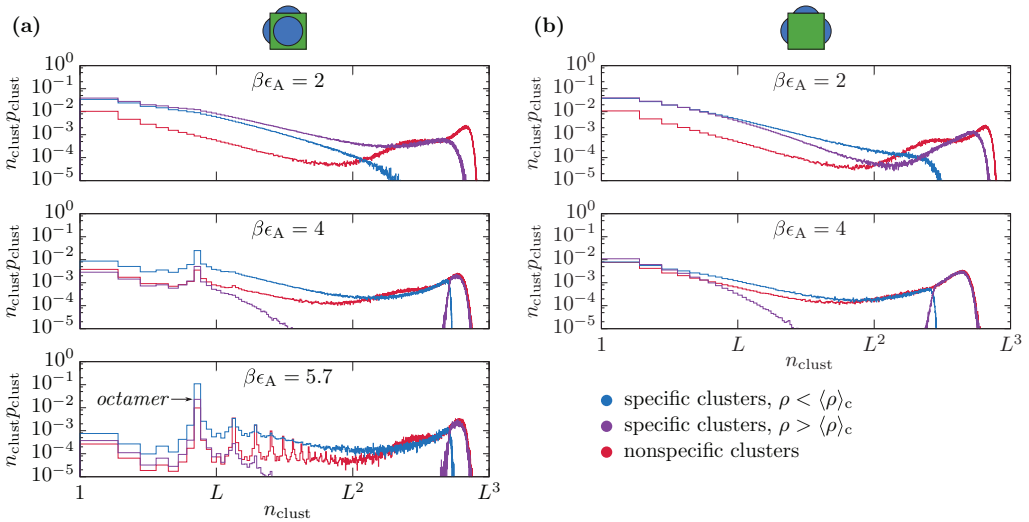


FIG. 8. Distributions of clusters consisting of n_{clust} monomers for protein models with three patches arranged (a) on mutually adjacent faces and (b) in a plane. The low-density specific, high-density specific, and nonspecific distributions are defined as in Figure 5. The peak associated with the assembly of passivated octamer complexes is apparent in (a) at $\beta\epsilon_A = 4$, as are larger clusters at $\beta\epsilon_A = 5.7$. In (b), a giant component appears close to the percolation point in the high-density incipient phase, and the specific and nonspecific cluster distributions converge.

interaction strength is increased further, the critical density actually increases, and the octamer peak persists in both the dilute and condensed incipient phases. Since the octamer itself is multiply connected, it is clear that loops of specific contacts make a significant contribution to the grand partition function in this regime; as a result, we find a marked disagreement between the predictions of the SAFT equations and the critical line of model (g). At the percolation point, $\beta\epsilon_A \simeq 6.45$, a sufficient number of monomers associate into a branched oligomer to form a giant component in the dense phase despite the high concentration of stable complexes. The critical density at the percolation point is approximately 0.463, suggesting that the stable complexes are filling in the voids that remain in the giant component.

Phase separation due to the formation of a specifically associated giant component in this protein model is a significant result, since it implies that excessively strong specific interactions can destabilize the protein solution. In the absence of appreciable nonspecific interactions, there is a large region in parameter space in which self-assembled octamers are the most common species in the solution. Yet given sufficiently attractive specific interactions, a high-density gel can become the most stable phase, as observed in Ref. 7. We stress that the formation of a gel beyond the percolation point occurs at equilibrium and is not related to kinetic arrest. In addition, we note that it is not sufficient to consider the equilibrium between monomers and a complex of a particular topology in cases where a partially unfolded complex can take part in either specifically or nonspecifically driven aggregation. Although a target complex could be stabilized for all values of $\beta\epsilon_A$ by introducing dihedral constraints, the possibility of aggregation through the exposure of a hydrophobic association site is likely to be relevant in real protein solutions.

3. Four-patch models

The critical behavior of both four-patch protein models is well described by mean-field arguments, even with $\beta\epsilon = 0$. As is characteristic of the mean-field regime, Figure 3(a) shows that these critical lines are nearly linear in $\beta\epsilon_A$ and that the critical densities in Figure 3(c) diverge from the lattice-gas critical density quadratically. Furthermore, there is little difference between the four-patch models: the percolation point for the planar arrangement (model (j)) occurs at $\beta\epsilon_A = 2.22$, while for the clustered configuration (model (i)) it is found at $\beta\epsilon_A = 2.37$. For comparison, if the associative interactions were perfectly averaged over all orientations, then we would expect the percolation point to occur at $\beta\epsilon_A = 1.99$ from the mean-field arguments presented in Sec. V A. The SAFT equations also give a nearly quantitative prediction for the percolation point of both models.

With more than fifty percent of the surface area occupied by associating patches, it is probably more appropriate to describe these four-patch models as nearly isotropic monomers than as canonical patchy particles. For these models, it is clear that nonspecific monomer condensation and phase separation driven solely by anisotropic interactions are simply the two limiting cases of the critical behavior of an associating fluid with both specific and nonspecific attractive interactions.

VI. DISCUSSION

The critical phenomena that we have described for this set of simplified protein models have significant implications for real protein solutions. While it is clear that monomers with only one effective interaction site never pose a substantial threat to the thermodynamic stability of a protein solution, this is not necessarily true for all complex-forming proteins. If the arrangement of protein-protein interaction sites allows for both the formation of both finite-size complexes and extended structures, then the exposure of a hydrophobic surface from within an otherwise stable complex may lead to disassembly and aggregation. Phase separation can thus occur in solutions with sufficiently strong specific interactions, even in the absence of significant nonspecific interactions. For instance, in the case model (g), our calculations show that aggregation due to specific contacts can drive phase separation in a solution in which pairwise monomer interactions are characterized by an approximately micromolar dissociation constant.

This systematic study provides a framework for rationalizing the phase behavior of a wide variety of protein solutions. For all protein models studied here, our results indicate that the critical behavior belongs to the Ising universality class, as suggested by both experimental measurements[45] and a previous simulation study of a smaller set of monomer geometries and interaction parameters.[46] In agreement with experimental observations, the critical behavior of our simple protein models depends strongly on both the geometry of the protein monomer and the distribution of oligomer sizes.[47,48] These calculations should also facilitate the interpretation of experiments relating complex assembly, stability and supersaturation, as suggested in Ref. 49.

The coarse-grained proteins studied in our simulations necessarily possess much higher symmetry than real globular proteins. Specific protein-protein contacts are typically composed of multiple residues, whose relative orientations afford a directionality to protein-protein interactions not unlike the dihedral constraints considered in this work. On the other hand, the enormous reduction in translational and rotational degrees of freedom imposed by the lattice artificially stabilizes the assembly of specific complexes. We therefore expect that the observed competition between protein-complex assembly and aggregation is relevant to off-lattice, lower symmetry models as well. Furthermore, we have shown that reducing the symmetry of a coarse-grained protein by introducing patch-antipatch interactions does not qualitatively change the phase behavior of the solution.

One limitation of studying a lattice model is the inability to distinguish between positionally ordered and disordered phases. In particular, it is not possible to describe crystallization within the present model; as a result, our calculations describe only a portion of the complete phase diagram. The critical lines examined here most likely indicate the onset of metastable liquid-liquid phase separation in solutions with very short-range interactions.[3,50,51] In qualitative agreement with our results, the experimentally determined phase

diagrams of multiple globular proteins in the metastable fluid show that the critical density is reduced in comparison to a reference fluid of spherical, isotropic monomers.[52] Fluctuations near a metastable critical point are also believed to play an important role in the two-step nucleation of protein crystals.[53–58] The present study may therefore have implications for the crystallization of anisotropic globular proteins.

A compensating advantage of examining a coarse-grained lattice model, however, is the ability to make generic predictions about a wide variety of associating fluids. By employing an extended law of corresponding states,[59] it is possible to use the lattice second-virial coefficient to map the phase diagrams appearing in Figure 3 onto colloidal fluids with an equivalent number of associating patches. We believe that our predictions for the critical phenomena of associating fluids may in this way be reproduced in off-lattice models of both spherical and aspherical anisotropic patchy particles. One potential pitfall in making such a correspondence lies in counting the number of distinct patches: the single-bond steric constraint assumed by SAFT and enforced by our lattice model may be too stringent in some cases, and the identification of an integer number of distinct interaction sites may not always be appropriate. Nevertheless, we believe that the rigid placement of associating patches on our monomer models is a reasonable approximation of real proteins, especially those that form specific complexes. As a result, we expect that the qualitative variations in phase behavior arising from the number and arrangement of specific interaction sites are directly relevant to naturally occurring protein solutions.

Finally, we note that the lattice variant of SAFT performs exceptionally well in protein models that primarily display dimer formation, linear polymerization or percolation via weakly embedded branched polymers. In particular, both the mean-field behavior near the lattice-gas critical point and the predicted percolation points of non-complexing monomers agree almost quantitatively with simulation results. The deviations in the critical densities from that of the undecorated lattice-gas are also qualitatively consistent, if generally overestimated. It is clear, however, that the present version of SAFT is not sufficient for predicting phase behavior in cases where loops of specific contacts and the assembly of passivated complexes are significant.

VII. CONCLUSIONS

We have investigated the critical behavior of fluids with both specific and nonspecific interactions by performing hundreds of complete mixed-field critical point calculations. We have provided simulation evidence that the critical behavior of fluids with short-range, attractive, anisotropic interactions belongs to the Ising universality class. Monomer and complex condensation, polymer collapse and demixing, and percolation through specific association can all be regarded as various limiting cases of this critical behavior, depending on the number of associating patches on each monomer. For each protein model considered in this work, there is a critical line that bounds the stable, homogeneous phase. If the strength of either the specific or nonspecific interactions is increased beyond this line, the protein solution may phase separate at equilibrium for a finite range of densities. In general, the specific or nonspecific driving force for phase separation and the oligomeric size distributions in the incipient phases vary continuously along each critical line.

We have shown that protein solutions require substantial nonspecific interactions in order to phase separate when associated in completely passivated complexes. Polydisperse mixtures of specifically associated oligomers that leave associating patches exposed to the solvent can phase separate either by collapse and condensation in the case of linear polymers, or by the formation of an infinite, specifically associated cluster in the case of branched polymers. In cases where complex assembly competes with specifically driven aggregation, the formation of finite-size complexes tends to suppress phase separation. Yet with sufficiently strong specific interactions, an associating fluid may still be vulnerable to phase separation in the absence of additional constraints that force the assembly of a unique complex. We believe that these qualitative features of phase separation in associating fluids are widely applicable to naturally occurring protein solutions.

ACKNOWLEDGMENTS

W.M.J. acknowledges support from the Gates Cambridge Trust and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1143678. D.W.O. appreciates support as a Visiting Fellow at Trinity College, Cambridge. D.F. acknowledges support of ERC Advanced Grant 227758 and EPSRC Programme Grant EP/I001352/1.

[1] J. Gsponer and M. M. Babu, Cell Rep. **2**, 1425 (2012).
[2] J. D. Gunton and A. Shiryayev, *Protein Condensation: Kinetic Pathways to Crystallization and Disease* (Cambridge University Press, 2007).
[3] J. A. Thomson, P. Schurtenberger, G. M. Thurston, and G. B. Benedek, Proc. Natl. Acad. Sci. USA **84**, 7079 (1987).
[4] F. Zhang, R. Roth, M. Wolf, F. Roosen-Runge, M. W. A. Skoda, R. M. J. Jacobs, M. Stzucki, and F. Schreiber, Soft Matter **8**, 1313 (2012).
[5] T. W. Han, M. Kato, S. Xie, L. C. Wu, H. Mirzaei, J. Pei, M. Chen, Y. Xie, J. Allen, G. Xiao *et al.*, Cell **149**, 768 (2012).
[6] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, and A. A. Hyman, Science **324**, 1729 (2009).
[7] 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.*, Nature (London) **483**, 336 (2012).
[8] C. D. Keating, Acc. Chem. Res. **45**, 2114 (2012).
[9] R. P. Sear, Faraday Discuss. **139**, 21 (2008).
[10] R. P. Sear, Soft Matter **3**, 680 (2007).
[11] C. P. Brangwynne, T. J. Mitchison, and A. A. Hyman, Proc. Natl. Acad. Sci. USA **108**, 4334 (2011).
[12] A. A. Hyman and C. P. Brangwynne, Developmental Cell **21**, 14 (2011).
[13] C. M. Roth and A. M. Lenhoff, Langmuir **11**, 3500 (1995).
[14] M. L. Broide, T. M. Tominc, and M. D. Saxowsky, Phys. Rev. E **53**, 6325 (1996).
[15] I. K. Voets, V. Trappe, and P. Schurtenberger, Phys. Chem. Chem. Phys. **14**, 2929 (2012).
[16] C. P. Brangwynne, Soft Matter **7**, 3052 (2011).
[17] W. M. Jacobs and D. Frenkel, J. Chem. Phys. **139**, 024108 (2013).
[18] V. G. Taratuta, A. Holschbach, G. M. Thurston, D. Blankschtein, and G. B. Benedek, J. Phys. Chem. **94**, 2140 (1990).
[19] W. G. Chapman, K. E. Gubbins, G. Jackson, and M. Radosz, Fluid Phase Equilib. **52**, 31 (1989).
[20] W. G. Chapman, K. E. Gubbins, G. Jackson, and M. Radosz, Ind. Eng. Chem. Res. **29**, 1709 (1990).
[21] M. S. Wertheim, J. Stat. Phys. **35**, 19 (1984).
[22] M. S. Wertheim, J. Stat. Phys. **42**, 459 (1986).
[23] M. L. Michelsen and E. M. Hendriks, Fluid Phase Equilib. **180**, 165 (2001).

[24]N. B. Wilding and A. D. Bruce, J. Phys.: Condens. Matter **4**, 3087 (1992).
[25]A. D. Bruce and N. B. Wilding, Phys. Rev. Lett. **68**, 193 (1992).
[26]N. B. Wilding, Z. Phys. B **93**, 119 (1993).
[27]M. Tsypin and H. Blöte, Phys. Rev. E **62**, 73 (2000).
[28]A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
[29]Z. Tan, E. Gallicchio, M. Lapelosa, and R. M. Levy, J. Chem. Phys. **136**, 144102 (2012).
[30]A. M. Ferrenberg, D. P. Landau, and R. H. Swendsen, Phys. Rev. E **51**, 5092 (1995).
[31]D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, 2001).
[32]B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[33]N. B. Wilding, Am. J. Phys. **69**, 1147 (2001).
[34]T. K. Haxton and S. Whitelam, Soft Matter **9**, 6851 (2013).
[35]E. Bianchi, J. Largo, P. Tartaglia, E. Zaccarelli, and F. Sciortino, Phys. Rev. Lett. **97**, 168301 (2006).
[36]J. Russo, J. M. Tavares, P. I. C. Teixeira, M. M. T. da Gama, and F. Sciortino, Phys. Rev. Lett. **106**, 085703 (2011).
[37]N. B. Wilding, J. Phys.: Condens. Matter **9**, 585 (1997).
[38]C. Vanderzande, *Lattice Models of Polymers* (Cambridge University Press, 1998).
[39]The nonspecific energy does not tend to zero in our variant of SAFT, since the ideal lattice entropy in Eq. (2a) includes an effective excluded volume interaction.
[40]A. Maritan, F. Seno, and A. L. Stella, Physica A **156**, 679 (1989).
[41]R. Dekeyser, E. Orlandini, A. L. Stella, and M. C. Tesi, Phys. Rev. E **52**, 5214 (1995).
[42]P. G. De Gennes, J. Phys. Lett. **36**, 55 (1975).
[43]N. B. Wilding, Phys. Rev. E **55**, 6624 (1997).
[44]A. Reinhardt, A. J. Williamson, J. P. Doye, J. Carrete, L. M. Varela, and A. A. Louis, J. Chem. Phys. **134**, 104905 (2011).
[45]P. Schurtenberger, R. A. Chamberlin, G. M. Thurston, J. A. Thomson, and G. B. Benedek, Phys. Rev. Lett. **63**, 2064 (1989).
[46]H. Liu, S. K. Kumar, and F. Sciortino, J. Chem. Phys. **127**, 084902 (2007).
[47]M. L. Broide, C. R. Berland, J. Pande, O. O. Ogun, and G. B. Benedek, Proc. Natl. Acad. Sci. USA **88**, 5660 (1991).
[48]N. Asherie, J. Pande, A. Lomakin, O. Ogun, S. R. A. Hanson, J. B. Smith, and G. B. Benedek, Biophys. Chem. **75**, 213 (1998).
[49]P. Ciryam, G. G. Tartaglia, R. I. Morimoto, C. M. Dobson, and M. Vendruscolo, Cell Rep. **5**, 781 (2013).
[50]M. H. J. Hagen and D. Frenkel, J. Chem. Phys. **101**, 4093 (1994).
[51]M. Dijkstra, Phys. Rev. E **66**, 021402 (2002).
[52]Y. Wang, A. Lomakin, R. F. Latypov, J. P. Laubach, T. Hideshima, P. G. Richardson, N. C. Munshi, K. C. Anderson, and G. B. Benedek, J. Chem. Phys. **139**, 121904 (2013).
[53]P. R. ten Wolde and D. Frenkel, Science **277**, 1975 (1997).
[54]V. Talanquer and D. W. Oxtoby, J. Chem. Phys. **109**, 223 (1998).
[55]O. Galkin and P. G. Vekilov, J. Am. Chem. Soc. **122**, 156 (2000).
[56]N. Kern and D. Frenkel, J. Chem. Phys. **118**, 9882 (2003).
[57]S. Whitelam, Phys. Rev. Lett. **105**, 088102 (2010).
[58]P. G. Vekilov, J. Phys.: Condens. Matter **24**, 193101 (2012).
[59]M. G. Noro and D. Frenkel, J. Chem. Phys. **113**, 2941 (2000).