1 Phase separation in biology

The biochemical reactions that sustain living things do not take place homogeneously within living cells. Instead, biochemistry is spatially organized in order to colocalize enzymes with their substrates, to create ion gradients that provide the free energy required to do mechanical and chemical work, and to control the mechanisms by which specific genes are transcribed, among many other critical processes. Membrane-bound organelles, such as the nucleus and mitochondria, represent one common form of compartmentalization within eukaryotic cells. However, spatial organization can also occur in the absence of membranes. Compartments that spontaneously form in this way are collectively referred to as *biomolecular condensates*. Over the past 15 years, a large number of biomolecular condensates have been identified and linked to a wide range of biological phenomena, including transcription, gene splicing, translation, protein-concentration buffering, stress response, and signal processing.

Many biomolecular condensates are believed to form via liquid–liquid phase separation (LLPS), the physical process by which complex fluid phases demix into distinct thermodynamic phases with different macromolecular concentrations. LLPS implies that a surface tension holds the phase-separated condensate together, while individual biomolecules—including proteins, RNAs, and other small molecules—exchange between the condensate and the surrounding fluid in dynamic equilibrium. These features render condensates distinct from irreversible aggregates, such as amyloid fibrils, or stoichiometric ribonucleic complexes, such as ribosomes. Moreover, phase-separated condensates represent a unique form of biological organization compared to traditional membrane-bound organelles, since the absence of a membrane allows for rapid assembly and disassembly in response to stimuli.

Because LLPS in biology involves complex macromolecules, concepts from polymer physics have proven essential to understanding how interactions among biomolecules give rise to the assembly of diverse biomolecular condensates. These interactions are commonly referred to as "multivalent," since biomolecules can associate through multiple interaction sites via a variety of forms of noncovalent bonding. Key players include conformationally heterogeneous proteins, such as intrinsically disordered proteins (IDPs), multidomain proteins containing intrinsically disordered regions (IDRs), and nucleic acids. In the context of IDPs, multivalency refers to the ability of an unfolded protein to engage in many contacts with nearby proteins in a condensed phase. Folded domains within multidomain proteins can also contribute to the multivalency required to drive LLPS, either through protein–protein interactions (PPIs) or, in the case of RNA binding domains (RBDs), through interactions with RNA. Finally, nucleic acid mixtures can phase separate under certain conditions due to intermolecular base-pairing and nonspecific association. The strengths of all these types of interactions are typically comparable to the thermal energy, as the protein and nucleic acid constituents of many biomolecular condensates can remain fluid on biologically relevant timescales.

The mechanisms by which biomolecular condensates form have primarily been studied through the lens of near-equilibrium thermodynamics. Within this view, the partitioning of biomolecules into phase-separated condensates is determined by equilibrium chemical potentials, while condensate (dis)assembly dynamics are governed by free-energy gradients close to equilibrium and/or transitions between metastable states. Predictions based on this near-equilibrium assumption generally hold up well when tested against both *in vitro* and *in vivo* experiments on LLPS. We therefore focus here on near-equilibrium approaches for describing biomolecular condensation. Emphasis is placed on simple equilibrium models that can be used to understand the general features of phase separation in biopolymeric mixtures.

2 Essential thermodynamics

2.1 Conditions for phase equilibria

Recall that for two phases to coexist at equilibrium, they must satisfy the following conditions:

- Equal temperatures (T), since the phases are in contact and can exchange thermal energy. Equal temperatures imply that there is no net heat transfer at equilibrium.
- Equal chemical potentials (μ), since molecules can move across the interfaces between phases. The chemical potential of every molecular species must be the same in all coexisting phases in order for the net diffusive fluxes to vanish.

- Equal pressures (P), since the interfaces between phases can move as the phases expand or contract. Equal pressures imply mechanical stability of the interfaces between coexisting phases at equilibrium.

A phase-separated mixture must also be stable with respect to density fluctuations in any phase, meaning that the free energy increases in response to perturbations away from equilibrium. This condition requires that the second derivative of the free energy with respect to the molecular concentrations is positive for all possible concentration fluctuations.

2.2 Chemical potential and Widom insertion

Of the three coexistence conditions listed above, the notion of equal chemical potentials is often the most poorly understood. How should we think about chemical potential in the context of phase separation?

Most fundamentally, the chemical potential is the reversible work required to insert a molecule into a system. It can therefore be defined as the derivative of the Helmholtz free energy, F, with respect to the number of particles, N,

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{V,T}. \tag{1}$$

In systems at constant temperature, the chemical potential therefore provides a thermodynamic driving force to change the concentration of a molecular species.

Further insight can be gained by considering how the free energy of a system changes when an additional molecule is introduced. We start from the relationship between the Helmholtz free energy and the classical partition function, Q, for a system containing N monomers,

$$F = -k_{\mathrm{B}}T \ln Q = -k_{\mathrm{B}}T \ln \left\{ \frac{[q(T)]^N}{N!\Lambda^{3N}} \int dr^N \exp\left[-\beta U(r^N)\right] \right\}, \tag{2}$$

where $q(T)$ is the internal partition function of an isolated monomer, Λ is its thermal wavelength, $\beta \equiv 1/k_{\mathrm{B}}T$, and $U(r^N)$ is the potential energy of a configuration of N monomers, r^N. The chemical potential of an ideal system, in which $U(r^N) = 0$, is

$$\mu^{\mathrm{id}} = \mu^{\circ}(T) - k_{\mathrm{B}}T \ln\left(\frac{V}{\Lambda^3 N}\right) = \mu^{\circ}(T) + k_{\mathrm{B}}T \ln(\Lambda^3 \rho), \tag{3}$$

where $\mu^{\circ}(T) \equiv -k_{\mathrm{B}}T \ln q(T)$. The excess chemical potential of a system with nonzero density can then be defined as $\mu^{\mathrm{ex}} \equiv \mu - \mu^{\mathrm{id}}$. By writing the derivative $(\partial F/\partial N)_{V,T}$ as a finite difference, the excess chemical potential can be expressed as an ensemble average evaluated in a system of N monomers,

$$\mu^{\mathrm{ex}} = \mu - \mu^{\mathrm{id}} = -k_{\mathrm{B}}T \ln\left\{\frac{Q(N+1,V,T)}{Q(N,V,T)}\right\} - \mu^{\mathrm{id}} \tag{4}$$

$$= -k_{\mathrm{B}}T \ln\left\{\frac{\int dr^{N+1} \exp\left[-\beta U(r^{N+1})\right]}{V \int dr^N \exp\left[-\beta U(r^N)\right]}\right\} \tag{5}$$

$$= -k_{\mathrm{B}}T \ln\left\{\frac{\int dr^N \left\{V^{-1} \int dr_{N+1} \exp\left[-\beta \Delta U_{N+1}(r_{N+1})\right]\right\} \exp\left[-\beta U(r^N)\right]}{\int dr^N \exp\left[-\beta U(r^N)\right]}\right\} \tag{6}$$

$$= -k_{\mathrm{B}}T \ln\left\langle V^{-1} \int dr_{N+1} \exp\left[-\beta \Delta U_{N+1}(r_{N+1})\right]\right\rangle_{N,V,T}. \tag{7}$$

This expression implies that the excess chemical potential is directly related to the ensemble-averaged change in the system potential energy, ΔU_{N+1}, when a "test particle" is inserted at a randomly chosen position r_{N+1} within the system. This expression, known as the *Widom insertion method*, provides a conceptually appealing and practically useful approach for computing chemical potentials.

Consider what this expression implies if we attempt to move a macromolecule from one phase to another. If the phase into which the macromolecule is being inserted is already dense in macromolecules and/or the inserted molecule is very large, then the chance that a random insertion will lead to a steric clash will be high,

and the excess chemical potential will be positive. However, if the interactions that the inserted molecule makes with other constituents of the phase more than compensate by reducing the potential energy, then the excess chemical potential may be negative. Because increasing the concentration of any molecular species always results in a positive ideal contribution to its chemical potential, achieving phase coexistence between phases with different macromolecular concentrations requires that the excess chemical potential of a molecule in the higher-concentration phase is sufficiently negative to result in equal chemical potentials at coexistence.

3 Flory–Huggins model

3.1 Derivation via Widom insertion

A widely used starting point for describing biomolecular phase separation is the Flory–Huggins model, a standard mean-field model of homopolymer phase behavior. Despite its simplicity, the Flory–Huggins model often provides a reasonably accurate description of phase coexistence in many macromolecular solutions. This model is also useful for exploring the phase behavior of multicomponent systems, which can be extraordinarily complex. While there are many ways to derive this model, the Widom approach is particularly attractive because of the physical intuition that it provides.

Let us consider a lattice with a lattice constant equal to the diameter of a monomer. No more than one monomer is allowed to reside at any single lattice site, and empty lattice sites represent the solvent. We define the volume fraction occupied by the monomers to be

$$\phi \equiv \rho v L, \tag{8}$$

where ρ is the number density (i.e., concentration) of polymers, v is the volume associated with a lattice site, and L is the degree of polymerization (i.e., the number of monomers per polymer). To apply the Widom method, we imagine inserting a polymer into a lattice that is already occupied by polymers at volume fraction ϕ. In this case, we evaluate the *exchange chemical potential* between the polymer and the solvent, since polymer insertion requires the removal of a volume vL of solvent. The excess contribution to the exchange chemical potential is

$$\mu^{\mathrm{ex}} = -k_{\mathrm{B}}T \ln \left[p_{\mathrm{overlap}} e^{-\infty} + (1 - p_{\mathrm{overlap}}) e^{-\beta \Delta U} \right] \tag{9}$$

$$= -kT \ln(1 - p_{\mathrm{overlap}}) + \Delta U, \tag{10}$$

where p_{overlap} is the probability that a randomly inserted polymer chain overlaps with one or more occupied lattice sites. If we assume that the monomer positions are uncorrelated (which is equivalent to neglecting the connectivity of the polymer chain), we obtain mean-field expressions for p_{overlap},

$$1 - p_{\mathrm{overlap}} = (1 - \phi)^L, \tag{11}$$

and the change in the energy due to replacing the volume Lv of solvent (S) with L monomers (P),

$$\Delta U = Lz \left[\phi(\epsilon_{\mathrm{PP}} - \epsilon_{\mathrm{SP}}) + (1 - \phi)(\epsilon_{\mathrm{SP}} - \epsilon_{\mathrm{SS}}) \right] \tag{12}$$

$$= Lz \left[(\epsilon_{\mathrm{PP}} - 2\epsilon_{\mathrm{SP}} + \epsilon_{\mathrm{SS}})\phi + (\epsilon_{\mathrm{SP}} - \epsilon_{\mathrm{SS}}) \right] \tag{13}$$

$$= -2k_{\mathrm{B}}TL\chi\phi + \mathrm{const.} \tag{14}$$

In the energy calculation, the constants ϵ_{PP}, ϵ_{SP}, and ϵ_{SS} represent the energy of nearest-neighbor monomer–monomer, solvent–monomer, and solvent–solvent contacts, respectively, on the lattice. The constant

$$\chi \equiv z \left[\frac{2\epsilon_{\mathrm{SP}} - (\epsilon_{\mathrm{PP}} + \epsilon_{\mathrm{SS}})}{2k_{\mathrm{B}}T} \right], \tag{15}$$

is known as the Flory χ parameter.

The exchange chemical potential of the polymer is found by summing the ideal contribution, $\mu^{\mathrm{id}} = k_{\mathrm{B}}T \ln \phi$, and the mean-field approximation for the excess contribution,

$$\mu = k_{\mathrm{B}}T \left[\ln \phi - L \ln(1 - \phi) - 2L\chi\phi \right]. \tag{16}$$

Note that the concentration-independent component of the ideal chemical potential can be ignored here, since it is the same in both phases. The osmotic pressure, P, can then be obtained from the Gibbs–Duhem relation,

$$P = \int d\rho\, \rho \left(\frac{\partial \mu}{\partial \rho}\right)_T = \int d\phi \left(\frac{\phi}{vL}\right) \left(\frac{\partial \mu}{\partial \rho}\right)_T. \tag{17}$$

In dimensionless form, the osmotic pressure is

$$\frac{Pv}{k_\mathrm{B}T} = L^{-1} \int d\phi \left[1 - \frac{L\phi}{1-\phi} - 2L\chi\phi\right] \tag{18}$$

$$= -\ln(1-\phi) + \left(\frac{1}{L} - 1\right)\phi - \chi\phi^2. \tag{19}$$

It is easy to see that this expression reduces to the pressure of an ideal solution in the limit $\phi \ll 1$,

$$\frac{Pv}{k_\mathrm{B}T} = \phi + \left(\frac{1}{L} - 1\right)\phi + \ldots \approx \frac{\phi}{L} = \rho v. \tag{20}$$

3.2 Phase behavior

The Flory–Huggins model predicts coexistence between a dilute phase, which is depleted in polymer, and a condensed phase, which is enriched in polymer, as a function of the Flory χ parameter. For two phases to be in coexistence, it must be possible to find two values of ϕ that have the same chemical potential at constant χ. The critical χ value, at which the model predicts a transition from single-phase to two-phase behavior, occurs where the coexistence volume fractions $\phi^{(\mathrm{dilute})}$ and $\phi^{(\mathrm{condens})}$ merge. To locate this point, we solve

$$0 = \frac{\partial \beta\mu}{\partial \phi} = \frac{1}{\phi} + \frac{L}{1-\phi} - 2L\chi, \tag{21}$$

and look for the value of χ at which the discriminant of this quadratic equation is zero,

$$0 = \mathrm{Discrim}\left[2L\chi\phi^2 - 2L\left(\chi - \frac{1}{2} + \frac{1}{2L}\right)\phi + 1\right] \tag{22}$$

$$= 4L^2\left(\chi - \frac{1}{2} + \frac{1}{2L}\right)^2 - 8L\chi. \tag{23}$$

Solving for χ yields the critical value

$$\chi_\mathrm{c} = \frac{(1+\sqrt{L})^2}{2L}. \tag{24}$$

Inserting this back into Eq. (21) then yields the critical volume fraction,

$$\phi_\mathrm{c} = \frac{1}{1+\sqrt{L}}. \tag{25}$$

In the absence of polymerization, $L = 1$, the Flory–Huggins model is equivalent to the *regular solution model*, for which $\chi_\mathrm{c} = 2$ and $\phi_\mathrm{c} = 1/2$. However, for long polymers, $L \gg 1$, the critical χ parameter tends to a constant, $\chi_\mathrm{c} \to 1/2$, and the critical volume fraction tends to zero, $\phi_\mathrm{c} \to 0$. The behavior in this limit may be surprising, since it seems to suggest that phase separation can occur at vanishingly low concentrations. However, this result may be understood more intuitively by considering the pervaded volume of a single chain in dilute solution. Since the Flory–Huggins model is mean-field, the configurational ensemble of a single isolated polymer is given by that of an ideal chain, for which the radius of gyration scales as $L^{1/2}$. The *overlap concentration*, at which the pervaded volumes of the isolated chains fill all space, thus scales as $\rho^* \sim 1/(vL^{3/2})$. In terms of the overlap volume fraction $\phi^* = vL/(vL^{3/2}) = L^{-1/2}$, we see that $\phi_\mathrm{c}/\phi^* \sim 1$. Thus, the critical point in this limit is located approximately at the volume fraction at which the chains first touch one another. The condensed phase is *semidilute*, meaning that the polymers have overlapping pervaded volumes, so that the chains interpenetrate in the condensed phase. This behavior should be contrasted with

that of a *polymer melt*, which implies that the solvent is completely excluded, such that $\phi^{\text{condens}} \approx 1$. Generally speaking, most naturally occurring biomolecular condensates can be described as semidilute.

We can approximate the binodal away from the critical point by noting that the dilute phase has a low polymer concentration, and thus $Pv/k_{\text{B}}T \approx 0$. Let us consider the case of long polymers, $L \gg 1$. In this limit, we are justified in Taylor expanding the osmotic pressure with respect to ϕ since the critical volume fraction is very low,

$$0 = \frac{\phi}{L} + \left(\frac{1}{2} - \chi\right)\phi^2 + \frac{\phi^3}{3} + \dots. \tag{26}$$

We can then ignore the L-dependent linear term when solving $Pv/k_{\text{B}}T \approx 0$. The condensed-phase volume fraction is thus

$$\phi^{(\text{condens})} \approx 3\left(\chi - \frac{1}{2}\right), \tag{27}$$

with the exchange chemical potential

$$\beta\mu \approx \ln\phi^{(\text{condens})} + L(1-2\chi)\phi^{(\text{condens})} + \frac{L}{2}(\phi^{(\text{condens})})^2 \tag{28}$$

$$= \ln\left[3\left(\chi - \frac{1}{2}\right)\right] - 6L\left(\chi - \frac{1}{2}\right)^2 + \frac{9L}{2}\left(\chi - \frac{1}{2}\right)^2 \tag{29}$$

$$= \ln\left[3\left(\chi - \frac{1}{2}\right)\right] - \frac{3L}{2}\left(\chi - \frac{1}{2}\right)^2. \tag{30}$$

Then, again invoking the approximation that the dilute phase is very dilute, we obtain

$$\phi^{(\text{dilute})} \approx \exp(\beta\mu) = \left[3\left(\chi - \frac{1}{2}\right)\right]\exp\left[-\frac{3L}{2}\left(\chi - \frac{1}{2}\right)^2\right]. \tag{31}$$

Note that the binodal is highly asymmetric about the critical point in this limit.

3.3 Application to biopolymers

Despite the fact that this derivation, strictly speaking, applies to homopolymers, the Flory–Huggins binodal model is routinely used to describe both *in vitro* and *in vivo* experimental measurements of (bio)heteropolymers. When used in this way, the effective χ parameter describes the average interactions among the different types of monomers in a heteropolymer. Consider the example of an IDP, which may contain up to 20 different types of amino acids. (In fact, naturally occurring IDPs are often *low-complexity sequences*, meaning that they are highly enriched in a relatively small number of amino acids.) If we ignore the ordering of the amino acids in the primary sequence and look only at the average amino-acid compositions of the chains, then we can use the same mean-field arguments to write the energetic contribution in Eq. (10) as

$$-k_{\text{B}}T\ln\left\langle\exp\left(\sum_{i=1}^{20} Lc_i\Delta U_i\right)\right\rangle \approx -2k_{\text{B}}TL\left(\sum_{i=1}^{20}\sum_{j=1}^{20} c_i c_j \chi_{ij}\right)\phi = -2k_{\text{B}}TL\chi_{\text{eff}}\phi, \tag{32}$$

where c_i represents the fraction of the L monomers in a heteropolymer that are amino acid type i, and $\sum_{i=1}^{20} c_i = 1$. The resulting χ_{eff} is the sequence composition-averaged effective χ parameter. While the approximation made in this expression is consistent with the mean-field framework, it is in principle possible to incorporate local (i.e., proximal with respect to the primary sequence) sequence patterns. A number of approaches along these lines have recently been proposed to predict or learn the effective χ parameters for naturally occurring IDPs.

4 Associating fluid theory

4.1 Associating polymers

The Flory–Huggins model assumes that the interactions between monomers are isotropic, meaning that they are the same for a pair of monomers that occupy nearest-neighbor lattice sites regardless of their

orientations. However, when describing many types of biopolymeric systems, it is desirable to account for directional interactions between discrete binding sites on some or all monomers. We will assume that each binding site can only participate in one interaction at a time, so that these interactions, which we refer to as *associative*, are mutually exclusive. For example, amino acids can interact via hydrogen bonds, which to a good approximation involve only two amino acids at a time. DNA hybridization and protein–protein interactions between specific interfaces on folded protein domains are other common examples of mutually exclusive interactions that can be viewed as associative interactions.

To construct a mean-field theory of associating polymers, we apply the Widom method to a lattice model in which a fraction f of the monomers in each polymer chain are decorated with associating binding sites,

$$\mu^{\mathrm{ex}} = -k_{\mathrm{B}}T \ln \left\{ p_{\mathrm{overlap}} e^{-\infty} + (1 - p_{\mathrm{overlap}}) \prod_{a=1}^{fL} \left[p_{\mathrm{adj},a} e^{-\beta \Delta U} + (1 - p_{\mathrm{adj},a}) e^{0} \right] \right\}. \tag{33}$$

As before, p_{overlap} is the probability that a randomly inserted polymer overlaps with existing monomers on the lattice. For each of the fL decorated monomers, we randomly choose its orientation and determine the probability, p_{adj}, that its binding site is adjacent to a nearest-neighbor available binding site. Under the mean-field assumption that the monomer orientations are uncorrelated, p_{adj} is given by

$$p_{\mathrm{adj}} = (\phi f) \left(\frac{X}{z} \right), \tag{34}$$

where ϕf is the probability that the nearest-neighbor lattice site is occupied by a monomer with a binding site, X is the conditional probability that the binding site is available (i.e., unbound), and $1/z$ is the conditional probability that it is oriented in the correct direction to form a new bond. The change in energy upon association is then ΔU. The excess chemical potential is thus

$$\mu^{\mathrm{ex}} = -k_{\mathrm{B}}TL \ln(1 - \phi) - k_{\mathrm{B}}TfL \ln \left[\frac{\phi f X}{z} \left(e^{-\beta \Delta U} - 1 \right) + 1 \right]. \tag{35}$$

We now need to solve self-consistently for the unbound fraction of binding sites, X. Given the assumptions that the monomers positions and orientations are uncorrelated, we simply need to establish the condition for chemical equilibrium in an ideal gas of binding sites. In terms of the concentrations of bound sites, [AA], and unbound sites, [A], we can write

$$\frac{[\mathrm{AA}]}{[\mathrm{A}]^2} = K = \frac{v q_{\mathrm{AA}}}{q_{\mathrm{A}}^2} = \frac{1}{2} \times \frac{v}{z} \left(e^{-\beta \Delta U} - 1 \right), \tag{36}$$

where K is the equilibrium constant, and q_{AA} and q_{A} are the internal partition functions of the bound and unbound sites. The final equality follows from the definition of ΔU (where bonds cannot form if $\Delta U = 0$), and the factor $1/2$ is due to the symmetry of an AA dimer. Because every binding site can either be bound or unbound, we can write

$$f\phi = v[\mathrm{A}] + 2v[\mathrm{AA}] = Xf\phi + (Xf\phi)^2 \frac{1}{z} \left(e^{-\beta \Delta U} - 1 \right), \tag{37}$$

leading to

$$X^{-1} = 1 + \frac{Xf\phi}{z} \left(e^{-\beta \Delta U} - 1 \right) = 1 + Xf\phi\Delta, \tag{38}$$

where we have introduced the association parameter $\Delta \equiv z^{-1}(e^{-\beta \Delta U} - 1)$. The excess chemical potential of the associating polymer solution can therefore be simplified to

$$\mu^{\mathrm{ex}} = -k_{\mathrm{B}}TL \ln(1 - \phi) + k_{\mathrm{B}}TfL \ln X. \tag{39}$$

4.2 Phase behavior

To evaluate the phase behavior of this model, we first need to solve the chemical equilibrium equation for the binding sites,

$$X^2 f\phi\Delta + X - 1 = 0. \tag{40}$$

The unbound fraction, which can differ between two coexisting phases, depends on the quantity $f\phi\Delta$,

$$X = \frac{-1 + \sqrt{1 + 4f\phi\Delta}}{2f\phi\Delta}. \tag{41}$$

When there is only one binding site type, as assumed here, one can substitute this expression for X into Eq. (39) to obtain an explicit expression in terms of ϕ that can then be used in phase-coexistence calculations.

It is informative to consider two limiting behaviors. First, if $f\phi\Delta \ll 1$, then the fraction of associated binding sites is small,

$$X \approx 1 - f\phi\Delta. \tag{42}$$

This is the "weak association" limit. Expanding the association contribution to the excess chemical potential, we find

$$\mu^{\mathrm{ex,assoc}} = k_{\mathrm{B}}TfL \ln X \approx k_{\mathrm{B}}TfL \ln(1 - f\phi\Delta) \approx -k_{\mathrm{B}}TL(f^2\Delta)\phi, \tag{43}$$

which looks precisely like the Flory–Huggins model if we identify $\chi = f^2\Delta/2$. Recall that for phase separation to occur in the Flory–Huggins model, we must have $\chi > 1/2$, so $f^2\Delta > 1$. For this condition to be satisfied when $f\phi\Delta \ll 1$, we must have $\phi/f \ll 1$ in both coexisting phases, implying that the critical volume fraction must be very low; if few monomers are decorated with binding sites, such that f is small, then the polymers must be longer still. Therefore, for phase separation of associating polymers to occur in this regime, we must have very long polymers with very many binding sites per polymer! Note that at equilibrium, the probability of any particular binding site being associated is low, even in the condensed phase. Furthermore, we know from the Flory–Huggins result that $\phi^{(\mathrm{condens})} \sim \chi = f^2\Delta$ away from the critical point. Thus, the association parameter must satisfy $f(f^2\Delta)\Delta \ll 1$, which implies that $\Delta \lesssim f^{-3/2}$ for phase separation to take place in the weak association limit.

In the opposite limit where $f\phi\Delta \gg 1$, nearly all the binding sites are associated at equilibrium. This is the "strong association" limit. We now have

$$X \approx \frac{1}{\sqrt{f\phi\Delta}}, \tag{44}$$

resulting in an association contribution to the excess chemical potential

$$\mu^{\mathrm{ex,assoc}} = k_{\mathrm{B}}TfL \ln X \approx -\frac{k_{\mathrm{B}}TfL}{2} \ln f\phi\Delta. \tag{45}$$

Applying the Gibbs–Duhem relation, the corresponding contribution to the dimensionless osmotic pressure is

$$\frac{P^{\mathrm{assoc}}v}{k_{\mathrm{B}}T} \approx -\int d\phi \left(\frac{\phi}{vL}\right) \frac{fL}{2\rho} = -\frac{fL\phi}{2}, \tag{46}$$

which is linear with respect to ϕ and, perhaps surprisingly, independent of Δ. Therefore, in the strong association limit, the phase behavior differs qualitatively from that of the Flory–Huggins model. In particular, the condensed phase volume fraction tends to a constant value $\phi^{(\mathrm{condens})} < 1$ if Δ becomes sufficiently large; by contrast, the Flory–Huggins model always predicts $\phi^{(\mathrm{condens})} \to 1$ as $\chi \to \infty$. It is also possible for the critical point to occur under conditions where most binding sites are already associated to form a system-spanning network, known as a physical gel phase. Such behavior is discussed below.

4.3 Validity of the mean-field assumption

Our discussion so far has assumed a mean-field description in the derivation of the excess chemical potential. However, it is important to consider whether the resulting model is in fact consistent with the implied configurational ensemble of the polymers. Note that polymers behave as ideal chains when the attractive and repulsive (including steric) contributions to the pairwise interactions among the monomers cancel, meaning that the ϕ^2 contribution to the osmotic pressure is small. By Taylor expanding Eq. (19), it is easy to see that this occurs when $\chi \approx 1/2$. The chains can then be treated as random walks, so that the positions and orientations of the monomers are uncorrelated, as assumed in the mean-field calculation of the excess chemical potential. The Flory–Huggins model of homopolymers therefore provides a consistent description

of the condensed phase in the phase-separated regime, since $\chi \gtrsim 1/2$ at the critical point and the chains interpenetrate in the semidilute condensed phase. The mean-field approximations are similarly valid in the condensed phase predicted by the associating polymer model when $f \approx 1$.

Problems can arise in the associating polymer model when $f \ll 1$, meaning that many undecorated monomers are interspersed among the monomers that contain binding sites. These "spacer regions" do not necessarily obey ideal chain statistics. Instead, if the monomers in the spacer regions only experience steric interactions, then they will behave as *swollen* chains, effectively pushing the decorated monomers farther apart and acting to suppress phase separation. To account for this, the associating polymer model should be modified to include isotropic interactions among the monomers as well, leading to

$$\mu^{\rm ex} = -k_{\rm B}TL\ln(1-\phi) + k_{\rm B}TfL\ln X - 2k_{\rm B}TL\chi\phi, \tag{47}$$

where $\chi \approx 1/2$ must be chosen to ensure that the entire heteropolymer behaves as an ideal chain. In other words, weakly attractive interactions among the undecorated monomers are required for the assumption of uncorrelated binding site positions and orientations to be valid when $f \ll 1$.

4.4 Physical gelation

Physical gelation occurs when polymers are reversibly cross-linked to form a system spanning cluster. In the present context, cross-linking refers to the formation of associative bonds between binding sites. Thus, a gel state is reached when we can identify a system-spanning giant cluster of molecules, each of which is connected to the cluster by one or more associative interactions.

To calculate the onset of gelation in our mean-field model, we estimate the conditions for a molecule in the giant cluster to bind with high probability to more than one other molecule. Specifically, for any molecule to be part of the giant cluster, we know that it must be interacting through at least one of its binding sites. The gel point therefore occurs when every molecule within the giant cluster associates, on average, with at least one additional molecule via its remaining binding sites. This condition implies

$$(1-X_{\rm gel})(fL-1) = 1 \implies 1 - X_{\rm gel} = \frac{1}{fL-1}, \tag{48}$$

assuming that the binding sites associate independently. Using the expression for chemical equilibrium among the binding sites, we can write the volume fraction at the gel point, $\phi_{\rm gel}$, as

$$\phi_{\rm gel} = \frac{1}{f\Delta}\frac{(fL-1)}{(fL-2)^2}. \tag{49}$$

Below $\phi_{\rm gel}$, there are many small clusters in a *sol* "phase", but the probability of finding a system-spanning cluster is essentially zero. Above $\phi_{\rm gel}$, the probability of there being a system-spanning cluster goes rapidly to one, and the system enters a *gel* "phase". Importantly, physical gelation is a continuous transition that can occur either independently of or concomitantly with phase separation. It is extremely important not to confuse a gel transition with phase coexistence! System-spanning clusters formed via physical gelation cannot be described as phase separated, because there is no coexistence between different regions of the system with different molecular concentrations.

4.5 Accounting for folding, oligomers, and non-pairwise interactions

While the mean-field assumption underlying the associating polymer model are often reasonable for describing a semidilute condensed phase, it is typically a poor approximation for describing the dilute phase. This is especially the case in the strong association limit. In this regime, we should also consider the possibility that the binding sites on a polymer might associate intramolecularly, or, especially if there are different types of binding sites on different polymer molecules, via the formation of finite-size oligomers. These association mechanisms typically dominate at low polymer concentrations, in which case many bonds can be formed with a minimal loss of translational entropy by involving a very small number of polymers.

In principle, one could describe either folding via intramolecular associations or the assembly of finite-size oligomers by accounting for correlations among the binding sites within the associating polymer framework.

However, a simpler approach that can also yield useful insights is to treat specific folded or oligomeric states as distinct molecular species. The relative populations of these species can then be determined by enforcing the conditions for chemical equilibrium among all possible "conformational states" of a molecule.

As an example, let us consider the scenario of intramolecular folding in the strong association regime. In this case, an isolated polymer in the dilute phase will minimize its free energy by forming intramolecular associative interactions. We refer to this as the folded (F) state. Polymers in the folded state cannot participate in associative interactions with other molecules, so the associative contribution to the excess chemical described previously applies only to the population of unfolded (U) molecules. Eq. (39) therefore becomes

$$\mu_{\mathrm{U}}^{\mathrm{ex}} = -k_{\mathrm{B}}TL\ln(1-\phi_{\mathrm{U}}-\phi_{\mathrm{F}}) + k_{\mathrm{B}}TfL\ln X_{\mathrm{U}} \tag{50}$$

$$\mu_{\mathrm{F}}^{\mathrm{ex}} = -k_{\mathrm{B}}TL\ln(1-\phi_{\mathrm{U}}-\phi_{\mathrm{F}}), \tag{51}$$

where ϕ_{U} and ϕ_{F} are the volume fractions of the U and F states. Chemical equilibrium among the binding sites on U-state polymers requires

$$X_{\mathrm{U}}^{-1} = 1 + X_{\mathrm{U}} f \phi_{\mathrm{U}} \Delta. \tag{52}$$

In addition, the ideal chemical potential of each species now depends on the contributions $\mu_{\mathrm{F}}^{\circ}(T)$ and $\mu_{\mathrm{U}}^{\circ}(T)$ due to the different internal partition functions of the folded and unfolded polymers. We define $\Delta G \equiv \mu_{\mathrm{F}}^{\circ} - \mu_{\mathrm{U}}^{\circ}$ to be the "internal" chemical potential difference between the F and U states implied by the "chemical reaction" U $\rightleftharpoons$ F. This reaction is at equilibrium when the chemical potentials of the two conformational states U and F are equal,

$$\mu_{\mathrm{U}} = \mu_{\mathrm{F}}. \tag{53}$$

Note that at dilute concentrations, Eq. (53) simplifies to an expression of chemical equilibrium for a reaction in an ideal solution, $\phi_{\mathrm{U}} \ll 1$ and $\phi_{\mathrm{F}} \ll 1$,

$$k_{\mathrm{B}}T\ln\phi_{\mathrm{U}} = k_{\mathrm{B}}T\ln\phi_{\mathrm{F}} + \Delta G, \tag{54}$$

leading to

$$\frac{\phi_{\mathrm{F}}}{\phi_{\mathrm{U}}} = e^{-\beta\Delta G}. \tag{55}$$

The phase behavior of this multistate model depends on the behavior of ΔG, which is a temperature-dependent quantity that reflects both the associative interactions in the folded state and the entropic penalty due to folding a polymer to satisfy the associative bonds internally. ΔG typically decreases as temperature is lowered, favoring associative bond formation over the maximization of the polymer configurational entropy. The chemical equilibrium between the U and F conformations thus tends to favor the F state in the dilute phase as temperature is lowered. More interestingly, in the strong association limit, each molecule can participate in $\approx fL/2$ associative interactions in both the dilute phase, via the F conformation, and in the condensed phase, via the U conformation. The equilibrium between the two phases then depends sensitively on the entropic contributions to the chemical potential in the each phase. In fact, it is possible to observe a re-entrant transition at constant polymer concentration, in which a phase-separated solution remixes at low temperatures. Re-entrance occurs in this case when the entropic cost of folding becomes sufficiently small, favoring the folded state and increasing the concentration of the dilute phase.

4.6 Application to biopolymers

The associating polymer model has been widely applied to describe the phase separation of IDPs, multidomain proteins, RNA-binding protein mixtures, and nucleic acid solutions. Applications to IDPs typically invoke the weak association limit, since the interactions between individual pairs of amino acids tend to be on par with or weaker than the thermal energy. The strong interaction limit is more appropriate when describing specific one-to-one protein–protein interactions between the folded domains of multidomain proteins, as well as associative interactions involving nucleic acid hybridization. In these latter cases, the equilibrium constants corresponding to protein–protein, protein–RNA/DNA, and RNA/DNA–RNA/DNA, associative interactions can be readily measured and used to predict the phase behavior. Key features of the associating polymer model in the strong interaction limit, including the existence of ultra low density condensed phases and re-entrant phase transitions due to the presence of dilute-phase oligomers, have been observed in various biopolymeric solutions.

5 Multicomponent phase behavior

5.1 An extremely brief survey of multicomponent thermodynamics

In a multicomponent fluid, phase coexistence is established by the conditions of equal temperatures, pressures, and chemical potentials for *all* components across *all* phases. In general, the coexistence concentrations in a multicomponent fluid are not specified uniquely without also prescribing the overall, or "parent", concentrations in the mixture, $\{\rho_i^{(\text{parent})}\}$ for $i = 1, \ldots, N$. The connection between the parent and coexisting-phase concentrations is provided by the conservation law

$$\rho_i^{(\text{parent})} = \sum_{\alpha=0}^{K} x^{(\alpha)} \rho_i^{(\alpha)}(\{\mu_j\}) \quad \forall i, \tag{56}$$

where α indexes the phases in a phase-separated state with $K+1$ phases, the concentrations $\{\rho_i^{(\alpha)}\}$ indicate coexisting phases with coexistence chemical potentials $\{\mu_j\}$, the volume fractions of the bulk phases are given by $\{x^{(\alpha)}\}$, and $\sum_{\alpha=0}^{K} x^{(\alpha)} = 1$. (This indexing convention is chosen because we are often interested in phase equilibria involving a solvent-majority phase, $\alpha = 0$.) Eq. (56) simplifies to the lever rule for binary mixtures.

In multicomponent mixtures, it is often desirable to characterize the extent to which a particular molecular species is inhomogeneously distributed among the coexisting phases at equilibrium. The *partition coefficient*, PC, is defined as the ratio of a molecule's concentration inside (in) and outside (out) of a particular phase and is directly related to the excess chemical potential of a molecular species i,

$$\text{PC}_i \equiv \frac{\rho_i^{(\text{in})}}{\rho_i^{(\text{out})}} = \exp\left(\beta\mu_{\text{ex},i}^{(\text{out})} - \beta\mu_{\text{ex},i}^{(\text{in})}\right). \tag{57}$$

Partition coefficients are experimentally accessible and biologically relevant quantities, since they quantify the tendency of specific biomolecules to partition spontaneously into phase-separated condensates.

Unlike binary mixtures, there is typically no unique critical point in a multicomponent fluid. Instead, multicomponent critical points lie on a temperature-and-concentration-dependent manifold with dimension one less than the number of non-solvent components in an incompressible fluid. Higher-order critical points, where more than two phases simultaneously merge into a single stable phase, are also possible.

5.2 Multicomponent Flory–Huggins model

Generalizing the derivation of the binary (i.e., one macromolecular component plus solvent) Flory–Huggins model leads to expressions for the chemical potentials and osmotic pressure in an incompressible fluid,

$$\frac{\mu_i}{k_\text{B}T} = \ln\phi_i - L_i\ln(1-\phi_\text{T}) + L_i\sum_{j=1}^{N}\epsilon_{ij}\phi_j \tag{58}$$

$$\frac{Pv}{k_\text{B}T} = -\ln(1-\phi_\text{T}) + \sum_{i=1}^{N}\frac{\phi_i}{L_i} - \phi_\text{T} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ij}\phi_i\phi_j, \tag{59}$$

respectively, where L_i is the degree of polymerization of polymeric species i and $\phi_\text{T} \equiv \sum_{i=1}^{N}\phi_i$ is the total macromolecular volume fraction. The solvent-occupied volume fraction, ϕ_0, is determined by the incompressibility constraint, $\sum_{i=0}^{N}\phi_i = 1$. Here we use dimensionless interaction parameters $\{\epsilon_{ij}\}$, which are negative when macromolecules attract one another. These interaction parameters are measured relative to a reference state of isolated macromolecules in solution, so that $\epsilon_{00} = \epsilon_{0i} = 0$ for all $i > 0$. "Homotypic" and "heterotypic" interactions are encoded in the on- and off-diagonal elements of $\{\epsilon_{ij}\}$, respectively. The contribution to the free-energy density from the pairwise interactions can also be written in terms of Flory χ parameters, $\chi_{ij} = \epsilon_{ij} - (\epsilon_{ii} + \epsilon_{jj})/2$, by extending the sums in the final terms in these equations to include the solvent (component 0) and replacing ϵ_{ij} with χ_{ij}. This change of variables introduces terms that are linear in $\{\phi_i\}$ into the free-energy density, which have no effect on the phase behavior. With this alternate notation, the on-diagonal elements $\{\chi_{ii}\}$ are zero by definition, and the homotypic interactions are encoded by the interactions with the solvent, $\{\chi_{i0}\}$.

Two non-solvent components are sufficient to reveal generic effects of homotypic versus heterotypic interactions. In such a mixture, two distinct types of phase transitions can occur: A "condensation" transition can occur if attractive heterotypic interactions are comparable to or stronger than any attractive homotypic interactions, while a "demixing" transition can occur if the heterotypic interactions are significantly less attractive than one or both of the homotypic interactions. Condensation transitions are analogous to LLPS in simple one-component-plus-solvent fluids, implying that the phase diagram can be fully described by projecting the concentrations onto the parent composition vector. By contrast, mixtures with dissimilar homotypic and heterotypic interaction strengths have more complex phase diagrams.

5.3 Random and designed pairwise interactions

The multicomponent Flory–Huggins model is a natural starting point for exploring how the presence of many distinct molecular components influence the phase behavior of a mixture. However, progress cannot be made without specifying the form of the interaction matrix, and yet limited systematic data exist for parameterizing heterotypic interactions. To deal with this lack of information, *random mixture models* have been introduced to model the $\{\epsilon_{ij}\}$ matrices based on assumed or inferred statistical properties of the interactions in a complex mixture. A proposed statistical distribution defines an ensemble of "random mixtures", such that each mixture in the ensemble corresponds to a particular realization of an $N \times N$ random interaction matrix.

To illustrate this approach, let us consider the problem of predicting instabilities in low-concentration polymer solutions. The central idea is that unstable concentration fluctuations can be determined from a linear stability analysis of the concentration fluctuations,

$$\beta v^{-1} \frac{\partial \mu_i}{\partial \rho_j} = \beta \frac{\partial \mu_i}{\partial \phi_j} L_j = \frac{\delta_{ij}}{v\rho_i} + \frac{L_i L_j}{1-\phi_{\mathrm{T}}} + L_i L_j \epsilon_{ij}. \tag{60}$$

By considering the special case of equimolar parent concentrations, $\rho_i^{(\mathrm{parent})} = \bar{\rho}^{(\mathrm{parent})}\, \forall i$, at low concentrations, we can determine the conditions for stability with respect to concentration fluctuations,

$$\beta v^{-1} \frac{\partial \mu_i}{\partial \rho_j} \approx \frac{\delta_{ij}}{v\bar{\rho}} + L_i L_j (\epsilon_{ij} - 1) \succ 0, \tag{61}$$

where the notation $\succ 0$ indicates that the left-hand side must be a positive definite matrix. According to this expression, instabilities occur when the minimum eigenvalue of the matrix $\{L_i L_j \epsilon_{ij}\}$ is less than $-1/v\bar{\rho}^{(\mathrm{parent})}$. Applying results from random matrix theory, it can be shown that the existence and nature of the dominant instability, which coincides with the minimum eigenvalue of $\{L_i L_j \epsilon_{ij}\}$, can be determined from the mean, b, and standard deviation, σ, of a Gaussian distribution of matrix elements in the limit of large N. If the standard deviation among the matrix elements is sufficiently small, such that $N^{1/2} b/\sigma \lesssim -1$, then the dominant instability involves concentration fluctuations that are parallel to the equimolar parent concentration vector. This type of instability is consistent with a condensation transition driven by similar homotypic and heterotypic interaction strengths. By contrast, if the standard deviation among the matrix elements is sufficiently large, such that $N^{1/2} b/\sigma \gtrsim -1$, then the dominant instability is orthogonal to the parent concentration vector, and individual components demix into phases with vastly different compositions. Importantly, these behaviors are self-averaging, meaning that the tendency of any particular random-mixture realization to undergo a condensation or demixing transition converges in probability as $N \to \infty$.

Applying analyses along these lines to predict instabilities and phase coexistence in multicomponent mixtures remains an active area of research. However, a key challenge is that the existence of an instability is a sufficient but not necessary condition for phase separation to occur at a particular set of parent concentrations. Sophisticated numerical methods are therefore required to identify the conditions for phase coexistence in multicomponent mixtures.

An alternative approach is instead to design the pairwise interaction matrix $\{\epsilon_{ij}\}$ to find combinations of homotypic and heterotypic interactions that yield coexisting phases with desired properties, such as specified molecular compositions. This *inverse design* approach entails working out constraints on the solution space of interaction matrices by enforcing the conditions of equal pressures, chemical potentials, and stability with respect to concentration fluctuations across K target phases. In practice, this inverse design approach provides a numerically tractable method for investigating the relationship between the properties of pairwise interaction matrices and multiphase coexistence in complex macromolecular mixtures.

6 Interfaces

6.1 Square-gradient functional

We now consider the effects of spatial inhomogeneities, which occur at the interfaces between fluid phases and between fluids and surfaces. In what follows, we will assume that the perturbation that gives rise to a spatial inhomogeneity is relatively weak, such that only long-wavelength variations need to be considered. More precisely, we assume

$$\frac{|\nabla\rho(r)|}{\rho_0} = \frac{1}{\xi} \ll \frac{1}{\xi_0}, \tag{62}$$

where ξ is the length scale associated with the perturbation and ξ_0 is the intrinsic correlation length of the bulk system. The intrinsic correlation length can be determined from the radial distribution function (RDF), which describes two-point correlations in a bulk fluid. At large intermolecular separations, the positions of molecules are uncorrelated, and the RDF approaches unity. The intrinsic correlation length is the characteristic distance over which the RDF deviates from unity. This coarse-graining is known as the *local density approximation*. Importantly, the local density approximation implies that an equilibrium free energy can be defined on a length scale $\gtrsim \xi_0$, which is necessarily larger than the size of an individual molecule.

Within the local density approximation, we can write the free energy of the system as a functional of the local number density $\rho(r)$,

$$\mathcal{F}[\rho] = \int dr\, f[\rho(r)]. \tag{63}$$

The integrand $f[\rho(r)]$ is the free-energy density within a ξ^3 volume of the fluid centered at r. For simplicity, let us assume that the inhomogeneity depends only on one spatial coordinate, z. We can then expand the free-energy density, $f[\rho(r)]$, in powers of the inverse range of the inhomogeneity, $1/\xi$. Since $d\rho(z)/dz \sim 1/\xi$, the higher-order derivatives of ρ correspond to higher powers of $1/\xi$,

$$f = f_0 + f_1 \frac{d\rho(z)}{dz} + f_{2'} \left(\frac{d\rho(z)}{dz}\right)^2 + f_{2''} \frac{d^2\rho(z)}{dz^2} + \dots \tag{64}$$

The first-order term must vanish by symmetry, since if we flip the sign of z, then we must retain the same physics with regard to f. Consequently, $f_1 = 0$. Next, we integrate the second-derivative term by parts and drop the boundary term,

$$\int_{-\infty}^{\infty} dz\, f_{2''}[\rho] \left(\frac{d^2\rho}{dz^2}\right) = f_{2''} \left(\frac{d\rho}{dz}\right)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dz \left(\frac{\delta f_{2''}}{\delta\rho} \frac{d\rho}{dz}\right) \left(\frac{d\rho}{dz}\right) = \int_{-\infty}^{\infty} dz\, f_{2'''} \left(\frac{d\rho}{dz}\right)^2. \tag{65}$$

We can therefore combine the second-order derivative into the square-gradient term to arrive at the *square-gradient functional*,

$$\mathcal{F}[\rho] = \int_{-\infty}^{\infty} dz \left[f_0 + f_2 \left(\frac{d\rho(z)}{dz}\right)^2 \right], \tag{66}$$

where f_0 and f_2 are, in general, both functionals of $\rho(z)$. This expression can be generalized to describe inhomogeneities in three dimensions by replacing $d\rho/dz$ with $\nabla\rho$,

$$\mathcal{F}[\rho] = \int dr \left[f_0 + f_2 \left(\nabla\rho(r)\right)^2 \right]. \tag{67}$$

Note that this functional is, strictly speaking, only valid in the long-wavelength limit for which density perturbations vary slowly in space.

The coefficient f_2 can be interpreted either in a microscopic sense, by matching the long-wavelength fluctuations to the two-point correlations in a microscopic model, or in a macroscopic sense, by relating this coefficient to the macroscopic surface tension between two phases. Here we will take the latter approach. Consider the formation of a planar interface, perpendicular to the z axis, between a dilute phase and a condensed phase. At $z = -\infty$, we assume that the system is in the dilute phase, while at $z = \infty$, the system is in the condensed phase. Both bulk phases have the same osmotic pressure P and exchange chemical

potential μ. Moreover, the number density profile $\rho(z)$ must have the same chemical potential *everywhere* in order to be at equilibrium. We therefore work in the grand canonical ensemble, in which we specify μ, V, and T, and seek to minimize the grand potential functional,

$$\Omega[\rho] = \int_{-\infty}^{\infty} dz \left[f_0 + f_2 \left(\frac{d\rho(z)}{dz} \right)^2 \right] - \int_{-\infty}^{\infty} dz\, \rho(z)\mu, \tag{68}$$

with respect to the density profile $\rho(z)$. In what follows, we shall assume that f_2 is a constant. Eq. (68) is minimized by solving the Euler–Lagrange equation

$$\frac{\delta f_0[\rho]}{\delta \rho(z)} - 2f_2 \frac{d^2\rho(z)}{dz^2} = \mu, \tag{69}$$

which can be rewritten in terms of the grand potential density $\omega \equiv f_0 - \mu\rho$,

$$\frac{\delta\{f_0[\rho] - \mu\rho(z)\}}{\delta\rho(z)} \equiv \frac{\delta\omega[\rho]}{\delta\rho(z)} = 2f_2 \frac{d^2\rho(z)}{dz^2}. \tag{70}$$

Integrating by parts, we obtain

$$-\omega[\rho] + f_2 \left(\frac{d\rho(z)}{dz} \right)^2 = P, \tag{71}$$

where the constant of integration, P, is chosen to enforce the boundary conditions

$$\lim_{z\to\pm\infty} \omega[\rho] = f_0\left(\rho^{(\text{dilute})}\right) - \mu\rho^{(\text{dilute})} = f_0\left(\rho^{(\text{condens})}\right) - \mu\rho^{(\text{condens})} = -P. \tag{72}$$

Integrating a second time yields

$$z = -f_2^{1/2} \int_{\rho(0)}^{\rho(z)} d\rho\, [P + \omega(\rho)]^{-1/2}. \tag{73}$$

Given a model for the grand potential density, we can solve Eq. (73) and invert it to obtain the equilibrium interfacial profile $\rho(z)$.

We can now evaluate the surface tension, which is equal to the total additional free energy per unit area relative to the bulk free energy, f_{B},

$$\gamma = \int_{-\infty}^{\infty} dz \left\{ f_0[\rho] + f_2 \left(\frac{d\rho}{dz} \right)^2 - f_{\text{B}} \right\} = \int_{-\infty}^{\infty} dz \left\{ \omega[\rho] + \mu\rho + f_2 \left(\frac{d\rho}{dz} \right)^2 - f_{\text{B}} \right\}, \tag{74}$$

where the bulk free-energy density is $f_{\text{B}} = f(\rho^{(\text{condens})})$ in the condensed phase and $f_{\text{B}} = f(\rho^{(\text{dilute})})$ in the dilute phase. Since the pressure is the same in both bulk phases, we can use the relation $-P = f_{\text{B}} - \mu\rho_{\text{B}}$, as well as the expression for $\omega[\rho]$ that minimizes the grand potential functional, Eq. (71), to write

$$\gamma = \int_{-\infty}^{\infty} dz \left\{ -P + \mu\rho(z) + 2f_2 \left(\frac{d\rho}{dz} \right)^2 - f_{\text{B}} \right\} \tag{75}$$

$$= \int_{-\infty}^{\infty} dz \left\{ \mu[\rho(z) - \rho_{\text{B}}] + 2f_2 \left(\frac{d\rho}{dz} \right)^2 \right\} \tag{76}$$

$$= 2f_2 \int_{-\infty}^{\infty} dz \left(\frac{d\rho}{dz} \right)^2 = 2f_2 \int_{-\infty}^{\infty} d\rho \left(\frac{d\rho}{dz} \right). \tag{77}$$

Finally, using the parametric solution for the density profile, Eq. (73), we obtain

$$\gamma = 2f_2^{1/2} \int_{\rho^{(\text{dilute})}}^{\rho^{(\text{condens})}} d\rho\, [P + \omega(\rho)]^{1/2}. \tag{78}$$

This equation is useful because it shows that the surface tension is governed by the free-energy surface, $f_0(\rho)$. To reiterate, this expression is valid assuming that the interface is broad compared to the intrinsic correlation length of the fluid and that the square-gradient coefficient f_2 is a constant.

If we assume that the grand potential per unit volume can be approximated by a quartic function of the density with minima at $\rho^{(\text{dilute})}$ and $\rho^{(\text{condens})}$,

$$\omega(\rho) = \frac{1}{2}C(\rho - \rho^{(\text{dilute})})^2(\rho - \rho^{(\text{condens})})^2 - P, \tag{79}$$

where C is a temperature-dependent constant, then we can solve the parametric equation for the density profile analytically,

$$z = -\left(\frac{m}{C}\right)^{1/2}\int_{\rho(0)}^{\rho(z)} \frac{d\rho}{(\rho^{(\text{condens})} - \rho)(\rho - \rho^{(\text{dilute})})} = -\zeta \ln\left[\frac{\rho(z) - \rho^{(\text{dilute})}}{\rho^{(\text{condens})} - \rho(z)}\right], \tag{80}$$

where $m = f_2/2$ and the characteristic length ζ is defined by

$$\zeta \equiv \frac{(m/C)^{1/2}}{\rho^{(\text{condens})} - \rho^{(\text{dilute})}}. \tag{81}$$

Rearranging the above equation yields the density profile as a function of z,

$$\rho(z) = \frac{1}{2}(\rho^{(\text{condens})} + \rho^{(\text{dilute})}) - \frac{1}{2}(\rho^{(\text{condens})} - \rho^{(\text{dilute})}) \tanh\left(\frac{z}{2\zeta}\right). \tag{82}$$

Note that this equation is anti-symmetric about the dividing surface, z_0. In reality, the fluctuations in the dilute phase are larger than in the condensed phase, which results in a shallower density profile on the dilute-phase side. The source of this discrepancy is our assumption that the square-gradient coefficient f_2 is independent of ρ.

6.2 Wetting and prewetting

In the context of biomolecular condensates, wetting occurs when a condensed phase makes contact with another liquid or solid phase at equilibrium. Wetting can either be partial, in which a condensed phase droplet has a nonzero contact angle with the substrate, or complete, in which the contact angle is zero. In an undersaturated solution, partial wetting results in the formation of a finite adsorbed layer at the substrate, whereas complete wetting causes the width of the adsorbed layer to diverge as the binodal is approached.

The local density approximation can be used to predict the transition between complete and partial wetting. To this end, we define the order parameter ξ, which describes the thickness of the wetted condensed-phase layer,

$$\xi = \int_0^\infty dz\, \left[\frac{\rho(z)}{\rho^{(\text{dilute})}} - 1\right]. \tag{83}$$

We then examine the wetting behavior by minimizing the square-gradient functional

$$\Omega[\rho] = \int_0^\infty dz \left\{\omega[\rho(z)] + f_2\left(\frac{d\rho}{dz}\right)^2 + \rho(z)\phi(z)\right\}, \tag{84}$$

where we have introduced an external field, $\phi(z)$, to account for the interactions between the fluid and the substrate. Since we assumed in the development of the square-gradient functional that thermodynamics applies on a coarse-grained length scale, we can treat the external field as a boundary condition at $z = 0$,

$$\rho(z)\phi(z) = \phi_0(\rho_\text{S})\delta(z), \tag{85}$$

where ρ_S is the density of the fluid at the surface (i.e., $z = 0$) and ϕ_0 represents the potential energy due to the interaction between the surface and the fluid. Consequently, Ω separates into two terms,

$$\Omega[\rho] = \Omega[\rho] + \Omega(\rho_\text{S}). \tag{86}$$

In what follows, we shall assume that $\phi_0(\rho_S)$ takes the phenomenological form

$$\phi_0(\rho_S) = \gamma_0 - \gamma_1\rho_S + \frac{1}{2}\gamma_2\rho_S^2, \tag{87}$$

where γ_1 and γ_2 are assumed to be positive. The parameter γ_1 accounts for the potential energy associated with substrate–fluid attractive forces, while γ_2 accounts for the reduction in cohesive forces within the fluid that result from the disruption of the local structure due to the presence of the substrate. We shall again assume a phenomenological grand potential density given by a quartic potential,

$$\omega(\rho) = \frac{1}{2}C(\rho - \rho^{(\text{dilute})})^2(\rho - \rho^{(\text{condens})})^2 - P. \tag{88}$$

Applying the result from Eq. (80) and the boundary condition at $z = 0$, the equilibrium density profile in the local density approximation is

$$z = -\zeta \ln\left[\left(\frac{\rho(z) - \rho^{(\text{dilute})}}{\rho^{(\text{condens})} - \rho(z)}\right)\left(\frac{\rho^{(\text{condens})} - \rho_S}{\rho_S - \rho^{(\text{dilute})}}\right)\right], \quad \text{where} \quad \zeta \equiv \frac{(m/C)^{1/2}}{\rho^{(\text{condens})} - \rho^{(\text{dilute})}}. \tag{89}$$

The thickness, ξ, is thus related to the condensed/dilute interfacial width, ζ, and the fluid density at the substrate interface, ρ_S, by

$$\xi = -\zeta \ln\frac{\rho^{(\text{condens})} - \rho_S}{\rho_S - \rho^{(\text{dilute})}} \equiv -\zeta \ln\Phi. \tag{90}$$

The equilibrium density profile in the local density approximation can therefore be written as,

$$\rho(z) = \rho^{(\text{dilute})} + \frac{\rho^{(\text{condens})} - \rho^{(\text{dilute})}}{1 + \exp[(z - \xi)/\zeta]}, \tag{91}$$

which guarantees that $\rho(0) = \rho_S$.

By substituting the equilibrium $\rho(z)$ back into Ω and integrating with respect to z, we can solve for the surface excess grand potential, $\Omega^{(s)}$, which is a function of ρ_S and $\phi_0(\rho_S)$,

$$\Omega^{(s)} \equiv \Omega[\rho] - \Omega(\rho^{(\text{dilute})}) = \Omega[\rho] - \Omega(\rho^{(\text{dilute})}) + \Omega(\rho_S) \tag{92}$$

$$= \gamma\left[3\left(\frac{\rho_S - \rho^{(\text{dilute})}}{\rho^{(\text{condens})} - \rho^{(\text{dilute})}}\right)^2 - 2\left(\frac{\rho_S - \rho^{(\text{dilute})}}{\rho^{(\text{condens})} - \rho^{(\text{dilute})}}\right)^3\right] + \gamma_0 - \gamma_1\rho_S + \frac{1}{2}\gamma_2\rho_S^2, \tag{93}$$

where γ is the surface tension at the condensed/dilute interface. Far from the critical point, the dilute phase is very dilute. Thus, when $\rho^{(\text{dilute})} \ll \rho^{(\text{condens})}$, the surface excess grand potential is well approximated by

$$\Delta\Omega^{(s)} = \gamma\left[\frac{3}{(1+\Phi)^2} - \frac{2}{(1+\Phi)^3} - 1 + 6(p_1 - p_2)\left(\frac{\Phi}{1+\Phi}\right) + 3p_2\left(\frac{\Phi}{1+\Phi}\right)^2\right], \tag{94}$$

where we have defined $\Delta\Omega^{(s)} \equiv \Omega^{(s)} - \Omega_0^{(s)}$, as well as the surface excess grand potential at complete wetting,

$$\Omega_0^{(s)} \equiv \gamma + \gamma_0 - \gamma_1\rho^{(\text{condens})} + \frac{1}{2}\gamma_2(\rho^{(\text{condens})})^2, \tag{95}$$

and the reduced adhesion and cohesion parameters, p_1 and p_2, respectively,

$$p_1 \equiv \frac{\gamma_1}{(mC)^{1/2}}\left(\rho^{(\text{condens})} - \rho^{(\text{dilute})}\right)^{-2} \quad \text{and} \quad p_2 \equiv \frac{\gamma_2}{(mC)^{1/2}}\left(\rho^{(\text{condens})} - \rho^{(\text{dilute})}\right)^{-1}. \tag{96}$$

Complete wetting occurs in the limit $\rho^{(\text{condens})} \to \rho_S$, in which case $\Phi \to 0$ and $\xi \to \infty$. On the other hand, partial wetting corresponds to finite ξ.

In order to establish the equilibrium configuration, what remains is to minimize $\Omega^{(s)} = \Omega_0^{(s)} + \Delta\Omega^{(s)}$ with respect to ρ_S; we shall assume that the other parameters ($\rho^{(\text{condens})}$ and ρ_S; C, m, γ_1, and γ_2) are fixed by

the Hamiltonian that describes the fluid–fluid and fluid–substrate interactions. Depending on the values of p_1 and p_2, the minimum of $\Omega^{(s)}$ may occur either at

$$\Phi = 0 \quad \text{or} \quad \Phi = \Phi^* \equiv \frac{\sqrt{(p_2+1)^2 - 4p_1} + p_2 + 1}{2p_1}. \tag{97}$$

If $p_2 < 1$, then the ξ that minimizes $\Delta\Omega^{(s)}$ jumps discontinuously from a finite value to infinity at $p_1 = (p_2+3)(3p_2+1)/16$. If $p_2 \geq 1$, then the ξ that minimizes $\Delta\Omega^{(s)}$ diverges continuously as $p_1 \to p_2$.

The analysis thus far has assumed that the condensed and dilute phases of the fluid are at coexistence. Thus, along the binodal, we predict a (continuous or discontinuous) transition from complete to partial wetting at $\chi_w > \chi_c$. (For convenience and consistency with the Flory–Huggins model discussed earlier, we characterize the distance from the critical point in terms of a Flory χ parameter.) However, as noted above, wetting can still occur in an undersaturated fluid (i.e., $\rho < \rho^{(\text{dilute})}$) if the fluid–substrate adhesive interactions are sufficiently attractive. In this case, a *finite*-thickness fluid layer wets the substrate. Thus, the following scenarios can occur *in an undersaturated fluid:*

- Below χ_w, the thickness diverges as $\rho \to \rho_{G,coex}$.
- Below χ_w, *if the wetting transition is discontinuous*, then a discontinuous *prewetting* transition between thin and thick wetted layers occurs at $\rho < \rho_{G,coex}$. The grand potential barrier between these states decreases as χ decreases, and the transition disappears at the critical prewetting point, χ_{pwc}.
- Above χ_w, the film thickness is always finite.