Distinguishing near- versus off-critical phase behaviors of intrinsically disordered proteins

Gaurav Mitra[a], Souradeep Ghosh[a], Kiersten M. Ruff[a], Ruoyao Zhang[a], Gaurav Chauhan[a,b], and Rohit V. Pappu[a,✉]

[a]Department of Biomedical Engineering and Center for Biomolecular Condensates, Washington University in St. Louis, St. Louis, MO, USA
[b]Department of Chemical Engineering, Indian Institute of Technology, Indore, India

Intrinsically disordered prion-like low complexity domains (PLCDs) drive phase transitions that underlie the biogenesis of different biomolecular condensates. There are suggestions of condensates formed by different proteins being in the vicinity of thermodynamic critical points. To enable quantitative adjudication of near- versus off-critical phase behaviors, we analyzed results from large-scale simulations of an archetypal PLCD that undergoes phase separation coupled to percolation. Mapping of critical points required the use of finite-size scaling and computation of Binder cumulants. We find that the binodal can be demarcated into three distinct regimes. Regime I is farthest from the critical point, with the dilute phase being akin to a gas of dispersed polymers. Regime II lies above the intersection of the overlap line and the dilute arm of the binodal. Here, dilute phases are characterized by heterogeneous cluster-size distributions with heavy tails. In Regimes I and II, dense phases are confined percolated networks. Regime III is closest to the critical point. Here, the dense phase becomes unconfined and the percolated network swells to become system-spanning. In Regime III, the dilute phase also forms a percolated, system-spanning network. Thus, Regime III comprises two interconnected, system-spanning networks. Strikingly, across the entire temperature range, conformational distributions of individual chains within dense and dilute phases closely follow Gaussian chain statistics. Changes to sequence-encoded interactions change critical temperatures and the locations of different regimes. Our results show that regime-specific cluster-size distributions within dilute phases are measurable signatures of near- versus off-critical phase behaviors of condensate-forming systems.

critical points | universality | intrinsically disordered proteins | phase separation | percolation

Correspondence: ***pappu@wustl.edu***

Introduction

Biochemical reactions in living cells are organized in space and time by the reversible formation of membraneless bodies known as biomolecular condensates (1, 2). Phase transitions drive condensate formation (3–5) and multivalence is a defining hallmark of proteins that drive condensation (1, 5, 6). For proteins, multivalence comes in different flavors (7) and this includes: multiple protein-protein interaction domains tethered by intrinsically disordered linkers (5, 8–10), oligomeric proteins comprising folded domains with or without intrinsically disordered regions (IDRs) (11–15), or intrinsically disordered proteins (IDPs) featuring multiple cohesive motifs that provide specificity of associations (16–20).

Phase separation coupled to percolation (PSCP) is the simplest process that explains the totality of the relevant phase transitions that underlie condensate formation (5, 6, 8, 9, 21–25). In this process, multivalence of specific motifs or domains enables reversible associations that lead to the formation of clusters of different sizes in solution. As concentrations increase above a threshold known as the gel point or percolation threshold (c_{perc}), there is a finite probability of forming clusters that can grow to become system-spanning. However, at temperatures below the critical temperature T_{c}, the growth of clusters in solution is typically suppressed because the system crosses the solubility limit (26–30). Hence, the overall free energy of the system is minimized by separation of the entire solution (31) into coexisting dense and dilute phases via a process known as phase separation. For systems where the associations are homotypic, the threshold concentration for phase separation is the saturation concentration or c_{sat}. If a blend of homotypic and heterotypic associations is involved, then the solubility limit is defined by the solubility product (32–34), with different components contributing differently (34, 35). For condensation via PSCP, the macromolecular concentration within dense phases will be above c_{perc} for temperatures away from the critical point, and the dense phase will be a percolated network that resembles a “confined physical gel" (9, 23, 29, 36, 37). In the context of condensate formation, PSCP was first described by Li et al., as “phase separation that is thermodynamically coupled to sol-gel transitions in the droplet phase" (5).

It has been proposed that condensates in living cells have the potential to be in the vicinity of thermodynamic critical points (38–45). Being close to the critical point would imply that condensates behave like systems that are on the edge of stability, characterized by large fluctuations of the order parameter. For condensates, the relevant order parameters are two-fold: they are the density differences of the condensate-driving macromolecules across the coexisting phases, and the degree of clustering and networking of molecules (46). The one-phase regimes for temperatures near the critical point versus far away from the critical point are expected to be very different (41, 47). Thus, the properties of the coexisting dilute phases are likely to be useful and measurable (48, 49) signatures of near- versus off-critical behaviors of condensates in cells. To set up the relevant expectations for a one-component system, we turned to simulations to map the complete phase behavior of A1-LCD, which is an archetypal intrinsically disordered protein. This is the intrinsically disordered prion-like low complexity domain (PLCD) from the protein hnRNP-A1 (Fig. S1) (50).

We first present the procedure used to delineate the critical

regime, distinguishing it from the off-critical regime. We then identify the location of the critical point and demonstrate how the percolation and overlap lines intersect the coexistence curve (binodal). We identify three distinct regimes, which are characterized by distinct, region-specific distributions of clusters in the coexisting dilute phases. Our findings pave the way to assess the differences between near-critical and off-critical phase behaviors.

Results

Details of the simulations and their scope. The coarse-grained lattice-based Monte Carlo (MC) simulations (53–55) employ the LaSSI simulation engine (46, 52), which is based on a generalization of the bond fluctuation model (56, 57). In the coarse-grained simulations, each residue is a single bead on a lattice. A single A1-LCD molecule comprises 137 beads, one for each of the residues. For the inter-residue potentials (Fig. S2), we used a short-range potential function that was parameterized using a Gaussian Process Bayesian Optimization procedure (52, 58). The parameters are transferable unto other PLCDs but not to other IDPs (52, 59). To map the critical region, we used finite-size scaling that relies on the quantification of density fluctuations via analysis of Binder cumulants (60, 61). We used a nine-step protocol (Fig. S3), which is followed sequentially to arrive at a complete characterization of the binodal of A1-LCD. The key considerations are as follows: Obtaining converged estimates of mean densities, density fluctuations, and higher-order moments, which are imperative for mapping critical regions and percolation lines, requires that the simulation cells encompass at least $\sim 6 \times 10^3$ A1-LCD molecules (see analysis in the *SI Appendix* and Fig. S4). The box sizes need to be large because of the high volume fractions near critical points, and the boxes need to be cubic with periodic boundaries to avoid confinement or symmetry-breaking artifacts. As a result, the number of MC steps required to obtain converged estimates exceeds 10^{11}. Based on these considerations, the simulation system includes 10^4 A1-LCD molecules for each simulation temperature T. For every value of T, we performed at least three independent MC simulations with different random seeds. Each simulation comprises 6×10^{11} MC steps. Additional simulations of equivalent length were performed when the Binder cumulant analysis needed to be deployed. For computing binodals as a function of simulation temperature, we used cubic boxes that were 240 lattice units (l.u.) to a side. In Å, this would translate to ~ 960 Å. For the finite-size scaling analysis, we quantified Binder cumulants using sub-boxes whose dimensions span the range of $L \approx 50 - 200$ l.u.

Delineating the critical and off-critical regimes. At a given temperature T, coexisting dense and dilute phases are delineated by an interface of width Δ. In the critical regime, the system is defined by large density fluctuations, and Δ will diverge (44, 62). For each T, the presence of a two-phase system in the simulation cell is evident from the temperature-specific radial density profiles $\phi(r)$ (Fig. 1A,B). The derivative of $\phi(r)$ with respect to r (Fig. 1A) helps delineate the location of the interface. The interface is the length scale over which the derivative is non-zero. For each T, the derivatives were fit to truncated Gaussian distributions. The width of the interface, $\Delta(T)$, is estimated as the full-width-at-half-maximum, and the midpoint of the interface, $r_{\text{mid}}(T)$, is the mode of the Gaussian (Fig. 1A, Fig. S5). Next, we inserted the computed values of $\Delta(T)$ and $r_{\text{mid}}(T)$ into a hyperbolic tangent, Fisk-Widom function (51, 52). This function is then fit to the T-dependent profiles for $\phi(r)$ to estimate $\phi_{\text{dense}}(T)$ and $\phi_{\text{dilute}}(T)$ (Fig. 1B).

The width $\Delta(T)$ shows the presence of two regimes as a function of temperature, with the crossover temperature being $T_{\text{crossover},\Delta} \approx 49.73°$C (Fig. 1C). Above, $T_{\text{crossover},\Delta}$, $\Delta(T)$ increases rapidly with T. This is characteristic of the divergence one expects for the width of the interface as the critical temperature T_c is approached (44, 62). Thus, we used $T_{\text{crossover},\Delta}$ to demarcate the critical and off-critical regimes, with $T > T_{\text{crossover},\Delta}$ being the critical regime.

The interfacial tension is the free energy penalty associated with increasing the area of the interface and it is computed as $\gamma(T) = \frac{k_B T}{(l\Delta(T))^2}$ (Fig. 1D). Here, $l = 4$ is a factor that converts 1 l.u. to 4 Å (see *SI Appendix*). The interfacial tension shows a sharp decrease above $T_{\text{crossover},\Delta}$ and it is consistent with the expectation that $\gamma(T)$ will approach zero as T approaches T_c (44). The calculated values of $\gamma(T)$ are consistent with measured estimates of interfacial tension for condensates (4, 40, 63–65) and for colloid-polymer systems (66).

Mapping the binodal and assessing its accuracy. The totality of $\phi_{\text{dilute}}(T)$ and $\phi_{\text{dense}}(T)$ points is used to generate the computed dilute and dense arms of the coexistence curve (binodal). We compared the binodals computed from simulations with 200 A1-LCD molecules (data from the previous work of Farag et al., (52)) versus 10^4 molecules. For the smaller system, we cannot obtain reliable estimates of the binodal for temperatures above $T_{\text{crossover},\Delta}$. This is because finite size considerations (see *SI Appendix*) require that tens of thousands of molecules be included in the simulation system to access large volume fractions while also sampling the relevant density fluctuations required to demarcate dense and dilute phases. These density fluctuations are quenched in small systems. Hence, above $T_{\text{crossover},\Delta}$, it becomes difficult to demarcate dense and dilute phases using radial density profiles for small systems such as those comprising 200 molecules. In contrast, the 10^4-chain system yields binodals that sample the vicinity of and go around the apparent critical point (Fig. 2A). We compared the accuracy of the computed binodals by converting simulation temperatures to degrees Kelvin and the volume fractions to molar units (see *SI Appendix*). Note that the experiments (open symbols) used absorbance for measuring off-critical coexistence points, and cloud point measurements to map the critical region (67). The cloud point measurements are noisy due to large density fluctuations. Hence, the experiments do not yield an accurate measurement of T_c, but they identify the near-critical regime. The location of the apparent critical point has been inferred in the past by fitting the data to

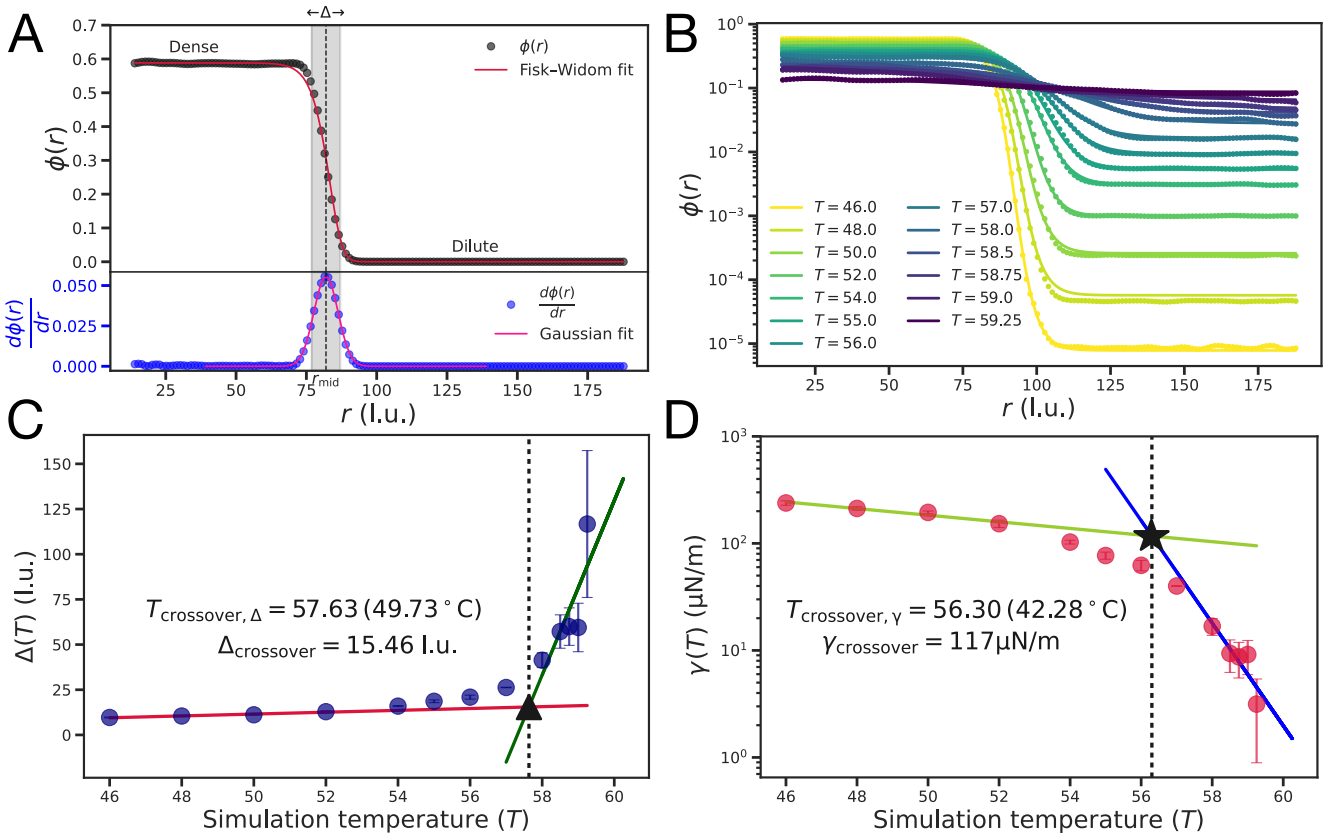

Fig. 1. Demarcation of the off-critical and near-critical regimes. (A) For a given simulation temperature, the derivative of the radial density profile $\phi(r)$ is fitted to a truncated Gaussian distribution, from which the interfacial width Δ is extracted as the full-width-at-half-maximum (FWHM). The midpoint of the interface r_{mid} is the mode of the Gaussian. We insert the computed values of Δ and r_{mid} into a hyperbolic tangent Fisk-Widom function (51, 52). This function is then fit to the $\phi(r)$ profile to estimate the coexisting dense and dilute-phase concentrations (ϕ_{dense} and ϕ_{dilute}). (B) Radial density profiles $\phi(r)$ as a function of r (in lattice units) at different temperatures, which are then fit using the Fisk-Widom function. (C) Temperature dependence of $\Delta(T)$ showing a crossover from the off-critical to the near-critical regime at $T_{\text{crossover},\Delta} \approx 57.60$ (49.7°C), with a corresponding interface width $\Delta_{\text{crossover}} \approx 15.5$ lattice units. (D) A similar crossover is observed for the interfacial energy per unit area $\gamma(T)$, with $T_{\text{crossover},\gamma} \approx 56.30$ (42.3°C) and $\gamma_{\text{crossover}} \approx 117\,\mu\text{N/m}$.

a modified field Flory-Huggins model that includes three-body interactions (17, 67, 68).

To put the comparisons between computed and measured binodals on a quantitative footing, we computed transfer free energies using $\Delta G_{\text{tr}} = -RT\ln\left(\frac{\phi_{\text{dense}}}{\phi_{\text{dilute}}}\right)$ in units of kJmol^{-1} (69). In the regime where the comparisons are feasible, the computed transfer free energies from both the 200-chain and 10^4-chain systems showed good agreement with the measurements (Fig. 2B).

Analysis of Binder cumulants to estimate T_c. The commonly used approach to estimate T_c is to impose the scaling of the order parameter that is in accord with the 3D Ising model (70). This is then used to estimate T_c via regression analysis (71–77). For a 200-chain system, density fluctuations are suppressed above $T_{\text{crossover},\Delta}$, which leads to nonexistent information regarding the binodal above this temperature. Imposing a scaling of the order parameter that is congruent with the 3D Ising model yields poor fits to data-free regions of binodals and incorrect estimates of the location of the critical point (see Fig. S6 and accompanying discussion in *SI Appendix*). While the situation improves, the problems are not alleviated if one uses binodals for the 10^4-chain system. The challenges come down to how points are chosen for regression analysis and to the number of parameters one seeks to extract from this analysis (Fig. S6). To avoid these problems, we used an approach based on Binder cumulants, which rests on computing density cumulants and the use of rigorous finite-size scaling to assess the behavior in the thermodynamic limit where the number of molecules $n \to \infty$ (60, 61, 78–81).

The cumulants quantify the mean density (first moment), density fluctuations (based on the second moment), and the anisotropic nature of the fluctuations (based on the fourth moment). The finite-size scaling uses cubic sub-boxes, each of a different length L. Each sub-box of length L is used to probe the mean density and cumulants on the length scale of the sub-boxes for all temperatures. A schematic of the workflow deployed to calculate and analyze Binder cumulants is shown in Fig. 3.

Probability distributions, $p(\phi)$ (Fig. S7), calculated within sub-boxes of different size L, are shown in Fig. 4(A-C) for three representative simulation temperatures. Below T_c, and for small sub-boxes, the density distributions are bimodal Fig. 4(A-C), reflecting the coexistence of dense and dilute phases. As L increases, the distributions approach a Gaussian form, which would be the canonical expectation for a macroscopically large system. The density distributions become unimodal and Gaussian for all sub-box sizes L as T_c is

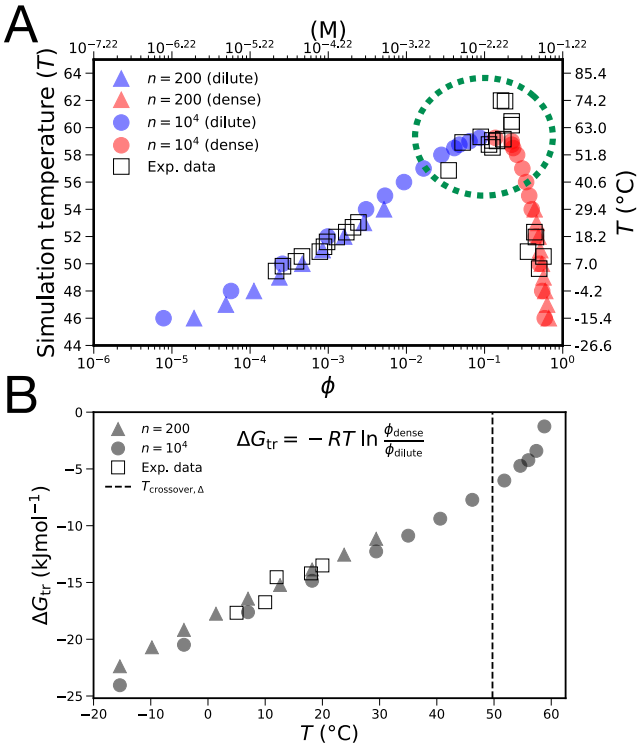

Fig. 2. Comparisons between computed and measured binodals. (A) Binodal of A1-LCD, showing coexistence concentrations of the dilute phase (blue) and dense phase (red). Results from simulations with 10^4 chains are shown as circles, and those from simulations with 200 chains are shown as triangles. The measured coexistence points (67) are shown as open squares. Note the large density fluctuations in the critical regime (indicated by the green dotted oval). In the simulations, the larger system size enables accurate mapping of the critical region, and the binodal is well-rounded near the apparent critical point. (B) Transfer free energies, calculated as $\Delta G_{\text{tr}} = -RT \ln(\phi_{\text{dense}}/\phi_{\text{dilute}})$, are plotted as a function of temperature (in °C). $R = 8.314\,\text{Jmol}^{-1}\text{K}^{-1}$. Triangles correspond to data from the 200-chain simulations, circles to data from 10^4-chain simulations, and squares to measurements. The dashed vertical line corresponds to $T_{\text{crossover},\Delta} \approx 49.7°$C.

approached and crossed (Fig. 4D).

For each sub-box of length L, the variance was computed using $(\Delta\phi_L(T))^2 = m_L^2(T)$ and the fourth-order cumulant was computed using: $U_L(T) = 1 - \langle m_L^4 \rangle / (3\langle m_L^2 \rangle^2)$, where $m_L = (\phi_L(T) - \langle \phi_L(T) \rangle)$ is the order parameter. When the variance and fourth-order cumulant are respectively plotted as a function of T for various sub-box sizes L, curves of $m_L^2(T)$ and $U_L(T)$ intersect at a common point that corresponds to the true T_c (Figs. 4E,F). Note that this analysis was performed for cumulants computed in sub-boxes of size $L \geq 160$. This choice ensures that all density fluctuations are fully accounted for in estimating T_c. In terms of simulation temperatures, the mean and median values of T_c were found to be 59.31 and 59.36, respectively. In degrees Celsius, these values translate to 59.14°C and 59.42°C, respectively (Figs. 4E-F). At the median value of our estimate for T_c we find that $U_L(T) \approx 0.47$. This is congruent with the value of $U^* = 0.466$, which is expected at $T = T_c$ for the 3D Ising model (82).

Using the median value of T_c, we calculated the critical volume fraction ϕ_c via the law of rectilinear diameters (Fig. S8). Usage of this law is valid because the interactions are short-range and the system is in 3-dimensions (83, 84). The temperature-dependent average of the dense and dilute phase volume fraction $\phi_{\text{diam}}(T)$, computed at different temperatures in the critical regime, is fit using linear regression and knowledge of T_c. Substituting $T_c = 59.36$ into the equation: $\phi_{\text{diam}}(T) = \frac{1}{2}[\phi_{\text{dense}}(T) + \phi_{\text{dilute}}(T)] = \phi_c + m(T - T_c)$, we obtained an estimate of $\phi_c \approx 0.12$ as the intercept. The slope was found to be $m = -0.018$ per unit T (Fig. S8).

Finally, having demarcated the critical and off-critical regimes and determined T_c and ϕ_c separately, we analyzed the scaling of the order parameter $\Delta\phi' = \left(\frac{\phi_{\text{dense}} - \phi_{\text{dilute}}}{\phi_c}\right)$ as a function of the distance from the critical point (ϕ_c, T_c). This is written as: $\Delta\phi' = A_o(1 - T/T_c)^\beta$. A regression analysis aimed at simultaneous extraction of β and A_0 (71) becomes problematic because of the unavoidable sparseness of the number of points in the critical regime. Additionally, the order parameter is not symmetric around the critical point. Instead, even in the vicinity of the critical point, the change in the volume fraction along the dilute arm is larger than the change along the dense arm, thus making $\Delta\phi'$ asymmetrical about T_c. Given these considerations, we extracted the value of β by analyzing the ratio $\frac{\ln(\Delta\phi_i/\Delta\phi_j)}{\ln(\Delta T_i/\Delta T_j)}$ for different pairs of points (ϕ_i, T_i) and (ϕ_j, T_j). Here, $\Delta\phi_i = (\phi_{\text{dense},i} - \phi_{\text{dilute},i})$ and $\Delta T_i = (T_c - T_i)$. Note that all points (ϕ_i, T_i) and (ϕ_j, T_j) were drawn from the critical regime defined by $T > T_{\Delta,\text{crossover}}$. This analysis yielded a set of values for β, the mean and median of which were $\beta_{\text{mean}} = 0.3245$ and $\beta_{\text{median}} = 0.3485$, respectively. The value of β should be 0.3264 for the 3D Ising model (85). Thus, through independent assessments, we conclude that the critical regime of A1-LCD belongs to the same universality class as the 3D Ising model, which is in accord with prior expectations (20, 44, 71–77, 86).

Locating the percolation line. The percolation threshold, ϕ_{perc}, is the concentration above which the A1-LCD molecules form a system-spanning network. We computed the threshold concentrations ϕ_{perc} for temperatures above and up to T_c. The locus of ϕ_{perc} values is the percolation line (Fig. S9). We also used the generalized bond percolation model (21) to compute the mean field percolation line that ignores chain connectivity and correlations (see *SI Appendix*). The mean field percolation line intersects the binodal at T_c (29). However, in accord with previous studies (36, 87), we find that the actual percolation line, computed via rigorous clustering analysis (Fig. S9) (46, 88), intersects the binodal at a temperature $T_p = 54.6°\text{C} < T_c$.

The overlap line. Next, we quantified the overlap concentration ϕ^* as a function of T (Fig. S10). Above the overlap concentration, the likelihood of polymers interacting via intermolecular interactions is higher than the likelihood of them interacting via intramolecular interactions (38, 89). The overlap line is the locus of ϕ^* values computed as a function of T. We find that the overlap line intersects the left arm of the binodal at a temperature we designate as T^*. Thus, for $T^* \leq T < T_p$, the dilute phase is actually a semidilute solution (38, 89, 90) with the chain concentration being above the overlap value. For WT A1-LCD, $T^* \approx 27.1°$C.

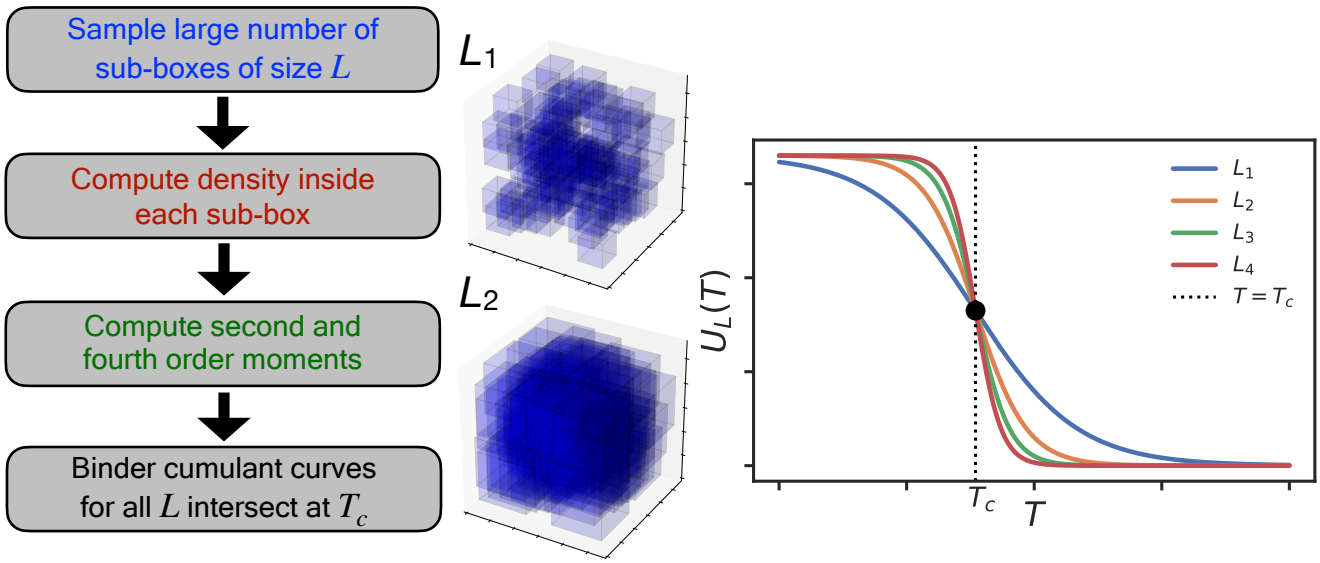

Fig. 3. Workflow for Binder Cumulant Analysis. Schematic illustrating the procedure used to obtain an accurate and independent estimate of the critical temperature T_c. This is based on computations of higher order moments of the order parameter $m_L = (\phi_L(T) - \langle\phi_L(T)\rangle)$, which quantifies the deviation of the density from its mean value within a cubic sub-box of length L at temperature T. A large number of cubic sub-boxes of length L are sampled. The variance and the fourth-order cumulant are obtained from m_L for each of the sampled sub-boxes. The curves for Binder cumulants, computed for different L values, intersect at T_c.

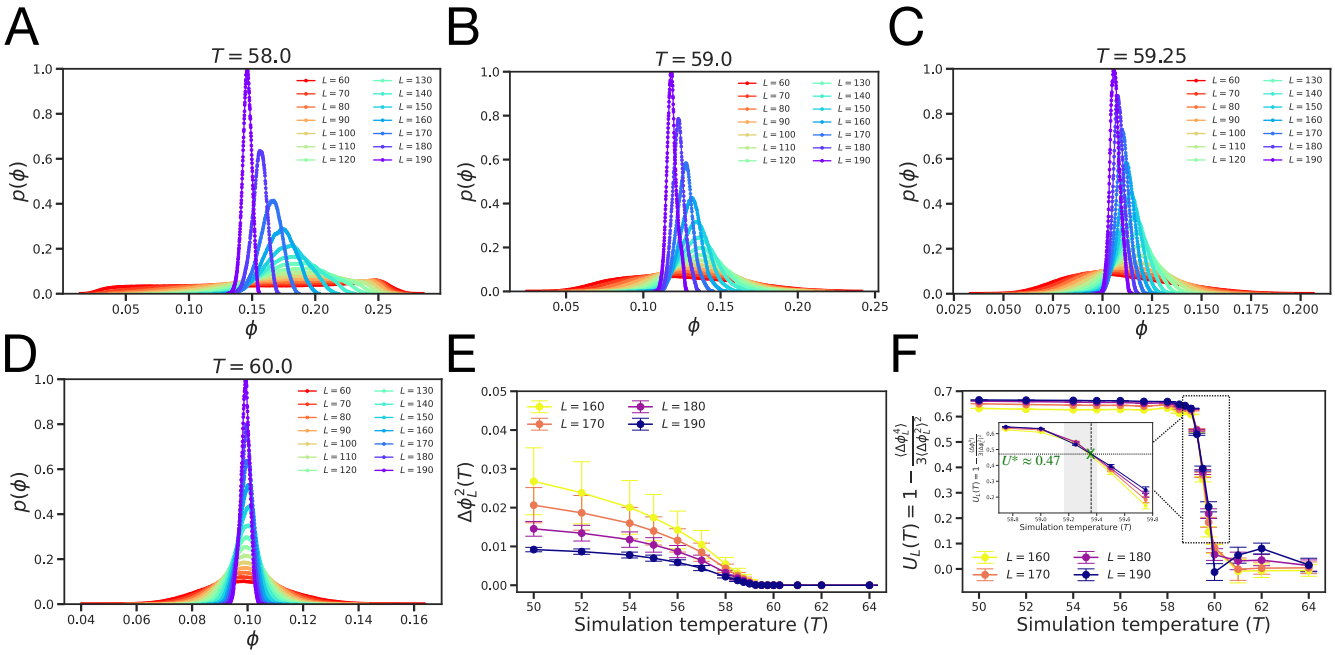

Fig. 4. Distributions of the order parameter and results from Binder cumulant analyses. (A–D) Probability distributions of the sub-box densities $p(\phi)$ at various simulation temperatures T. These distributions are used to calculate the second- and fourth-order moments of the density fluctuations. Below T_c, and for small sub-boxes, the density distributions are bimodal, reflecting the coexistence of dense and dilute phases. As L increases, the distributions approach a Gaussian form. The density distributions become unimodal and Gaussian for all sub-box sizes L as T_c is approached and crossed. (E) The variance of the order parameter plotted as a function of the simulation temperature T for various sub-box sizes L. (F) Binder cumulant $U_L(T)$ as a function of T for different sub-box sizes L. The curves intersect at a median critical temperature of $T_c \approx 59.36$, as shown in the inset. The shaded gray region (in the inset) indicates the range between the upper and lower bounds of the pairwise crossing temperatures, and the green cross marks the intersection point corresponding to the median T_c. At the median value of T_c, the value we obtain for U_L corresponds to $U^* = 0.466$, which is the value that has been determined for U_L at the T_c of the 3D Ising model (82).

The binodal has three regimes. The binodal of A1-LCD (Fig. 5) can be demarcated into three different regimes. These are: Regime I defined by $T < T^*$, Regime II defined by $T^* \leq T < T_p$, and Regime III defined by $T_p \leq T < T_c$ (Fig. 5). In Fig. 5, the label of the left ordinate is T_r, which we define as $T_r = (T_c - T)/T_c$, where T is the simulation temperature. When $T < T_c$, the value of T_r is positive, whereas when $T > T_c$, the value of T_r is negative. In both cases, the magnitude of T_r is a measure of the distance from the critical point. Thus, for A1-LCD, in Regime I, $T_r > 0.1$; in Regime II, $0.02 < T_r \leq 0.1$; and in Regime III, $0 < T_r \leq 0.02$.

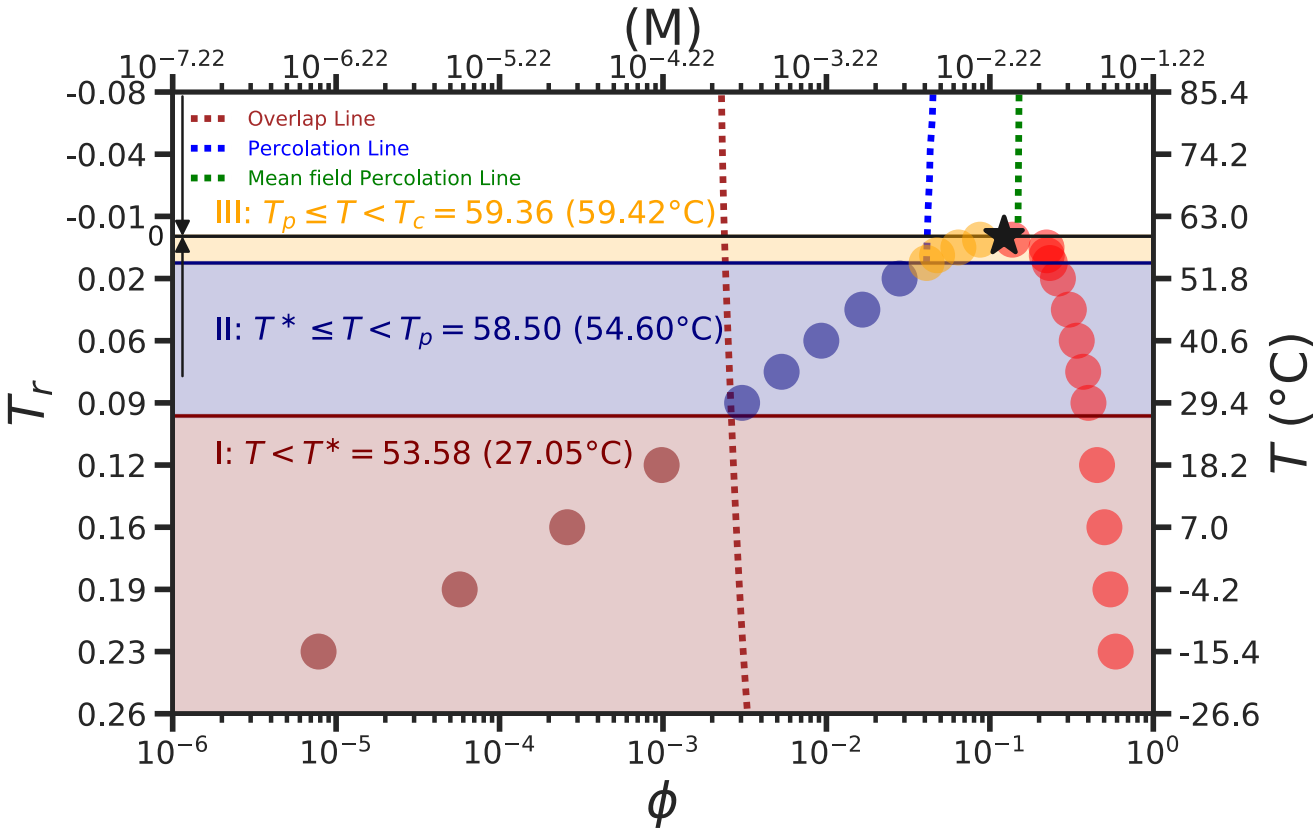

Fig. 5. Fully annotated binodal for WT A1-LCD. The overlap and the percolation lines are indicated by maroon and blue dashed lines, respectively. The mean field percolation line is mapped using the generalized bond percolation model (shown by the green dashed line) (21). The overlap and the percolation lines intersect the left arm of the binodal and divide the binodal into three distinct regimes. The critical point ($\phi_c \approx 0.12, T_c \approx 59.36$), corresponding to $T_r = 0$ is marked by an asterisk. Red circles show the dense phase points, while the dilute phase is shown by circles whose color corresponds to the regime to which the points belong. These regimes are: (i) the low-temperature regime (maroon), at $T < T^*$ (or, $T_r > 0.1$); (ii) the intermediate regime (navy blue), for $T^* \leq T < T_p$ (or, $0.02 < T_r \leq 0.1$); (iii) the high-temperature regime closest to criticality (orange), for $T_p \leq T < T_c$ (or, $0 < T_r \leq 0.02$). The two arrows facing each other are indicative of approaching the critical point ($T_r = 0$) from T_r below and above.

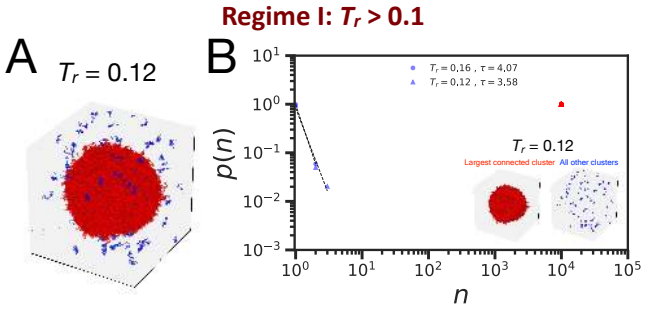

Fig. 6. Regime I of the binodal defined by $T_r > 0.1$. (A) Snapshots showing the dense and dilute phases for Regime I. The dense phase that is representative of the largest connected cluster is colored in red, while the dilute phase is shown in blue. (B) Cluster-size distributions for the dilute phase (shown by blue symbols) and dense phase (shown by red symbols, representing the largest connected cluster). The legend shows the dilute phase τ values obtained at different T_r.

The dense phase is a percolated network. Across the three regimes of the binodal, the concentration of the dense phase changes minimally, by less than a factor of 4, across the entire range of temperatures. Direct calculations of the percolation line and the full binodal show that the right arm of the binodal, which is the locus of ϕ_{dense} values, is located to the right of the percolation line. Thus, $\phi_{\text{dense}} > \phi_{\text{perc}}$ for all $T < T_c$. That $\phi_{\text{dense}} > \phi_{\text{perc}}$ for all $T < T_c$ is proof positive that the phase transition of A1-LCD is consistent with PSCP (52, 67).

Clustering in dilute phases is different across the regimes. In contrast to the dense phase, the concentrations of coexisting dilute phases decrease by more than four orders of magnitude between T_c and the lowest simulation temperature. Demarcation of the binodal into three regimes derives from the intersection of the dilute arm of the binodal by the overlap and percolation lines. Thus, we reasoned that properties of coexisting dilute phases must differ across the three regimes. To test this hypothesis, we analyzed cluster-size distributions in coexisting dilute phases as a function of T.

The probability $p(n)$ of forming clusters containing n molecules can be described using a power-law relation of the form: $p(n) = n^{-\tau}$ (91). The magnitude of τ is a measure of the extent of clustering. Higher values of τ are indicative of smaller clusters being formed, whereas smaller values of τ imply the converse. We analyzed cluster-size distributions that form in coexisting dilute phases as a function of temperature. Using linear regression analysis to analyze these distributions, which required the fitting of $\log_{10} p(n)$ to $-\tau \log_{10} n$, we extracted the temperature dependence of τ (Fig. S11).

In Regime I, the cluster size distribution in the coexisting dilute phase decays steeply as n increases. The distribution is

tailless, and the values of τ that we extract lie in the interval $3.6 \leq \tau \leq 4.1$ (Fig. 6). For all temperatures in the dilute phase of Regime I, the number of molecules within clusters does not go above 5. Thus, the coexisting dilute phase in Regime I is akin to a "gas" of A1-LCD molecules. The single largest cluster is localized to the region of high density (Fig. 6), and this is concordant with the dense phase being akin to a confined physical gel (29). The confinement comes from the high interfacial tension between the coexisting dilute and dense phases, which is the highest in Regime I with $\gamma \approx 200 - 250\,\mu\mathrm{N/m}$ in Regime I.

In Regime II, the coexisting dilute phase is a semidilute solution (38, 89, 90). Individual A1-LCD molecules can overlap with one another, forming finite-sized clusters in the dilute phase. The cluster-size distributions are characterized by exponents that decrease from $\tau = 3.26$ for $T \approx T^*$ to $\tau \approx 2.18$ at $T \approx T_p$ (Fig. 7, Fig. S11). Across Regime II, the cluster-size distributions have heavy tails with exponential cutoffs.

As the concentration of the coexisting dilute phase increases above the overlap line, cluster formation in the dilute phase becomes easier, as evidenced by the decreasing values of τ (Fig. S11). As a result, larger clusters form, and their abundance, quantified by the values of $p(n)$, increases with increasing T. Additionally, in the cluster size range of $\sim 50 < n < 100$, the distributions are characterized by flat tails.

In addition to the formation of heterogeneous distributions of clusters in the dilute phase, Regime II is characterized by an order of magnitude decrease in the interfacial tension which goes from $\approx 100\,\mu\mathrm{N/m}$ at $T \approx T^*$ to less than $10\,\mu\mathrm{N/m}$ as T_p is approached. However, the single largest cluster, which coincides with the dense phase, remains confined for all the temperatures corresponding to Regime II (Fig. 7).

In Regimes I and II, the dense phases are best described as "confined physical gels" (29). This is because $\phi_{\mathrm{dense}} > \phi_{\mathrm{perc}}$, and the fact that the single largest clusters localize to the dense phase (Figs. 6, 7) across the range of temperatures corresponding to the two regimes. The percolated nature of dense phases has been recently corroborated via direct imaging. Quick-freeze deep-etch electron microscopy of a variant of A1-LCD (92) has provided clear evidence of a percolated network-like organization within condensate interiors. Additionally, single-fluorogen imaging of A1-LCD condensates (93) has revealed the inhomogeneous organization of molecules within condensate interiors, thus verifying the prediction of their small-world network structures from LaSSI simulations (52).

Percolation is a continuous networking transition that is different from density transitions (9, 26, 88, 94, 95). It is defined by a distinct critical point that need not coincide with (ϕ_c, T_c), which is the critical point for density transitions. In 3 dimensions, $\tau = \tau_p = 2.18$ is the Fisher exponent that defines the percolation threshold (88). $\phi_p = \phi_{\mathrm{perc}}$ is the value of ϕ where the percolation line intersects the binodal, and hence (ϕ_p, T_p) is the critical point for percolation. As T approaches T_p, we find that τ approaches and crosses the Fisher exponent, $\tau_p = 2.18$ (Fig. S11). The mapping of the percolation line was performed without consideration of the critical exponent that defines the percolation threshold. Thus, analysis of the cluster-size distribution provides independent corroboration of T_p being the temperature where the percolation threshold is crossed.

Taken together, we find that along the dilute arm of the binodal, Regime III ($0 < T_r \leq 0.02$) is bracketed by two critical points (ϕ_p, T_p) and (ϕ_c, T_c). This leads to its designation as the critical regime. In Regime III, both the dense and dilute phases become system-spanning (Fig. 8). Due to the large density fluctuations and low interfacial tension, the dense phase percolated network is no longer confined. It swells to become system-spanning. This swelling of the dense phase network is concordant with inferences made by Smokers and Spruijt from their analysis of how volumes of dense phases change as the critical point is approached and crossed (45). Above T_p, the networks corresponding to the coexisting phases swell, and the two percolated networks become interconnected (Fig. 8).

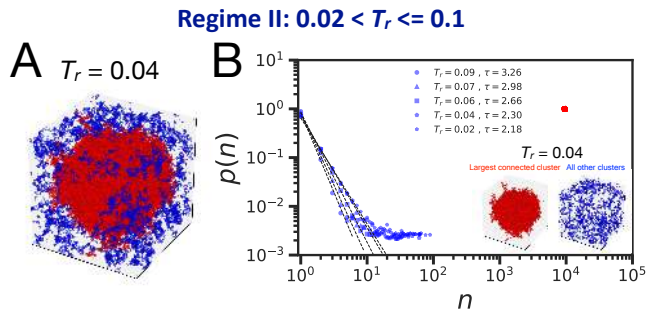

Fig. 7. Regime II of the binodal defined by $0.02 < T_r \leq 0.1$. (A) Snapshots showing the dense and dilute phases for Regime II. (B) Cluster-size distributions for the dilute phase (shown by blue symbols) and dense phase (shown by red symbols). The cluster-size distribution for the dilute phase features a heavy tail of low probability. The legend shows the dilute phase τ values obtained at different T_r.

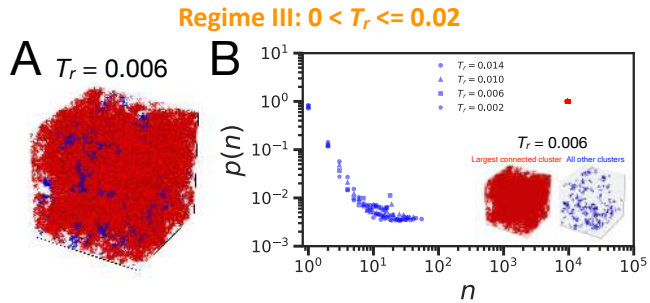

Fig. 8. Regime III of the binodal defined by $0 < T_r \leq 0.02$. (A) Snapshot showing the largest cluster (red) and smaller clusters (blue) that form in Regime III. (B) Distributions of sizes for clusters that lie outside the single largest cluster within Regime III (shown by blue symbols).

Conformations of individual A1-LCD molecules within dense and dilute phases. Conformations of A1-LCD molecules within dense phases are more expanded than in dilute phases (52). (Fig. 9A, Figs. S12, S13). The R_g values in the dense phase are essentially invariant with increasing T, but those of the dilute phase increase with increasing T (Fig. 9A). The scaling of segmental distances within individual molecules shows fractal behavior (Figs. S14, S15), whereby the sizes of segments ($R(s)$) scale with s, the number of residues within a segment, according to the relation $R(s) \sim s^{\nu}$. This behavior is observed in both dense and dilute

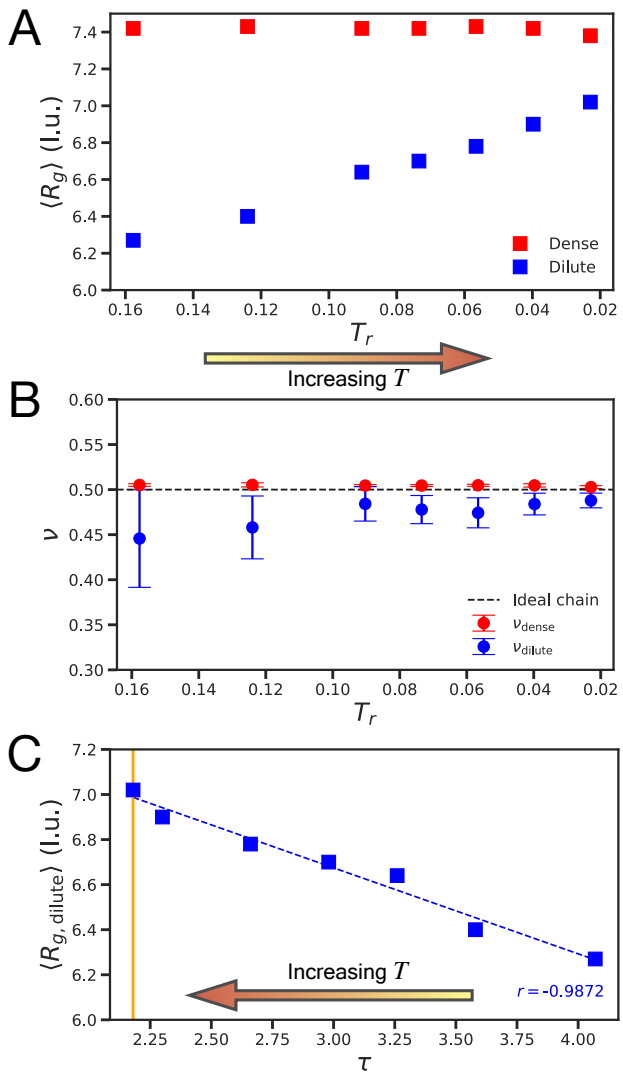

Fig. 9. Temperature dependence of conformational properties of molecules within dilute and dense phases. (A) Dense and dilute phase R_g values as a function of T_r. (B) ν values from analysis of the scaling of $R(s)$ versus s for individual A1-LCD molecules in the dense versus dilute phases. (C) The R_g values of individual A1-LCD molecules in the dilute phase are negatively correlated with τ. The vertical orange line corresponds to $\tau \approx 2.18$, which is the critical exponent for percolation ($d = 3$).

phases. Accordingly, we used scaling analysis to ascertain the value of ν for individual chains extracted from the two phases. For each chain, we selected a series of segment pairs, each comprising s_x and s_y residues, and quantified the distribution of ν values using: $\nu = \left(\frac{\ln(R(s_x))-\ln(R(s_y))}{\ln(s_x)-\ln(s_y)}\right)$ (96). Here, $40 \leq s_x, s_y \leq 120$. These bounds on s_x and s_y help avoid the inclusion of locally stiff segments at lower s values and dangling chain ends at higher s values. The scaling analysis yielded a distribution of ν values (Fig. S16) for individual A1-LCD molecules in dense and dilute phases. The mean values and standard deviations are shown in Fig. 9B.

Within the dense phase, $\nu \approx 0.5$, and this behavior does not change with temperature. Thus, within the dense phase, individual A1-LCD molecules are essentially Gaussian chains (89). The thermal blob (ξ_T), which is the length scale over which Gaussian statistics are followed (97), will span the entirety of a chain. The mesh size ξ_M, which is set by the size of the thermal blob (97), corresponds to the size of an individual chain molecule. Hence, the dense phase may be viewed as a “fractal aggregate" (98) comprising Gaussian chains that individually swell vis-à-vis the dilute phase to enable the high degree of reversible crosslinking (52) that gives rise to an intra-dense-phase percolated network and measurable viscoelasticity (99, 100).

Within the dilute phase, the mean value of ν increases from ≈ 0.44, which is in accord with estimates of ν from small-angle x-ray scattering measurements (17, 67), to ≈ 0.5 as T increases (Fig. 9B, Fig. S16). Thus, even in dilute phases, A1-LCD molecules closely follow Gaussian chain statistics. Deviations from $\nu = 0.5$ (Fig. S16) are in line with expectations for finite-sized polymers (101, 102). Thus, to zeroth order, one can think of dilute phases as comprising Gaussian chains that undergo some degree of compaction (103) as T decreases below T_c. The R_g values in the dilute phase are negatively correlated with τ (Fig. 9C). Chain expansion (larger R_g values) and increased clustering (lower τ values) are correlated in the dilute phase.

The preceding discussion brings us to the topic of the theta temperature (T_θ). This is the point where the intra-chain two-body interaction coefficient $B_{2,\text{intra}} = 0$. In the Flory picture, where intra-chain correlations are ignored, $B_{2,\text{intra}} = 0$ is also the point where the chain is unperturbed and follows Gaussian chain statistics (89). Imposing this mean field expectation as an axiom leads to the erroneous inference that $T_\theta < T_c$. Note that $B_{2,\text{intra}}$ will be negative below T_c. Thus it follows that $T_\theta > T_c$ for a finite-sized polymer (90, 104). Congruence with Gaussian chain statistics cannot be used as a way to identify the T_θ temperature. Instead, mapping of the full binodal, with accurate estimation of ϕ_c and T_c, helps with estimating T_θ via extrapolation. This will require knowledge of how T_c varies as a function of chain length N (105). The upshot is that congruence with Gaussian chain statistics does not imply congruence with the theta temperature (102). Instead, observation of Gaussian chain statistics for individual chains in the dense and dilute phases points to the counterbalancing of intra- and inter-molecular interactions at the level of a single chain. The behaviors we observe may be rationalized by the sequence-encoded hierarchy of interactions that characterize A1-LCD and other PLCDs (Fig. S2) (52, 59, 67). Tyr, Phe, Arg residues form a set of stickers, and the strengths of these inter-sticker interactions range between $\approx 2.5 - 4.0\ k_BT$ (52, 67). These inferences from measurements have been recently validated using simulations based on a polarizable force-field (95). The stickers, which are distributed uniformly and non-randomly along the linear sequence of A1-LCD (17, 106), constitute roughly 20% of the A1-LCD sequence. The remaining residues are spacers. They provide cohesive interactions of $\approx 1 - 2\ k_BT$. These sequence-encoded hierarchies of interactions that are weak or of intermediate strength with respect to k_BT, ensure that the chain does not form globules in dilute solutions. As a result, the mean field picture of a gas of globules in the dilute phase coexisting with a melt-like dense phase ($\phi \approx 1$) (71, 89) is not

realized for A1-LCD. Instead, individual A1-LCD molecules follow Gaussian chain statistics in dense and dilute phases, and this implies that distinct interactions within the coexisting phases are realized while maximizing conformational entropy at the level of individual molecules (89, 97).

Quantifying sequence-encoded impacts on phase behaviors. Bremer et al. (67) designed and characterized a series of variants of A1-LCD. The calculations of full binodals, deploying the full extent of Binder cumulant analysis, become prohibitively expensive to perform across all variants. However, based on favorable comparisons with experiments (Fig. 10A), we used the previously computed binodals in the off-critical regime that were based on simulations of 200-chain systems (52) to understand how sequence changes affect key features of the binodals. We fit the computed binodals to a modified Flory-Huggins (mFH) model (17, 67, 68) (*SI Appendix*). The inferred T_c values, which we refer to as $T_{c,\mathrm{mFH}}$, are formally incorrect because $T_c \neq T_{c,\mathrm{mFH}}$ (Table S1), and the finite-sized nature of the simulations with 200 molecules causes T_c to be underestimated. However, these values can be used as proxies for comparative assessments of how changes to the A1-LCD sequence influence the driving forces for PSCP. We analyzed extant simulation results by fitting the computed binodals to the modified Flory-Huggins model to extract estimates of $T_{c,\mathrm{mFH}}$ and T^*_{mFH}. Larger values of $T_{c,\mathrm{mFH}}$ imply stronger driving forces for PSCP when compared to the wild-type A1-LCD, and lower values of $T_{c,\mathrm{mFH}}$ imply the converse (75, 76).

An increase in $T_{c,\mathrm{mFH}}$ will be accompanied by a leftward shift in the left arm of the binodal, whereas a decrease in $T_{c,\mathrm{mFH}}$ will be accompanied by a rightward shift of the dilute arm of the binodal (Fig. 10B). The overlap lines change minimally across the variants (Fig. S17). This is because all A1-LCD variants are of identical lengths. However, T^*_{mFH} will change with sequence because of the sequence-specific leftward or rightward shifts of the left arm of the binodal, which cause an increase or decrease in $T_{c,\mathrm{mFH}}$ (Figs. S18, S19). For all variants, there is a strong positive correlation between T^*_{mFH} and $T_{c,\mathrm{mFH}}$ (Fig. 10C). The implication is that enhancing the driving forces for PSCP of A1-LCD-like PLCDs will lead to an increase in T_c and T^*. Since the strongest interactions are on the order of 2.5 - 4.0 k_BT (52, 59, 67, 107), PLCDs and related IDPs that drive phase separation via homotypic interactions at room temperatures or 37°C, are likely to be poised to access or fall out of Regime II through subtle changes to sequence-encoded associations or the valence of stickers (67).

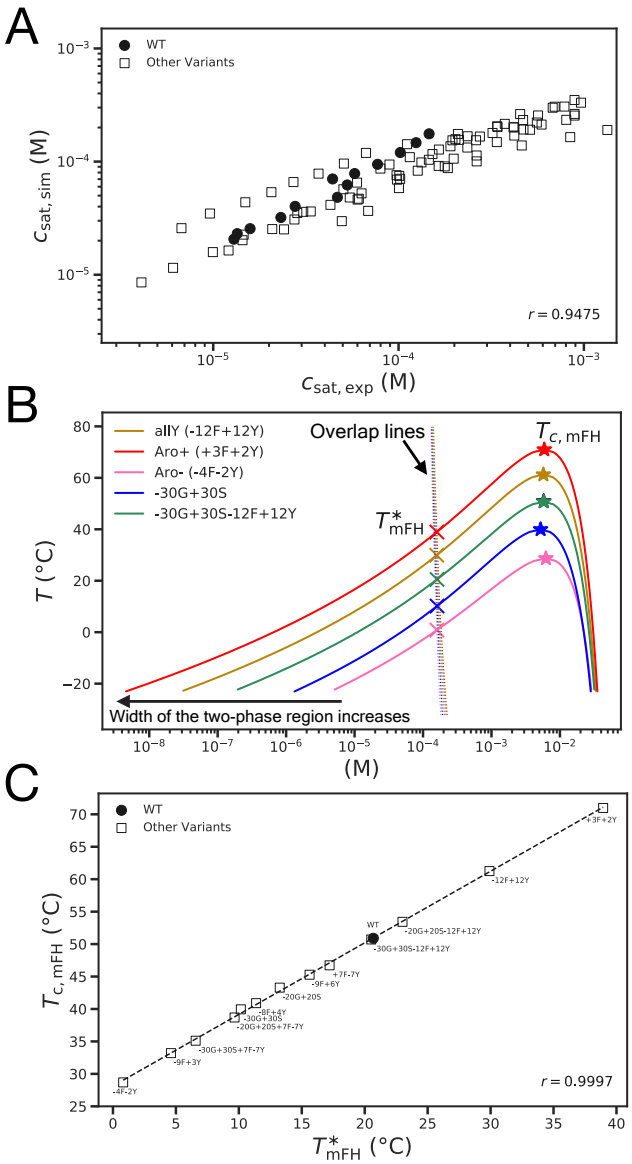

Fig. 10. Assessing the effects of sequence variants of A1-LCD on T^*_{mFH} and $T_{c,\mathrm{mFH}}$. (A) Comparison of measured and computed c_{sat} values. Correlations between measured and computed c_{sat} values are strong and positive, with Pearson correlation coefficients ranging from 0.98 to 0.99 (the global Pearson correlation coefficient across all the different variants and temperatures is 0.9475). (B) Coexistence curves obtained by fitting the modified Flory-Huggins model to the computed binodals for five A1-LCD variants – Aro+ or +3F+2Y (red curve), allY or -12F+12Y (gold-colored curve), -30G+30S-12F+12Y (green curve), -30G+30S (blue curve), and Aro- or -4F-2Y (pink curve). The binodals are annotated by the variant-specific overlap lines, and this shows where the different overlap lines intersect the dilute arms of the corresponding binodals. (C) Correlation plot of T^*_{mFH} and $T_{c,\mathrm{mFH}}$. The Pearson correlation coefficient is 0.9997.

Connecting to experimental observations. Simulation temperatures are useful proxies for effective inter-residue interaction coefficients (B_2) whereby $T_r \propto -B_2$ (6, 38, 75). Lan et al., (49) reported that in unstressed cells, the negative elongation factor (NELF) forms nucleoplasmic clusters that are consistent with what we observe for A1-LCD in the dilute phases of Regime II. Conversely, in stressed cells, the phosphorylation of NELF leads to the formation of macrophases. Setting B_2^* to be the value of the interaction coefficient that corresponds to T^*, we would propose that $|B_2| < |B_2^*|$ for both unphosphorylated NELF and phosphorylated NELF. At endogenous expression levels, unphosphorylated NELF is in the one-phase regime where it forms clusters corresponding to Regime II. Conversely, phosphorylation moves the system into the two-phase regime by lowering B_2, and the coexisting dilute phase is akin to what we observe for Regime II. Signatures of being in Regime III will be the formation of interconnected networks of coexisting dense and dilute phases. This behavior has been observed for the WNK1 kinase that forms condensates in response to osmotic stress (108). The

observed behavior was referred to as spinodal decomposition. Indeed, Regime III is expected to coincide with and be below the spinodal. Thus, the dynamics of phase separation in Regime III will likely follow that of viscoelastic phase separation (109).

Materials and Methods

Key aspects of the methods are summarized in this section. Other methodological details are in the *SI Appendix*.

LaSSI simulations. The simulations were performed following the approach of Farag et al.(52, 59). Each amino acid residue corresponds to a single lattice bead. Beads occupy positions on a cubic lattice, and the energy of interaction is $-\varepsilon_{ij}$ (Fig. S2) for beads of type i and j if the distance between the beads on the lattice is 1, $\sqrt{2}$, or $\sqrt{3}$ l.u. apart. Otherwise, the interaction energy is zero. The values of $-\varepsilon_{ij}$ were those of Farag et al. (52).

The nine-step simulation and analysis protocol. **Step 1** involves setting up simulations of the appropriate size and length (see the analysis in *SI Appendix*). **Step 2** is the extraction of $\Delta(T)$ and r_{mid} from the radial density profiles. **Step 3** uses the T-dependent radial density profiles and the estimated T-dependent values of $\Delta(T)$ and r_{mid} to extract $\phi_{\text{dense}}(T)$ and $\phi_{\text{dilute}}(T)$ using the Fisk-Widom function. **Step 4** delineates the critical and off-critical regimes by analyzing the variation of $\Delta(T)$ with T, quantifying $T_{\text{crossover},\Delta}$, and designating the region above $T_{\text{crossover},\Delta}$ as the critical region. **Step 5** deploys the analysis of Binder cumulants, which requires considerable additional sampling, to estimate T_c. **Step 6** uses the known T_c to calculate ϕ_c using the method of rectilinear diameters. **Step 7** deploys sampling at high densities to extract the temperature-dependent percolation thresholds, the locus of which constitutes the percolation line. **Step 8** estimates the overlap line by performing single-chain simulations at different temperatures to estimate $\phi^*(T)$. **Step 9** involves the parsing of the binodal into distinct regimes based on the values of T^*, T_p, and T_c.

Estimation of the coexistence concentrations and the interface width. To determine the interfacial width Δ and the midpoint of the interface r_{mid}, we analyzed the derivatives of the radial density profile $\phi(r)$. The derivative profile $d\phi(r)/dr$ exhibits a peak centered at the location of the interface. We fit this peak to a truncated Gaussian function (110),

$$\frac{d\phi(r)}{dr} = A \exp\left[-\frac{(r-r_{\text{mid}})^2}{2\sigma^2}\right]$$

where r_{mid} is the midpoint of the interface and σ is the Gaussian width parameter. The full width at half maximum (FWHM) of the fitted Gaussian defines the interfacial width, $\Delta = 2\sqrt{2\ln 2}\,\sigma$. The dilute- and dense-phase concentrations used to construct the binodal are obtained by fitting the radial density profile $\phi(r)$ to the Fisk–Widom function, which interpolates smoothly between the two bulk concentrations and accounts for the interfacial width. The functional form is

$$\phi(r) = \left(\frac{\phi_{\text{dilute}} + \phi_{\text{dense}}}{2}\right) - \left(\frac{\phi_{\text{dense}} - \phi_{\text{dilute}}}{2}\right) \tanh\left(\frac{r - r_{\text{mid}}}{\Delta/2}\right)$$

where ϕ_{dilute} and ϕ_{dense} denote the dilute- and dense-phase concentrations, respectively. The midpoint r_{mid} and interfacial width Δ are those extracted from the derivative fit. The coexistence concentrations ϕ_{dilute} and ϕ_{dense} reported for the binodal are obtained by averaging over three independent replicate simulations for each condition.

Binder cumulants. We used the Binder cumulant analysis to estimate T_c. This involves sampling a large number of sub-boxes of size L within the full simulation box of size L_{sim} and computing the density distribution and moments of the distribution, which include the mean density, the variance, and kurtosis within each sub-box. The variance and kurtosis are then used to evaluate the Binder cumulants. Analysis of the L-dependent cumulants as a function of T leads to an estimate of T_c as the temperature of intersection of the different cumulant curves.

Calculation of the overlap volume fraction from single chain simulations. The overlap volume fraction ϕ^*, which is a function of temperature, is obtained from single-chain simulations by first computing the root-mean-square end-to-end distance,

$$R_{ee}^{\text{rms}} = \sqrt{\langle R_{ee}^2 \rangle}\,,$$

and then estimating the typical coil volume as $(R_{ee}^{\text{rms}})^3$. The overlap volume fraction is defined as

$$\phi^*(T) = \frac{N\,r^3}{\left(R_{ee}^{\text{rms}}(T)\right)^3}\,,$$

where $N = 137$ is the number of residues in a chain and $r = 0.5$ l.u., the radius of individual residues/beads.

Percolation threshold as a function of temperature. To determine the percolation threshold ϕ_{perc} at a given simulation temperature T, we compute the percolation order parameter defined as the fraction of chains belonging to the largest connected cluster, following the approach used in (46).

$$f_{\text{lc}}(\phi) \;=\; \frac{n_{\text{largest}}}{n}$$

Here, $n = 10^4$ is the total number of A1-LCD molecules in the system. The resulting $f_{\text{lc}}(\phi)$ vs. ϕ curve exhibits a sigmoidal shape, characteristic of a percolation transition. Accordingly, we fit the data to a logistic function of the form

$$f_{\text{lc}}(\phi) \;=\; \frac{1}{1+\exp\left[-k\left(\phi-\phi_{\text{perc}}\right)\right]},$$

where k is a parameter describing the steepness of the transition and ϕ_{perc} denotes the percolation threshold. The value of ϕ_{perc} is identified with the inflection point of the fitted logistic curve, corresponding to $f_{\text{lc}}(\phi) = 0.5$. We deploy this

computation of ϕ_{perc} only for state points corresponding to temperatures that lie outside the two-phase region of the binodal for WT A1-LCD.

ACKNOWLEDGEMENTS

We are grateful to Tanja Mittag for stimulating discussions and collaborations, and to Michael Rubinstein for helpful discussions and insights. Computations were performed using facilities of the Washington University Research Computing and Informatics Facility (RCIF) and the facilities of the SEAS High Performance Research Computing. The RCIF is funded by NIH S10 program grants: 1S10OD025200-01A1 and 1S10OD030477-01. This work was supported by grants from the US-Air Force Office of Scientific Research (FA9550-20-1-0241), the US National Institutes of Health (R01NS121114), the US National Science Foundation (MCB-2227268), and the Center for Biomolecular Condensates in the James McKelvey School of Engineering at Washington University in St. Louis.

Bibliography

1. Salman F Banani, Hyun O Lee, Anthony A Hyman, and Michael K Rosen. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.*, 18(5):285–298, May 2017.
2. Yongdae Shin and Clifford P Brangwynne. Liquid phase condensation in cell physiology and disease. *Science*, 357(6357):eaaf4382, September 2017.
3. Clifford P. Brangwynne, Christian R. Eckmann, David S. Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A. Hyman. Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324(5935): 1729–1732, June 2009. ISSN 1095-9203. doi: 10.1126/science.1172046.
4. Clifford P. Brangwynne, Timothy J. Mitchison, and Anthony A. Hyman. Active liquid-like behavior of nucleoli determines their size and shape in xenopus laevis oocytes. *Proceedings of the National Academy of Sciences*, 108(11):4334–4339, February 2011. ISSN 1091-6490. doi: 10.1073/pnas.1017150108.
5. Pilong Li, Sudeep Banjade, Hui-Chun Cheng, Soyeon Kim, Baoyu Chen, Liang Guo, Marc Llaguno, Javoris V. Hollingsworth, David S. King, Salman F. Banani, Paul S. Russo, Qiu-Xing Jiang, B. Tracy Nixon, and Michael K. Rosen. Phase transitions in the assembly of multivalent signalling proteins. *Nature*, 483(7389):336–340, March 2012. ISSN 1476-4687. doi: 10.1038/nature10879.
6. Rohit V Pappu, Samuel R Cohen, Furqan Dar, Mina Farag, and Mrityunjoy Kar. Phase transitions of associative biomacromolecules. *Chem. Rev.*, 123(14):8945–8987, July 2023.
7. Jeong-Mo Choi, Alex S. Holehouse, and Rohit V. Pappu. Physical principles underlying the complex biology of intracellular phase transitions. *Annual Review of Biophysics*, 49(1): 107–133, May 2020. ISSN 1936-1238. doi: 10.1146/annurev-biophys-121219-081629.
8. Salman F. Banani, Allyson M. Rice, William B. Peeples, Yuan Lin, Saumya Jain, Roy Parker, and Michael K. Rosen. Compositional control of phase-separated cellular bodies. *Cell*, 166 (3):651–663, July 2016. ISSN 0092-8674. doi: 10.1016/j.cell.2016.06.010.
9. Tyler S Harmon, Alex S Holehouse, Michael K Rosen, and Rohit V Pappu. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *eLife*, 6, November 2017. ISSN 2050-084X. doi: 10.7554/elife.30294.
10. Dragomir Milovanovic, Yumei Wu, Xin Bian, and Pietro De Camilli. A liquid phase of synapsin and lipid vesicles. *Science*, 361(6402):604–607, August 2018. ISSN 1095-9203. doi: 10.1126/science.aat5671.
11. Diana M. Mitrea, Christy R. Grace, Marija Buljan, Mi-Kyung Yun, Nicholas J. Pytel, John Satumba, Amanda Nourse, Cheon-Gil Park, M. Madan Babu, Stephen W. White, and Richard W. Kriwacki. Structural polymorphism in the n-terminal oligomerization domain of npm1. *Proceedings of the National Academy of Sciences*, 111(12):4466–4471, March 2014. ISSN 1091-6490. doi: 10.1073/pnas.1321007111.
12. David W. Sanders, Nancy Kedersha, Daniel S.W. Lee, Amy R. Strom, Victoria Drake, Joshua A. Riback, Dan Bracha, Jorine M. Eeftens, Allana Iwanicki, Alicia Wang, Ming-Tzo Wei, Gena Whitney, Shawn M. Lyons, Paul Anderson, William M. Jacobs, Pavel Ivanov, and Clifford P. Brangwynne. Competing protein-rna interaction networks control multiphase intracellular organization. *Cell*, 181(2):306–324.e28, April 2020. ISSN 0092-8674. doi: 10.1016/j.cell.2020.03.050.
13. Jordina Guillén-Boixet, Andrii Kopach, Alex S. Holehouse, Sina Wittmann, Marcus Jahnel, Raimund Schlüßler, Kyoohyun Kim, Irmela R.E.A. Trussina, Jie Wang, Daniel Mateju, Ina Poser, Shovamayee Maharana, Martine Ruer-Gruß, Doris Richter, Xiaojie Zhang, Young-Tae Chang, Jochen Guck, Alf Honigmann, Julia Mahamid, Anthony A. Hyman, Rohit V. Pappu, Simon Alberti, and Titus M. Franzmann. Rna-induced conformational switching and clustering of g3bp drive stress granule assembly by condensation. *Cell*, 181(2):346–361.e17, April 2020. ISSN 0092-8674. doi: 10.1016/j.cell.2020.03.049.
14. Samantha K. Powers, Alex S. Holehouse, David A. Korasick, Katherine H. Schreiber, Natalie M. Clark, Hongwei Jing, Ryan Emenecker, Soeun Han, Eric Tycksen, Ildoo Hwang, Rosangela Sozzani, Joseph M. Jez, Rohit V. Pappu, and Lucia C. Strader. Nucleo-cytoplasmic partitioning of arf proteins controls auxin responses in arabidopsis thaliana. *Molecular Cell*, 76(1):177–190.e5, October 2019. ISSN 1097-2765. doi: 10.1016/j.molcel.2019.06.044.
15. SoRi Jang, Zhao Xuan, Ross C. Lagoy, Louise M. Jawerth, Ian J. Gonzalez, Milind Singh, Shavanie Prashad, Hee Soo Kim, Avinash Patel, Dirk R. Albrecht, Anthony A. Hyman, and Daniel A. Colón-Ramos. Phosphofructokinase relocalizes into subcellular compartments with liquid-like properties in vivo. *Biophysical Journal*, 120(7):1170–1186, April 2021. ISSN 0006-3495. doi: 10.1016/j.bpj.2020.08.002.
16. Jie Wang, Jeong-Mo Choi, Alex S. Holehouse, Hyun O. Lee, Xiaojie Zhang, Marcus Jahnel, Shovamayee Maharana, Régis Lemaitre, Andrei Pozniakovsky, David Drechsel, Ina Poser, Rohit V. Pappu, Simon Alberti, and Anthony A. Hyman. A molecular grammar governing the driving forces for phase separation of prion-like rna binding proteins. *Cell*, 174(3): 688–699.e16, July 2018. ISSN 0092-8674. doi: 10.1016/j.cell.2018.06.006.
17. Erik W. Martin, Alex S. Holehouse, Ivan Peran, Mina Farag, J. Jeremias Incicco, Anne Bremer, Christy R. Grace, Andrea Soranno, Rohit V. Pappu, and Tanja Mittag. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science*, 367(6478):694–699, February 2020. ISSN 1095-9203. doi: 10.1126/science.aaw8653.
18. Jamie A. Greig, Tu Anh Nguyen, Michelle Lee, Alex S. Holehouse, Ammon E. Posey, Rohit V. Pappu, and Gregory Jedd. Arginine-enriched mixed-charge domains provide cohesion for nuclear speckle condensation. *Molecular Cell*, 77(6):1237–1250.e4, March 2020. ISSN 1097-2765. doi: 10.1016/j.molcel.2020.01.025.
19. Nicola Galvanetto, Miloš T. Ivanović, Aritra Chowdhury, Andrea Sottini, Mark F. Nüesch, Daniel Nettels, Robert B. Best, and Benjamin Schuler. Extreme dynamics in a biomolecular condensate. *Nature*, 619(7971):876–883, July 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06329-5.
20. Gregory L. Dignon, Wenwei Zheng, Young C. Kim, Robert B. Best, and Jeetain Mittal. Sequence determinants of protein phase behavior from a coarse-grained model. *PLOS Computational Biology*, 14(1):e1005941, January 2018. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1005941.
21. Jeong-Mo Choi, Anthony A. Hyman, and Rohit V. Pappu. Generalized models for bond percolation transitions of associative polymers. *Physical Review E*, 102(4), October 2020. ISSN 2470-0053. doi: 10.1103/physreve.102.042403.
22. Tanja Mittag and Rohit V. Pappu. A conceptual framework for understanding phase separation and addressing open questions and challenges. *Molecular Cell*, 82(12):2201–2214, June 2022. ISSN 1097-2765. doi: 10.1016/j.molcel.2022.05.018.
23. Tharun Selvam Mahendran, Gable M. Wadsworth, Anurag Singh, Ritika Gupta, and Priya R. Banerjee. Homotypic rna clustering accompanies a liquid-to-solid transition inside the core of multi-component biomolecular condensates. *Nature Chemistry*, 17(8):1236–1246, July 2025. ISSN 1755-4349. doi: 10.1038/s41557-025-01847-3.
24. Zeyu Shen, Bowen Jia, Yang Xu, Jonas Wessén, Tanmoy Pal, Hue Sun Chan, Shengwang Du, and Mingjie Zhang. Biological condensates form percolated networks with molecular motion properties distinctly different from dilute solutions. *eLife*, 12, June 2023. ISSN 2050-084X. doi: 10.7554/elife.81907.
25. Zhitao Liao, Bowen Jia, Dongshi Guan, Xudong Chen, Mingjie Zhang, and Penger Tong. Emergent mechanics of a networked multivalent protein condensate. *Nature Communications*, 16(1), June 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-60345-9.
26. Fumihiko Tanaka. Theory of thermoreversible gelation. *Macromolecules*, 22(4):1988–1994, April 1989. ISSN 1520-5835. doi: 10.1021/ma00194a077.
27. Fumihiko Tanaka and Masahiko Ishida. Thermoreversible gelation of hydrated polymers. *Journal of the Chemical Society, Faraday Transactions*, 91(16):2663, 1995. ISSN 1364-5455. doi: 10.1039/ft9959102663.
28. Fumihiko Tanaka. Phase formation of associating polymers: gelation, phase separation and microphase formation. *Advances in Colloid and Interface Science*, 63:23–40, January 1996. ISSN 0001-8686. doi: 10.1016/0001-8686(95)00275-8.
29. Alexander N Semenov and Michael Rubinstein. Thermoreversible gelation in solutions of associative polymers. 1. statics. *Macromolecules*, 31(4):1373–1385, February 1998.
30. Fumihiko Tanaka. Theoretical study of molecular association and thermoreversible gelation in polymers. *Polymer Journal*, 34(7):479–509, July 2002. ISSN 1349-0540. doi: 10.1295/polymj.34.479.
31. Yifan Dai, Zhen-Gang Wang, and Richard N. Zare. Unlocking the electrochemical functions of biomolecular condensates. *Nature Chemical Biology*, 20(11):1420–1433, September 2024. ISSN 1552-4469. doi: 10.1038/s41589-024-01717-y.
32. Aniruddha Chattaraj, Michael L Blinov, and Leslie M Loew. The solubility product extends the buffering concept to heterotypic biomolecular condensates. *eLife*, 10, July 2021. ISSN 2050-084X. doi: 10.7554/elife.67176.
33. Aniruddha Chattaraj and Leslie M. Loew. The maximum solubility product marks the threshold for condensation of multivalent biomolecules. *Biophysical Journal*, 122(9):1678–1690, May 2023. ISSN 0006-3495. doi: 10.1016/j.bpj.2023.03.036.
34. Dan Deviri and Samuel A. Safran. Physical theory of biological noise buffering by multicomponent phase separation. *Proceedings of the National Academy of Sciences*, 118(25): e2100099118, 2021. doi: 10.1073/pnas.2100099118.
35. Daoyuan Qian, Hannes Ausserwoger, Tomas Sneideris, Mina Farag, Rohit V. Pappu, and Tuomas P. J. Knowles. Dominance analysis to assess solute contributions to multicomponent phase equilibria. *Proceedings of the National Academy of Sciences*, 121(33):e2407453121, 2024. doi: 10.1073/pnas.2407453121.
36. Daniel W. Cheong and Athanassios Z. Panagiotopoulos *. Phase behaviour of polyampholyte chains from grand canonical monte carlo simulations. *Molecular Physics*, 103 (21-23):3031–3044, 2005. doi: 10.1080/00268970500186045.
37. Gable M Wadsworth, Walter J Zahurancik, Xiangze Zeng, Paul Pullara, Lien B Lai, Vaishnavi Sidharthan, Rohit V Pappu, Venkat Gopalan, and Priya R Banerjee. RNAs undergo phase transitions with lower critical solution temperatures. *Nat. Chem.*, 15(12):1693–1704, December 2023.
38. Ming-Tzo Wei, Shana Elbaum-Garfinkle, Alex S Holehouse, Carlos Chih-Hsiung Chen, Marina Feric, Craig B Arnold, Rodney D Priestley, Rohit V Pappu, and Clifford P Brangwynne. Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat. Chem.*, 9(11):1118–1125, November 2017.
39. Miguel A Munoz. Colloquium: Criticality and dynamical scaling in living systems. *Reviews of Modern Physics*, 90(3):031001, 2018.
40. Shunsuke F Shimobayashi, Pierre Ronceray, David W Sanders, Mikko P Haataja, and Clifford P Brangwynne. Nucleation landscape of biomolecular condensates. *Nature*, 599 (7885):503–506, 2021.
41. Mason Rouches, Sarah L. Veatch, and Benjamin B. Machta. Surface densities prewet a near-critical membrane. *Proceedings of the National Academy of Sciences*, 118(40), October 2021. ISSN 1091-6490. doi: 10.1073/pnas.2103401118.
42. Anatol W Fritsch, Andrés F Diaz-Delgadillo, Omar Adame-Arana, Carsten Hoege, Matthäus Mittasch, Moritz Kreysing, Mark Leaver, Anthony A Hyman, Frank Jülicher, and Christoph A

Weber. Local thermodynamics govern formation and dissolution of caenorhabditis elegans p granule condensates. *Proceedings of the National Academy of Sciences*, 118(37): e2102772118, 2021.

43. Isabella R. Graf and Benjamin B. Machta. Thermodynamic stability and critical points in multicomponent mixtures with structured interactions. *Phys. Rev. Res.*, 4:033144, Aug 2022. doi: 10.1103/PhysRevResearch.4.033144.
44. Andrew G. T. Pyo, Yaojun Zhang, and Ned S. Wingreen. Proximity to criticality predicts surface properties of biomolecular condensates. *Proceedings of the National Academy of Sciences*, 120(23):e2220014120, May 2023. ISSN 1091-6490. doi: 10.1073/pnas.2220014120.
45. Iris B. A. Smokers and Evan Spruijt. Quantification of biomolecular condensate volume reveals network swelling and dissolution regimes during phase transition. *Biomacromolecules*, 26(1):363–373, December 2024. ISSN 1526-4602. doi: 10.1021/acs.biomac.4c01201.
46. Jeong-Mo Choi, Furqan Dar, and Rohit V Pappu. LASSI: A lattice model for simulating phase transitions of multivalent proteins. *PLoS Comput. Biol.*, 15(10):e1007028, October 2019.
47. Sarah L. Veatch, Olivier Soubias, Sarah L. Keller, and Klaus Gawrisch. Critical fluctuations in domain-forming lipid mixtures. *Proceedings of the National Academy of Sciences*, 104 (45):17650–17655, 2007. doi: 10.1073/pnas.0703513104.
48. Arjun Narayanan, Anatoli Meriin, J Owen Andrews, Jan-Hendrik Spille, Michael Y Sherman, and Ibrahim I Cisse. A first order phase transition mechanism underlies protein aggregation in mammalian cells. *Elife*, 8:e39695, 2019.
49. Chenyang Lan, Juhyeong Kim, Svenja Ulferts, Fernando Aprile-Garcia, Sophie Weyrauch, Abhinaya Anandamurugan, Robert Grosse, Ritwick Sawarkar, Aleks Reinhardt, and Thorsten Hugel. Quantitative real-time in-cell imaging reveals heterogeneous clusters of proteins prior to condensation. *Nature Communications*, 14(1), August 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-40540-2.
50. Amandine Molliex, Jamshid Temirov, Jihun Lee, Maura Coughlin, Anderson P. Kanagaraj, Hong Joo Kim, Tanja Mittag, and J. Paul Taylor. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*, 163 (1):123–133, September 2015. ISSN 0092-8674. doi: 10.1016/j.cell.2015.09.015.
51. S. Fisk and B. Widom. Structure and free energy of the interface between fluid phases in equilibrium near the critical point. *The Journal of Chemical Physics*, 50(8):3219–3227, 04 1969. ISSN 0021-9606. doi: 10.1063/1.1671544.
52. Mina Farag, Samuel R Cohen, Wade M Borcherds, Anne Bremer, Tanja Mittag, and Rohit V Pappu. Condensates formed by prion-like low-complexity domains have small-world network structures and interfaces defined by expanded conformations. *Nat. Commun.*, 13(1):7722, December 2022.
53. Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 06 1953. ISSN 0021-9606. doi: 10.1063/1.1699114.
54. Marshall N. Rosenbluth and Arianna W. Rosenbluth. Monte carlo calculation of the average extension of molecular chains. *The Journal of Chemical Physics*, 23(2):356–359, 02 1955. ISSN 0021-9606. doi: 10.1063/1.1741967.
55. Jörn Ilja Siepmann and Daan Frenkel. Configurational bias monte carlo: a new sampling scheme for flexible chains. *Molecular Physics*, 75(1):59–70, 1992. doi: 10.1080/00268979200100061.
56. I. Carmesin and Kurt Kremer. The bond fluctuation method: a new effective algorithm for the dynamics of polymers in all spatial dimensions. *Macromolecules*, 21(9):2819–2823, September 1988. ISSN 1520-5835. doi: 10.1021/ma00187a030.
57. J. Scott Shaffer. Effects of chain topology on polymer dynamics: Bulk melts. *The Journal of Chemical Physics*, 101(5):4205–4213, 09 1994. ISSN 0021-9606. doi: 10.1063/1.467470.
58. Kiersten M. Ruff, Tyler S. Harmon, and Rohit V. Pappu. Camelot: A machine learning approach for coarse-grained simulations of aggregation of block-copolymeric protein sequences. *The Journal of Chemical Physics*, 143(24), November 2015. ISSN 1089-7690. doi: 10.1063/1.4935066.
59. Mina Farag, Wade M Borcherds, Anne Bremer, Tanja Mittag, and Rohit V Pappu. Phase separation of protein mixtures is driven by the interplay of homotypic and heterotypic interactions. *Nat. Commun.*, 14(1):5527, September 2023.
60. Kurt Binder. Finite size scaling analysis of ising model block distribution functions. *Zeitschrift für Physik B Condensed Matter*, 43(2):119–140, 1981.
61. Kurt Binder. Computer simulations of critical phenomena and phase behaviour of fluids. *Molecular Physics*, 108(14):1797–1815, 2010.
62. B. Widom. Intermolecular forces and the nature of the liquid state. *Science*, 157(3787): 375–382, 1967. doi: 10.1126/science.157.3787.375.
63. Chi W. Pak, Martyna Kosno, Alex S. Holehouse, Shae B. Padrick, Anuradha Mittal, Rustam Ali, Ali A. Yunus, David R. Liu, Rohit V. Pappu, and Michael K. Rosen. Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Molecular Cell*, 63(1):72–85, July 2016. ISSN 1097-2765. doi: 10.1016/j.molcel.2016.05.042.
64. Marina Feric, Nilesh Vaidya, Tyler S. Harmon, Diana M. Mitrea, Lian Zhu, Tiffany M. Richardson, Richard W. Kriwacki, Rohit V. Pappu, and Clifford P. Brangwynne. Coexisting liquid phases underlie nucleolar subcompartments. *Cell*, 165(7):1686–1697, June 2016. ISSN 0092-8674. doi: 10.1016/j.cell.2016.04.047.
65. Holly H. Cheng, James V. Roggeveen, Huan Wang, Howard A. Stone, Zheng Shi, and Clifford P. Brangwynne. Micropipette aspiration reveals differential rna-dependent viscoelasticity of nucleolar subcompartments. *Proceedings of the National Academy of Sciences*, 122(22), May 2025. ISSN 1091-6490. doi: 10.1073/pnas.2407423122.
66. Dirk G. A. L. Aarts, Matthias Schmidt, and Henk N. W. Lekkerkerker. Direct visual observation of thermal capillary waves. *Science*, 304(5672):847–850, 2004. doi: 10.1126/science.1097116.
67. Anne Bremer, Mina Farag, Wade M. Borcherds, Ivan Peran, Erik W. Martin, Rohit V. Pappu, and Tanja Mittag. Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. *Nature Chemistry*, 14(2):196–207, December 2021. ISSN 1755-4349. doi: 10.1038/s41557-021-00840-w.
68. Mina Farag, Alex S. Holehouse, Xiangze Zeng, and Rohit V. Pappu. Fireball: A tool to fit protein phase diagrams based on mean-field theories for polymer solutions. *Biophysical Journal*, 122(12):2396–2403, June 2023. ISSN 0006-3495. doi: 10.1016/j.bpj.2023.05.007.
69. Giulio Tesei, Thea K. Schulze, Ramon Crehuet, and Kresten Lindorff-Larsen. Accurate model of liquid–liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. *Proceedings of the National Academy of Sciences*, 118(44): e2111696118, 2021. doi: 10.1073/pnas.2111696118.
70. Ernst Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, February 1925. ISSN 0044-3328. doi: 10.1007/bf02980577.
71. Kevin S. Silmore, Michael P. Howard, and Athanassios Z. Panagiotopoulos. Vapour–liquid phase equilibrium and surface tension of fully flexible lennard–jones chains. *Molecular Physics*, 115(3):320–327, 2017. doi: 10.1080/00268976.2016.1262075.
72. Gregory L. Dignon, Wenwei Zheng, Robert B. Best, Young C. Kim, and Jeetain Mittal. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences*, 115(40):9929–9934, September 2018. ISSN 1091-6490. doi: 10.1073/pnas.1804177115.
73. Jerelle A. Joseph, Aleks Reinhardt, Anne Aguirre, Pin Yu Chew, Kieran O. Russell, Jorge R. Espinosa, Adiran Garaizar, and Rosana Collepardo-Guevara. Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nature Computational Science*, 1(11):732–743, November 2021. ISSN 2662-8457. doi: 10.1038/s43588-021-00155-3.
74. D. Janka Bauer, Lukas S. Stelzl, and Arash Nikoubashman. Single-chain and condensed-state behavior of hnrnpa1 from molecular simulations. *The Journal of Chemical Physics*, 157(15):154903, 10 2022. ISSN 0021-9606. doi: 10.1063/5.0105540.
75. Yaxin An, Michael A. Webb, and William M. Jacobs. Active learning of the thermodynamics-dynamics trade-off in protein condensates. *Science Advances*, 10(1):eadj2448, 2024. doi: 10.1126/sciadv.adj2448.
76. M Julia Maristany, Anne Aguirre Gonzalez, Jorge R Espinosa, Jan Huertas, Rosana Collepardo-Guevara, and Jerelle A Joseph. Decoding phase separation of prion-like domains through data-driven scaling laws. January 2025. doi: 10.7554/elife.99068.2.
77. Subhadip Biswas and Davit A Potoyan. Decoding biomolecular condensate dynamics: an energy landscape approach. *PLOS Computational Biology*, 21(2):1–19, 02 2025. doi: 10.1371/journal.pcbi.1012826.
78. Michael E. Fisher and Michael N. Barber. Scaling theory for finite-size effects in the critical region. *Phys. Rev. Lett.*, 28:1516–1519, Jun 1972. doi: 10.1103/PhysRevLett.28.1516.
79. Brézin, E. An investigation of finite size scaling. *J. Phys. France*, 43(1):15–22, 1982. doi: 10.1051/jphys:0198200430101500.
80. Daniel Nilsson and Anders Irbäck. Finite-size scaling analysis of protein droplet formation. *Phys. Rev. E*, 101:022413, Feb 2020. doi: 10.1103/PhysRevE.101.022413.
81. Anders Malthe-Sørenssen. *Finite Size Scaling*, page 85–99. Springer International Publishing, 2024. ISBN 9783031599002. doi: 10.1007/978-3-031-59900-2_6.
82. Alan M. Ferrenberg, Jiahao Xu, and David P. Landau. Pushing the limits of monte carlo simulations for the three-dimensional ising model. *Phys. Rev. E*, 97:043301, Apr 2018. doi: 10.1103/PhysRevE.97.043301.
83. N. D. Mermin. Lattice gas with short-range pair interactions and a singular coexistence-curve diameter. *Phys. Rev. Lett.*, 26:957–959, Apr 1971. doi: 10.1103/PhysRevLett.26.957.
84. John A. Zollweg and George W. Mulholland. On the law of the rectilinear diameter. *The Journal of Chemical Physics*, 57(3):1021–1025, 08 1972. ISSN 0021-9606. doi: 10.1063/1.1678352.
85. G. S. Pawley, R. H. Swendsen, D. J. Wallace, and K. G. Wilson. Monte carlo renormalization-group calculations of critical behavior in the simple-cubic ising model. *Phys. Rev. B*, 29: 4030–4040, Apr 1984. doi: 10.1103/PhysRevB.29.4030.
86. Antonia Statt, Helena Casademunt, Clifford P. Brangwynne, and Athanassios Z. Panagiotopoulos. Model for disordered proteins with strongly sequence-dependent liquid phase behavior. *The Journal of Chemical Physics*, 152(7):075101, 02 2020. ISSN 0021-9606. doi: 10.1063/1.5141095.
87. Xinxiang Chen, Jude Ann Vishnu, Pol Besenius, Julian König, and Friederike Schmid. Sol–gel transition in heteroassociative rna-protein solutions: A quantitative comparison of coarse-grained simulations and the semenov–rubinstein theory. *Macromolecules*, 58(6): 3331–3342, March 2025. ISSN 1520-5835. doi: 10.1021/acs.macromol.4c03065.
88. Naeem Jan and Dietrich Stauffer. Random site percolation in three dimensions. *International Journal of Modern Physics C*, 09(02):341–347, 1998. doi: 10.1142/S0129183198000261.
89. M. Rubinstein and R.H. Colby. *Polymer Physics*. OUP Oxford, 2003. ISBN 9780198520597.
90. M. Muthukumar. Thermodynamics of polymer solutions. *The Journal of Chemical Physics*, 85(8):4722–4728, 10 1986. ISSN 0021-9606. doi: 10.1063/1.451748.
91. Douglas T Li, Paul E Rudnicki, and Jian Qin. Distribution cutoff for clusters near the gel point. *ACS Polymers Au*, 2(5):361–370, 2022.
92. Min Kyung Shinn, Dylan T. Tomares, Vicky Liu, Avnika Pant, Yuanxin Qiu, Andreas Vitalis, You Jin Song, Yuna Ayala, Gregory W. Strout, Kiersten M. Ruff, Matthew D. Lew, Kannanganattu V. Prasanth, and Rohit V. Pappu. Nuclear speckle proteins form intrinsic and malat1-dependent microphases. *Cell*, page In press, 2025. doi: 10.1101/2025.02.26.640430.
93. Tingting Wu, Matthew R. King, Yuanxin Qiu, Mina Farag, Rohit V. Pappu, and Matthew D. Lew. Single-fluorogen imaging reveals distinct environmental and structural features of biomolecular condensates. *Nature Physics*, 21(5):778–786, March 2025. ISSN 1745-2481. doi: 10.1038/s41567-025-02827-7.
94. Max Kolb and Monique A. V. Axelos. *Gelation Transition versus Percolation Theory*, pages 255–261. Springer Netherlands, Dordrecht, 1990. ISBN 978-94-009-2157-3. doi: 10.1007/978-94-009-2157-3_21.
95. Xiangze Zeng and Rohit V. Pappu. Backbone-mediated weakening of pairwise interactions enables percolation in peptide-based mimics of protein condensates. *Communications Chemistry*, 8(1), April 2025. ISSN 2399-3669. doi: 10.1038/s42004-025-01502-5.
96. Wenli Meng, Nicholas Lyle, Bowu Luan, Daniel P. Raleigh, and Rohit V. Pappu. Experiments and simulations show how long-range contacts can form in expanded unfolded proteins with negligible secondary structure. *Proceedings of the National Academy of Sciences*, 110 (6):2123–2128, 2013. doi: 10.1073/pnas.1216979110.
97. P.G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, 1979. ISBN 9780801412035.
98. Lech Gmachowski. Fractal aggregate model of chain with mixed statistics. *Polymer*, 48(15):

4316–4321, 2007. ISSN 0032-3861. doi: https://doi.org/10.1016/j.polymer.2007.05.042.
99. Ibraheem Alshareedah, Wade M Borcherds, Samuel R Cohen, Anurag Singh, Ammon E Posey, Mina Farag, Anne Bremer, Gregory W Strout, Dylan T Tomares, Rohit V Pappu, Tanja Mittag, and Priya R Banerjee. Sequence-specific interactions determine viscoelasticity and ageing dynamics of protein condensates. *Nat. Phys.*, 20(9):1482–1491, September 2024.
100. Samuel R Cohen, Priya R Banerjee, and Rohit V Pappu. Direct computations of viscoelastic moduli of biomolecular condensates. *J. Chem. Phys.*, 161(9):095103, September 2024.
101. David Shirvanyants, Sergey Panyukov, Qi Liao, and Michael Rubinstein. Long-range correlations in a polymer chain due to its connectivity. *Macromolecules*, 41(4):1475–1485, January 2008. ISSN 1520-5835. doi: 10.1021/ma071443r.
102. A. L. Kholodenko and Karl F. Freed. Theta point ("tricritical") region behavior for a polymer chain: Transition to collapse. *The Journal of Chemical Physics*, 80(2):900–924, 01 1984. ISSN 0021-9606. doi: 10.1063/1.446748.
103. Isaac C. Sanchez. Phase transition behavior of the isolated polymer chain. *Macromolecules*, 12(5):980–988, September 1979. ISSN 1520-5835. doi: 10.1021/ma60071a040.
104. B. Widom. Theory of phase equilibrium. *The Journal of Physical Chemistry*, 100(31): 13190–13199, January 1996. ISSN 1541-5740. doi: 10.1021/jp9536460.
105. Peter Grassberger and Rainer Hegger. Simulations of three-dimensional θ polymers. *The Journal of Chemical Physics*, 102(17):6881–6899, 05 1995. ISSN 0021-9606. doi: 10.1063/1.469126.
106. Megan C. Cohan, Min Kyung Shinn, Jared M. Lalmansingh, and Rohit V. Pappu. Uncovering non-random binary patterns within sequences of intrinsically disordered proteins. *Journal of Molecular Biology*, 434(2):167373, 2022. ISSN 0022-2836. doi: https://doi.org/10.1016/j.jmb.2021.167373.
107. Timothy J. Nott, Evangelia Petsalaki, Patrick Farber, Dylan Jervis, Eden Fussner, Anne Plochowietz, Timothy D. Craggs, David P. Bazett-Jones, Tony Pawson, Julie D. Forman-Kay, and Andrew J. Baldwin. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Molecular Cell*, 57(5):936–947, March 2015. ISSN 1097-2765. doi: 10.1016/j.molcel.2015.01.013.
108. Cary R. Boyd-Shiwarski, Daniel J. Shiwarski, Shawn E. Griffiths, Rebecca T. Beacham, Logan Norrell, Daryl E. Morrison, Jun Wang, Jacob Mann, William Tennant, Eric N. Anderson, Jonathan Franks, Michael Calderon, Kelly A. Connolly, Muhammad Umar Cheema, Claire J. Weaver, Lubika J. Nkashama, Claire C. Weckerly, Katherine E. Querry, Udai Bhan Pandey, Christopher J. Donnelly, Dandan Sun, Aylin R. Rodan, and Arohan R. Subramanya. Wnk kinases sense molecular crowding and rescue cell volume via phase separation. *Cell*, 185 (24):4488–4506.e20, November 2022. ISSN 0092-8674. doi: 10.1016/j.cell.2022.09.042.
109. Hajime Tanaka. Viscoelastic phase separation. *Journal of Physics: Condensed Matter*, 12 (15):R207, 2000.
110. Gaurav Chauhan, Edward G. Wilkinson, Yaning Yuan, Samuel R. Cohen, Masayuki Onishi, Rohit V. Pappu, and Lucia C. Strader. Active transport enables protein condensation in cells. *Science Advances*, 11(21):eadv7875, 2025. doi: 10.1126/sciadv.adv7875.
111. Hagen Hofmann, Andrea Soranno, Alessandro Borgia, Klaus Gast, Daniel Nettels, and Benjamin Schuler. Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences*, 109(40):16155–16160, 2012.
112. Kingshuk Ghosh, Jonathan Huihui, Michael Phillips, and Austin Haider. Rules of physical mathematics govern intrinsically disordered proteins. *Annual Review of Biophysics*, 51(1): 355–376, 2022.
113. Ruoyao Zhang, Gaurav Mitra, Souradeep Ghosh, and Rohit V. Pappu. Computational rheometry of viscoelastic networks: From random graphs to biomolecular condensates. *bioRxiv*, 2025. doi: 10.1101/2025.08.20.671282.
114. Andrés R. Tejedor, Anne Aguirre Gonzalez, M. Julia Maristany, Pin Yu Chew, Kieran Russell, Jorge Ramirez, Jorge R. Espinosa, and Rosana Collepardo-Guevara. Chemically informed coarse-graining of electrostatic forces in charge-rich biomolecular condensates. *ACS Central Science*, 11(2):302–321, February 2025. ISSN 2374-7951. doi: 10.1021/acscentsci.4c01617.
115. Fan Cao, Sören von Bülow, Giulio Tesei, and Kresten Lindorff-Larsen. A coarse-grained model for disordered and multi-domain proteins. *Protein Science*, 33(11), October 2024. ISSN 1469-896X. doi: 10.1002/pro.5172.
116. Sören von Bülow, Giulio Tesei, Fatima Kamal Zaidi, Tanja Mittag, and Kresten Lindorff-Larsen. Prediction of phase-separation propensities of disordered proteins from sequence. *Proceedings of the National Academy of Sciences*, 122(13), March 2025. ISSN 1091-6490. doi: 10.1073/pnas.2417920122.
117. D Janka Bauer, Lukas S Stelzl, and Arash Nikoubashman. Single-chain and condensed-state behavior of hnrnpa1 from molecular simulations. *The Journal of Chemical Physics*, 157(15), 2022.
118. Yaxin An, Michael A Webb, and William M Jacobs. Active learning of the thermodynamics-dynamics trade-off in protein condensates. *Science Advances*, 10(1):eadj2448, 2024.
119. Antonia Statt, Helena Casademunt, Clifford P Brangwynne, and Athanassios Z Panagiotopoulos. Model for disordered proteins with strongly sequence-dependent liquid phase behavior. *The Journal of Chemical Physics*, 152(7), 2020.
120. R. P. Brent. An algorithm with guaranteed convergence for finding a zero of a function. *The Computer Journal*, 14(4):422–425, 01 1971. ISSN 0010-4620. doi: 10.1093/comjnl/14.4.422.

1. Conversion from simulation to experimental units

Scaling factors are required to convert the computed coexistence curves to the simulation temperatures into degree Kelvin and volume fractions into molar units. For the latter, we followed Wei et al. (38) and Farag et al. (52). A volume fraction of 1.0 corresponds to a mass concentration of 1310 mg/ml. Following Farag et al. (52), we used a multiplicative factor of 5.6 to convert the simulation temperature in LaSSI (which is written in units of $k_B = 1$) to degrees Kelvin. This factor was derived as follows: Binodals extracted from LaSSI simulations were converted into mass concentrations along the abscissa, and the ordinate was multiplied by a conversion factor that maximized the overlap with the measured binodals for A1-LCD (17, 52). This conversion, which was derived for A1-LCD, was then applied without further modification to more than 30 different sequences derived from the A1-LCD parent (52, 67) and to the LCD of FUS (59). A fixed density scaling factor (0.6) was used to convert volume fraction from lattice units to volume fraction as measured experimentally or, equivalently, from simulation concentrations in (M) to experimental concentrations measured in (M) (17, 52). Lengths in lattice units were converted to real units using a lattice spacing of 4 Å per lattice unit. This spacing was chosen because it approximately corresponds to the average distance between adjacent C_α atoms ($3.8 - 4.0$ Å) (111, 112). Consequently, one lattice site represents a single residue both in volume and in bond length, yielding physically realistic chain dimensions. Together, these conversions enable direct comparison between lattice simulations and experimental data.

2. Details of simulation setup and considerations for mapping the binodal of WT A1-LCD

System size considerations for the LaSSI simulations. Previously, a 200-chain system was used to map the binodal for WT A1-LCD (52). For this system, we were unable to obtain clearly demarcated dense and dilute coexistence concentrations for simulation temperatures greater than $T = 54.0$ (Fig. S4). This shows that, for a system with 200 molecules, we cannot map binodals beyond $T = T_{\text{crossover},\Delta}$ unless we use simulation boxes that are considerably smaller than the contour length of WT A1-LCD (137 l.u.), to a side. To discern the optimal box size and numbers of A1-LCD molecules (n) that will be needed to map critical regions, we first estimated the critical volume fraction (ϕ_c) by fitting the computed binodal from the simulations with 200 molecules. For this, we used the modified Flory-Huggins model (68).

The estimated value of the critical volume fraction was found to be $\phi_c \approx 0.1$. The contour length of each A1-LCD molecule is 137 l.u. Therefore, to achieve the estimated volume fraction using 200 molecules when the system is quenched to be just below the apparent critical point, the dimensions of the cubic box would have to be $L_{\text{sim}} \approx 65$ l.u. to a side. However, this is less than half the chain contour length. This becomes a problem because the simulation volume of a cubic box can only accommodate density fluctuations that are smaller than the contour length of a single chain. Additionally, as $T \to T_c$, the correlation length ξ_c becomes comparable to the box size L_{sim}. Therefore, a box comprising 200 molecules with $L_{\text{sim}} \approx 65$ l.u. poses two challenges: First, the estimates of density fluctuations will be noisy due to severe finite-size effects resulting from the use of small numbers of molecules to mimic a macroscopic system. Second, the long-wavelength density fluctuations will be suppressed by the small box size that is necessitated by the use of only 200 molecules. In contrast, for 10^4 molecules, the box length will need to increase to $L_{\text{sim}} \approx 239$ l.u. to achieve a volume fraction of $\phi \approx 0.1$. This box size accommodates many independent correlation volumes thereby enabling the sampling of a broader spectrum of density and conformational fluctuations. The dilute and dense arms that are extracted from radial profiles converge smoothly, generating the expected rounded binodal near T_c.

To estimate the appropriate numbers of molecules and box sizes to use, we use the following considerations. For fixed $\phi \approx \phi_c$, the adequate box length will scale as $L_{\text{sim}} \propto n^{1/3}$, where n is the number of chain molecules. Since $n \propto L_{\text{sim}}^3$, the minimum number of molecules that would be required to map phase behaviors close to criticality would require that the simulation box size $L_{\text{sim, target}}$ at $\phi \approx \phi_c$ be between $1.5 \times N$ and $2 \times N$, where $N = 137$ is the contour length for A1-LCD. This gives

$$n_{\text{min}} \;=\; 200 \left(\frac{L_{\text{sim,target}}}{L_{\text{sim,200}}} \right)^3,$$

where $L_{\text{sim,200}} = 65$ l.u. and $L_{\text{sim,target}} = 1.5 \times 137$ l.u. or $L_{\text{sim,target}} = 2 \times 137$ l.u.. Substituting these values, we obtain

$$n_{\text{min, 1.5N}} = 200 \left(\frac{205.5}{65} \right)^3 \approx 6.3 \times 10^3,$$

$$n_{\text{min, 2N}} = 200 \left(\frac{274}{65} \right)^3 \approx 1.5 \times 10^4.$$

Hence, the minimum number of A1-LCD molecules required to ensure that $L_{\text{sim,target}} \gtrsim (1.5 - 2) \times N$ lies in the range

$$6.3 \times 10^3 \;\lesssim\; n_{\text{min}} \;\lesssim\; 1.5 \times 10^4.$$

The simulations with $n \approx 10^4$ lies within the prescribed range, and this should be adequate to map the phase behavior in the vicinity of the critical point. In our simulations, we used a cubic box of size $L_{\text{sim}} = 240$ l.u. for a system of 10^4 A1-LCD molecules.

As an aside, we note that the findings reported here should also be accessible via off-lattice, molecular dynamics (MD) simulations. In fact, a recent study (113) showed that structures within dense phases of A1-LCD molecules share similarities across three different engines that include LaSSI, MPIPI-recharged (114), and CALVADOS (69, 115). Based on these comparisons, we propose that the findings reported are likely to be consistent across different simulation protocols that use similar, single-bead per residue models. The key to mapping critical points is the avoidance of finite size artifacts and the use of finite size scaling methods (60, 80). It might also be essential to revisit the use of slab geometries that are in vogue across many MD simulations (20, 69, 72–76, 116). Slabs cause a smoothening of density profiles and this is achieved by suppressing fluctuations that are the defining hallmarks of critical regions. Assuming finite size artifacts are avoided and cubic boxes that do not break symmetries *a priori* are used, then the extraction of cluster-size distributions across different temperature regimes should be feasible using MD simulations that can be deployed using different engines.

Derivative of T-dependent $\phi(r)$ profiles to extract $\Delta(T)$ and $r_{\mathrm{mid}}(T)$. Previously, Farag et al., (52) used the Fisk-Widom equation and estimated four parameters, ϕ_{dilute}, ϕ_{dense}, Δ, and r_{mid}, simultaneously from fits to the radial density profiles $\phi_T(r)$. However, this approach yielded confounding observations regarding the chain-length (N) dependence for the scaling of Δ with T for homopolymers. Specifically, Farag et al. observed an N-dependence for $\Delta(T)$ for $T < T_{\text{crossover},\Delta}$, and a convergence above this temperature. In hindsight, the correct behavior should be the opposite of what was observed and reported as being confounding (52). Closer scrutiny revealed that the erroneous observation stems from the extraction of four parameters from a single fit. Such a fit is weighted toward the high density of points in the plateau regions, which define the densities of the coexisting phases. To eliminate this error, we followed Chauhan et al., and implemented a two-step approach (110). We first used the derivatives of $\phi_T(r)$ to estimate $\Delta(T)$ and $r_{\text{mid}}(T)$. Next, we use the derived values of $\Delta(T)$ and $r_{\text{mid}}(T)$ in the Fisk-Widom function to estimate ϕ_{dilute} and ϕ_{dense}.

3. Comparative analyses of different approaches used to map critical points

Following Silmore et al., (71), it has become common practice to assume that the order parameter defined by $\Delta\phi = (\phi_{\mathrm{dense}} - \phi_{\mathrm{dilute}})$ scales as $\left(\frac{T_c - T}{T_c}\right)^{\beta}$, with β set to be ≈ 0.33, being the exponent that is expected for the 3D Ising model. The procedure that is followed by most authors is to use the law of rectilinear diameters and the assumption of the scaling of the order parameter to be congruent with the 3D Ising model, thereby fixing β, and using a regression analysis to estimate T_c, ϕ_c, and a pre-factor A_0 (20, 72, 73, 76, 77, 117, 118). Note that the scaling of the order parameter is only valid in the vicinity of the critical point (104). Accordingly, some authors exercise caution in the use of the regression analysis (119), although this will be affected by the paucity of points that one sample in the vicinity of the critical point from simulations of finite-sized systems. To work around this paucity of points, most authors enforce the scaling of the order parameter across the entirety of the computed binodal.
An alternative approach has been to fit computed or measure binodals to a modified Flory-Huggins model (17, 67, 68). This method is reasonable in the off-critical, mean field regime. However, it generates a sharper scaling of the order parameter in the vicinity of the critical point. Note that prior usage of the modified Flory-Huggins model has been restricted to extracting the two- and three-body interaction coefficients and to qualitative rather than quantitative estimates of T_c and ϕ_c. Note also that contrary to certain assertions (73, 76), the experiments (67) do not report direct measurements of T_c and ϕ_c. Instead, they report estimates obtained by fitting the measurements to the modified Flory-Huggins model (17, 67).
Here, we discuss the approaches used in the literature and assess their performance in terms of estimating T_c and ϕ_c, and the quality of numerical fits to computed binodals for the 200- and 10^4-chain systems. We first discuss the approach that is common practice in the computational literature. In this approach, the reduced temperature is defined as

$$T_r = 1 - \frac{T}{T_c}, \tag{1}$$

where T_c is the critical temperature. According to the law of rectilinear diameters, the average of the coexisting phase volume fractions varies linearly with temperature:

$$\frac{\phi_{\mathrm{dense}}(T) + \phi_{\mathrm{dilute}}(T)}{2} = \phi_c - m T_c T_r, \tag{2}$$

where ϕ_c is the critical volume fraction and m is the slope. Next, the difference between the dense and dilute arms is taken to scale as a power law with Ising critical exponent $\beta \approx 0.33$:

$$\Delta\phi(T) = \phi_{\mathrm{dense}}(T) - \phi_{\mathrm{dilute}}(T) = A_o T_r^{\beta}, \tag{3}$$

where A_o is a non-universal amplitude. Combining these two expressions, the binodal branches are given by

$$\phi_{\mathrm{dense}}(T) = \phi_c - m T_c T_r + \tfrac{1}{2} A_o T_r^{\beta}, \tag{4}$$

$$\phi_{\mathrm{dilute}}(T) = \phi_c - m T_c T_r - \tfrac{1}{2} A_o T_r^{\beta}. \tag{5}$$

The system of equations combined with setting $\beta \approx 0.33$, and knowledge of the computed values of $\phi_{\text{dense}}(T)$ and $\phi_{\text{dilute}}(T)$ across the entire binodal are used to estimate m, T_c, ϕ_c, and A_0. We assessed the quality of the fits obtained using the approach laid out above. The analysis was performed using data from simulations of 200 chains, which lacks information regarding the binodal for temperatures above $T_{\text{crossover},\Delta}$, and for data for simulations of the 10^4-chain system. We also used the FIREBALL package (68) to obtain fits to the modified Flory-Huggins model. The results are shown in Fig. S6.

Given that the concentrations along the dilute arm of the binodal span several orders of magnitude, we show the results of the fits using a logarithmic scale for the concentration (Fig. S6A) and linear scale for the concentration (Fig. S6B). For the 200-chain system, we observe the following results: The dilute arm is fit very poorly by imposing the scaling of the order parameter from the 3D Ising model to apply across the entire binodal (Fig. S6A). This is congruent with the fact that this scaling should only apply in the vicinity of the critical point. However, as noted in the main text, the near critical regime cannot be sampled in simulations based on 200 molecules, because the system is too small for adequate delineation between coexisting phases in the critical regime. Given the lack of data for temperatures above $T_{\text{crossover},\Delta}$, the estimated T_c is very low (Fig. S6A, S6B) when compared to what we obtain using the Binder cumulant analysis for the system comprising 10^4 chains. In contrast, the binodal obtained by fitting the data to the modified Flory-Huggins model does a better job in terms of reproducing the binodal. However, the estimated T_c is still lower than what it should be, although the estimate is closer to the true value when compared to the approach based on the 3D Ising model.

Next, we assessed the performance of the two approaches using data from 10^4-chain system. Fitting the entire binodal to the scaling of the order parameter that is based on the 3D Ising model improves the estimation of T_c. However, the reproduction of the binodal is very poor and this is expected because the scaling is invalid away from the critical regime (Fig. S6C, S6D). When we restrict the analysis to the near critical regime, i.e., to temperatures above $T_{\text{crossover},\Delta}$, the estimated T_c does not change much, but even here the symmetry expected of the order parameter is not present in the actual data, and hence the fit to the binodal in the vicinity of the critical point remains poor (Fig. S6D). As noted in the main text, equations requiring the simultaneous estimation of multiple parameters (four in this case) will fare poorly when there is a sparse number of points. As for the modified Flory-Huggins model, the fits, in general, look better than what we obtain based on the 3D Ising model. However, the value of β will always be 0.5, and this, as we have shown in the main text, is incorrect.

In summary, the preceding analysis establishes why direct estimation of the critical parameters and / or characterizations of near-critical behaviors (44) are likely to be problematic if one uses approaches that are in vogue in the computational literature. Of course, as crude estimates of T_c and for assessing how this parameter varies with changes to sequence, the approach of regression is likely to be useful. Even for this exercise, we would recommend the use of fits to the modified Flory-Huggins model, as we have deployed, as opposed to the use of regression based on assuming scaling of the order parameter based on the 3D Ising model (76, 118). At least with the modified Flory-Huggins model, one gets a more accurate estimation of the transfer free energy, and this provides greater confidence with regard to comparative assessments of sequence-specific effects on phase behaviors.

4. Binder cumulant analysis using sub-box sampling

The order parameter of interest at a given temperature T is the deviation of the density inside a cubic sub-box of length L from its mean value:

$$m_L = \phi_L(T) - \langle \phi_L(T) \rangle, \tag{6}$$

We define the second and fourth central moments of the sub-box order parameter as $\mu_2(L;T) = \langle m_L^2(T) \rangle$ and $\mu_4(L;T) = \langle m_L^4(T) \rangle$, respectively. The dimensionless Binder cumulant is then

$$U_L(T) = 1 - \frac{\mu_4(L;T)}{3\mu_2(L;T)^2}, \tag{7}$$

or equivalently, $U_L = 1 - \kappa/3$, with κ denoting the kurtosis of $p(m_L)$. In the one-phase regime above T_c, $p(m_L)$ should be Gaussian and unimodal, corresponding to $U_L \to 0$. In the two-phase regime, $p(m_L)$ becomes bimodal, and $U_L \to 2/3$. To choose the optimal range of sub-box sizes, we used: $R_{ee} < L \lesssim 0.8\, L_{\text{sim}}$. Here L_{sim} is the length of the full simulation box ($L_{\text{sim}} = 240$ l.u. for a system of 10^4 A1-LCD molecules). This choice of upper and lower bounds for L ensures that each sub-box is larger than the characteristic polymer size (R_{ee}) to avoid intra-chain correlations, while remaining small enough relative to L_{sim} to provide sufficient statistical sampling of density fluctuations.

Sampling of sub-boxes and calculation of sub-box densities. 10^4 cubic sub-boxes of size L were placed uniformly at random within the simulation box. The coordinates of the origin of each sub-box (x_0, y_0, z_0) were drawn independently from a uniform distribution: $x_0, y_0, z_0 \sim \mathcal{U}(0, L_{\text{sim}} - L)$, ensuring that all sub-boxes remained entirely within the simulation boundaries. Random uniform draws produce overlapping sub-boxes, which provide an unbiased and statistically dense sampling. The simulation coordinates were discretized onto a cubic lattice of size L_{sim}^3, producing a binary occupancy grid $\phi(\mathbf{r})$ that equals 1 for occupied sites and 0 otherwise. The list of bead coordinates was ordered sequentially by chain. For each randomly placed

sub-box V_i of volume L^3, the local density was calculated as $\rho_i = \frac{1}{L^3}\sum_{\mathbf{r}\in V_i}\phi(\mathbf{r})$, corresponding to the fraction of occupied lattice sites within that region. The sub-box densities were used to construct the probability distribution $p_L(\phi)$ and to compute the second and fourth-order moments of the order parameter.

Pairwise Binder Cumulant crossings to determine T_c. Binder cumulant curves $U_L(T)$ from different sub-box sizes L intersect near the fixed-point value U^* (60). To estimate the critical temperature T_c, we employed the standard Binder cumulant crossing analysis across sub-boxes of size L. For each sub-box of size L, the discrete values of $U_L(T)$ obtained from simulations were interpolated using a monotonic piecewise cubic Hermite interpolant (PCHIP), yielding continuous $U_L(T)$. For every pair of sub-box sizes $(L_1, L_2) \geq 160$ l.u., we computed the difference $D_{L_1,L_2}(T) = U_{L_1}(T) - U_{L_2}(T)$, and located all temperatures T satisfying $D_{L_1,L_2}(T) = 0$. For this, we used Brent's root-finding method (120) within the overlap of their temperature ranges. The resulting set of pairwise crossing temperatures $\{T_c^{(L_1,L_2)}\}$ can be treated as a distribution of finite-size estimates for the critical temperature. The median of all filtered pairwise crossings thus provides a robust estimate of T_c with well-defined uncertainty bounds (the upper and lower values of this set).

5. Derivation of the percolation threshold from mean-field predictions $\phi_{\mathrm{perc,MF}}(T)$

To compute the percolation threshold from a mean field model, we used the generalized bond percolation model of Choi et al., (21). In this model, the mean field connectivity parameter between stickers of types 'i' and 'j' is defined as:

$$\lambda_{ij}(T) = v_{ij}\, e^{-\beta\varepsilon_{ij}}, \tag{8}$$

where v_{ij} is the effective bond volume (configurational volume for forming an i–j contact), ε_{ij} is the pairwise interaction energy in units of $k_{\mathrm{B}}T$, and $\beta = 1/k_BT$ is the inverse temperature. The total connectivity D arising from all sticker types is given by (21),

$$D(T) = \sum_i \lambda_{ii} n_i^2 + 2\sum_{i<j}\lambda_{ij} n_i n_j \tag{9}$$

where n_i is the number of stickers of type 'i' per chain. The percolation threshold is calculated using

$$c_{\mathrm{perc}}(T) = \frac{1}{D(T)} \tag{10}$$

For WT A1-LCD, the relevant sticker types are Y (tyrosine), F (phenylalanine), and R (arginine), where R interacts primarily with the aromatic residues. Expanding $D(T)$ gives

$$\begin{aligned} D(T) &= \lambda_{YY}(T)n_Y^2 + \lambda_{FF}(T)n_F^2 + 2\lambda_{YF}(T)n_Y n_F + 2\lambda_{RAro}(T)\, n_R n_{Aro} \\ &= \lambda_{YY}(T)n_Y^2 + \lambda_{FF}(T)n_F^2 + 2\lambda_{YF}(T)n_Y n_F + 2\lambda_{RAro}(T)\, n_R(n_Y + n_F) \end{aligned}$$

Hence,

$$c_{\mathrm{perc}}(T) = \frac{1}{\lambda_{YY}(T)n_Y^2 + \lambda_{FF}(T)n_F^2 + 2\lambda_{YF}(T)n_Y n_F + 2\lambda_{RAro}(T)n_R(n_Y + n_F)} \tag{11}$$

If each chain consists of N beads and each bead occupies a volume v_b ($v_b = 1$ l.u.^{3} in LaSSI units), then the percolation volume fraction is

$$\phi_{\mathrm{perc}}(T) = c_{\mathrm{perc}}(T)\, N v_b \tag{12}$$

Substituting $c_{\mathrm{perc}}(T)$ gives the final expression:

$$\phi_{\mathrm{perc}}(T) = \frac{N v_b}{\lambda_{YY}(T)n_Y^2 + \lambda_{FF}(T)n_F^2 + 2\lambda_{YF}(T)n_Y n_F + 2\lambda_{RAro}(T)n_R(n_Y + n_F)} \tag{13}$$

Now,

$$\lambda_{ij}(T) = v_{ij}\, e^{-\varepsilon_{ij}/k_BT} \tag{14}$$

Since $k_B = 1$ in LaSSI, the percolation threshold explicitly depends on temperature as

$$\phi_{\mathrm{perc}}(T) = \frac{N v_b}{v_{YY}e^{-\varepsilon_{YY}/T}n_Y^2 + v_{FF}e^{-\varepsilon_{FF}/T}n_F^2 + 2v_{YF}e^{-\varepsilon_{YF}/T}n_Y n_F + 2v_{RAro}e^{-\varepsilon_{RAro}/T}n_R(n_Y + n_F)} \tag{15}$$

If all bond volumes are taken as being equal ($v_{ij} = 1$ l.u.^{3}) and ε_{ij} are the LaSSI contact energies, the final form of the expression is

$$\phi_{\mathrm{perc}}(T) = \frac{N}{e^{-\varepsilon_{YY}/T}n_Y^2 + e^{-\varepsilon_{FF}/T}n_F^2 + 2e^{-\varepsilon_{YF}/T}n_Y n_F + 2e^{-\varepsilon_{RAro}/T}n_R(n_Y + n_F)} \tag{16}$$

For the WT A1-LCD sequence ($\mathrm{N} = 137$ residues), the relevant sticker types and their counts per chain are:

$$n_Y = 7 \quad \text{(Tyrosine)},$$
$$n_F = 12 \quad \text{(Phenylalanine)},$$
$$n_R = 10 \quad \text{(Arginine)}$$

6. Assigning chains to dense versus dilute phases

Chains were classified as belonging to dense versus dilute phases based on their spatial proximity to the center of mass (COM) of the largest connected cluster. The COM of each chain was computed first. The Euclidean distance between the COM of each chain and that of the largest cluster was then calculated; chains within a predefined dense cutoff radius ($= 60\,\mathrm{l.u.}$) were assigned to the dense phase, those beyond a dilute cutoff radius ($= 120\,\mathrm{l.u.}$) were assigned to the dilute phase, and those lying between the two cutoffs were ignored in the current analysis. The cutoff values were chosen such that, across all simulation temperatures examined, the dense and dilute phase regions correspond to the plateau regions of the respective radial density profiles (Fig. 1B), thereby ensuring unambiguous classification.

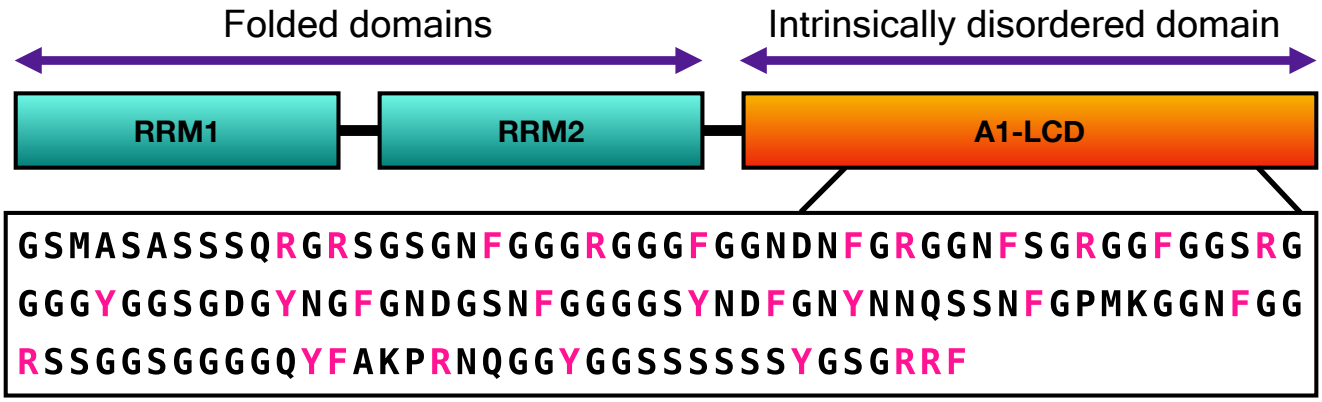

Fig. S1. Architecture of the hnRNP-A1 protein and amino acid sequence of A1-LCD. hnRNP-A1 comprises the folded RNA Recognition Motifs and the PLCD (Prion-Like Low Complexity Domain), which is designated as A1-LCD. The amino acid sequence of A1-LCD, which has 137 residues, is also shown in the box. The sticker residues in the A1-LCD sequence– tyrosine (Y), phenylalanine (F), and arginine (R) are colored in magenta to delineate them from the rest of the residues (colored in black).

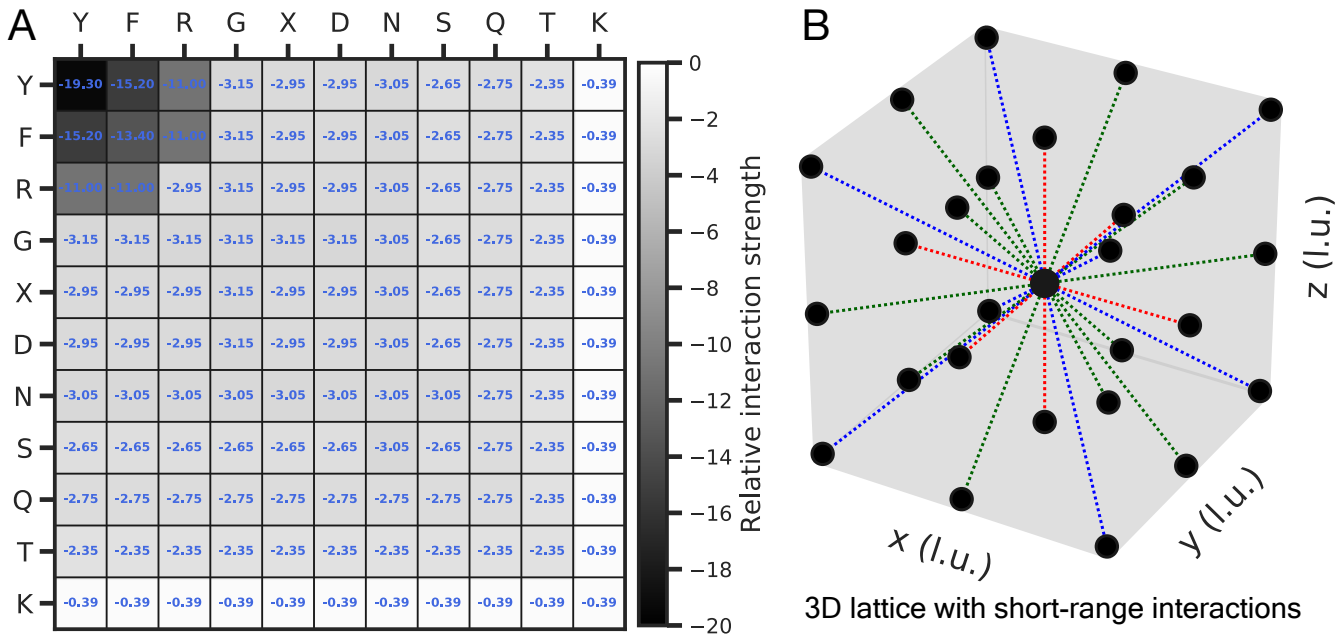

Fig. S2. Interaction energies and definition of a contact in the bond fluctuation model framework. (A) Matrix showing the pairwise interaction strengths used in the LaSSI computational model. The interaction energies are those of Farag et al. (52). The energies were parameterized to reproduce the sequence-specificity of measured radii of gyration (67) and the parameterization scheme was based on the Gaussian Process Bayesian Optimization approach of Ruff et al. (58). "Aro" is used to indicate either Y or F. "X" is used to indicate any amino acid for which a specific interaction is not defined. The values for the relative interaction strengths for each residue pair are shown inside the squares of the matrix grid. The energies are unitless. Moves are accepted or rejected by referencing to thermal energy k_BT, where $k_B = 1$ and T is the simulation temperature. The temperatures considered range from 44 to 60. (B) Schematic of a 3D lattice with short-range interactions where each bead (representing a single residue of the A1-LCD sequence) has 26 possible neighbors at distances of 1 l.u. (colored by red dotted lines), $\sqrt{2}$ l.u. (colored by green dotted lines), or $\sqrt{3}$ l.u. respectively (colored by blue dotted lines).

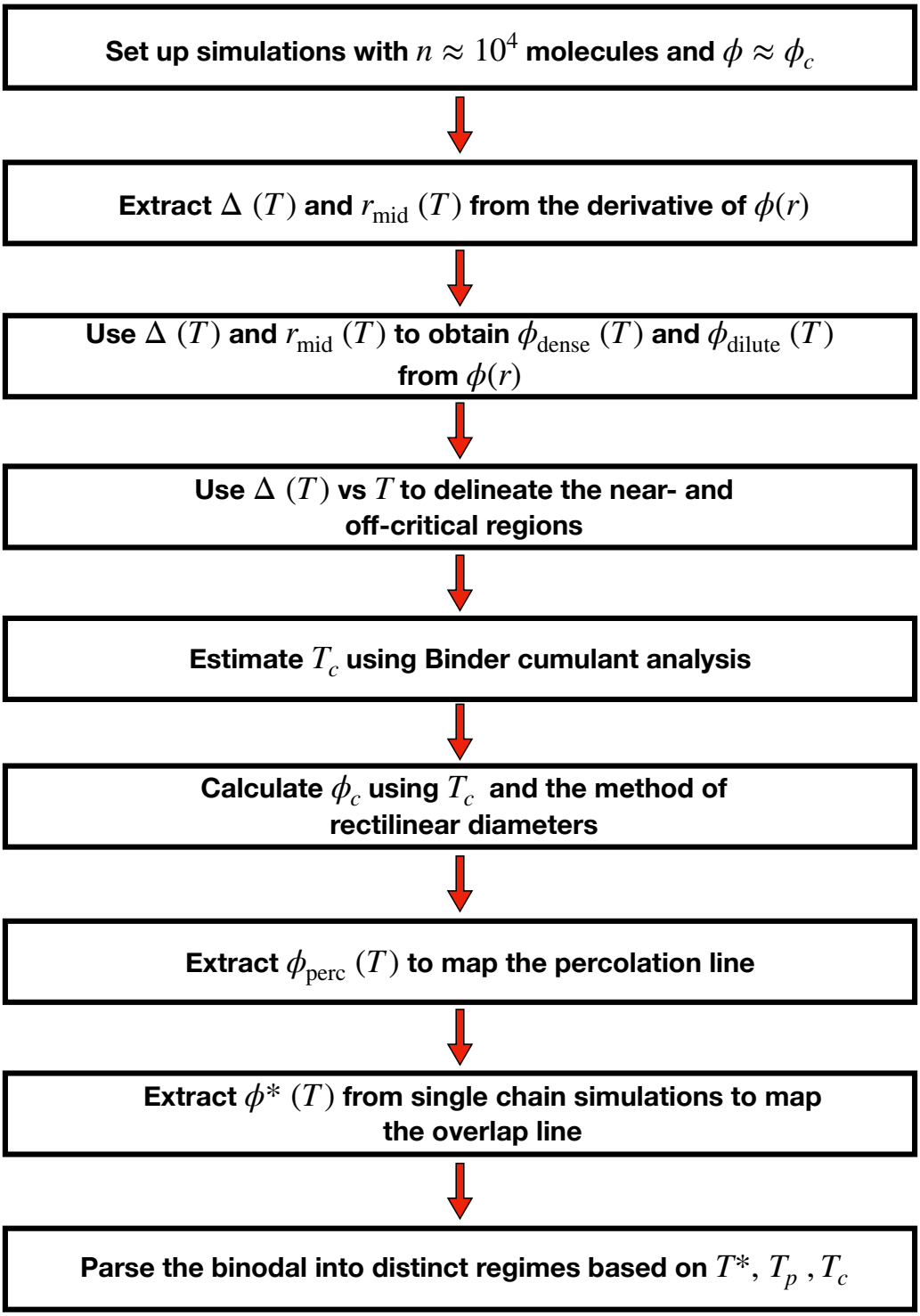

Fig. S3. Flowchart showing the nine-step simulation and analysis protocol to map the full binodal for WT A1-LCD. Details of the protocol are provided in the Materials and Methods section of the main text.

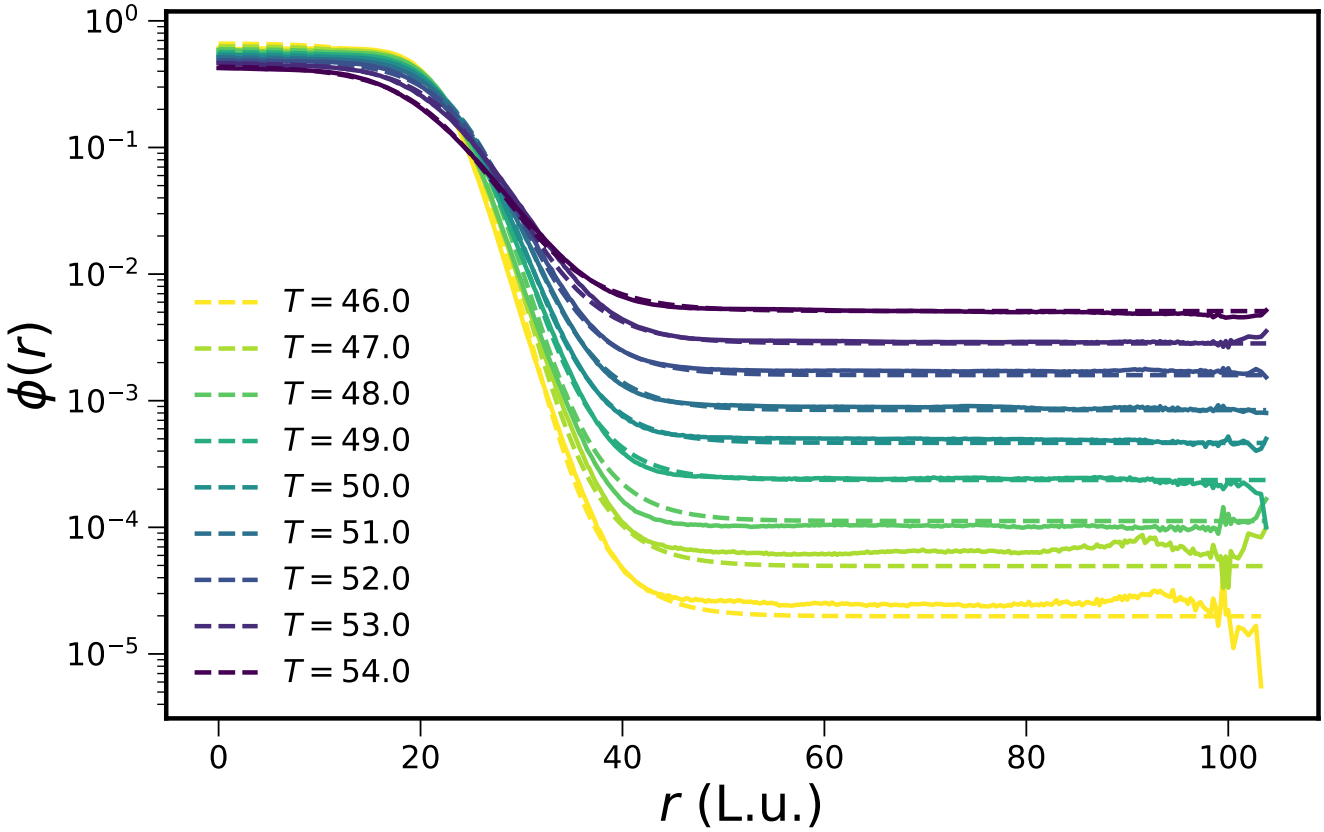

Fig. S4. T-dependent radial density profiles $\phi_T(r)$ for the 200-chain system of WT A1-LCD. Radial density profiles for the system containing 200 molecules of WT A1-LCD, as a function of the simulation temperature T, were used to extract the dense and dilute coexistence concentrations by fitting them to the Fisk-Widom function.

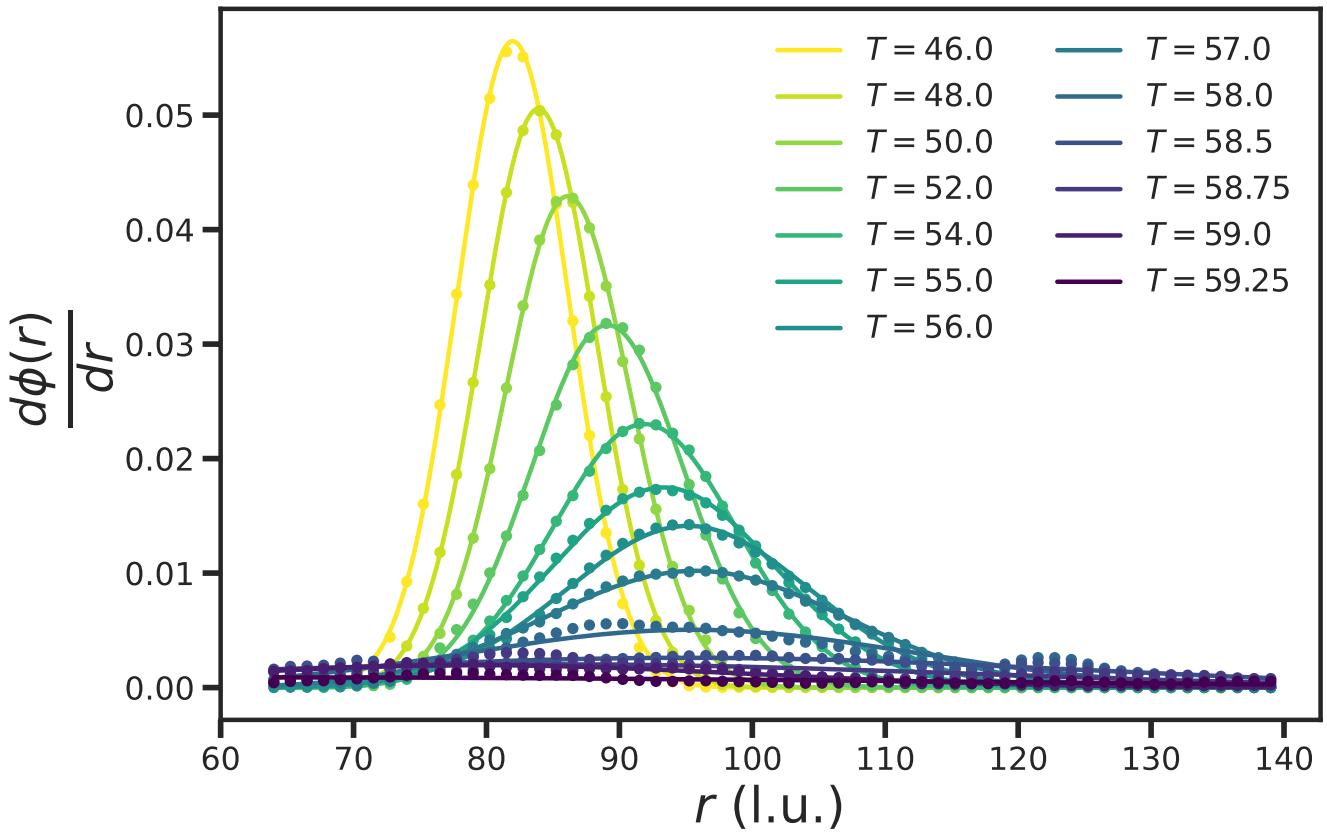

Fig. S5. Derivatives of the radial density profiles, $d\phi(r)/dr$, at different temperatures for the 10^4-chain system of WT A1-LCD. The derivative of the radial density profile at each given temperature is fit to a truncated Gaussian distribution, from which the interfacial width $\Delta(T)$ is extracted as the full-width-at-half-maximum (FWHM), given by $2\sigma\sqrt{2\ln 2}$, where σ is the standard deviation. The midpoint of the interface $r_{\mathrm{mid}}(T)$ is the mode of the Gaussian.

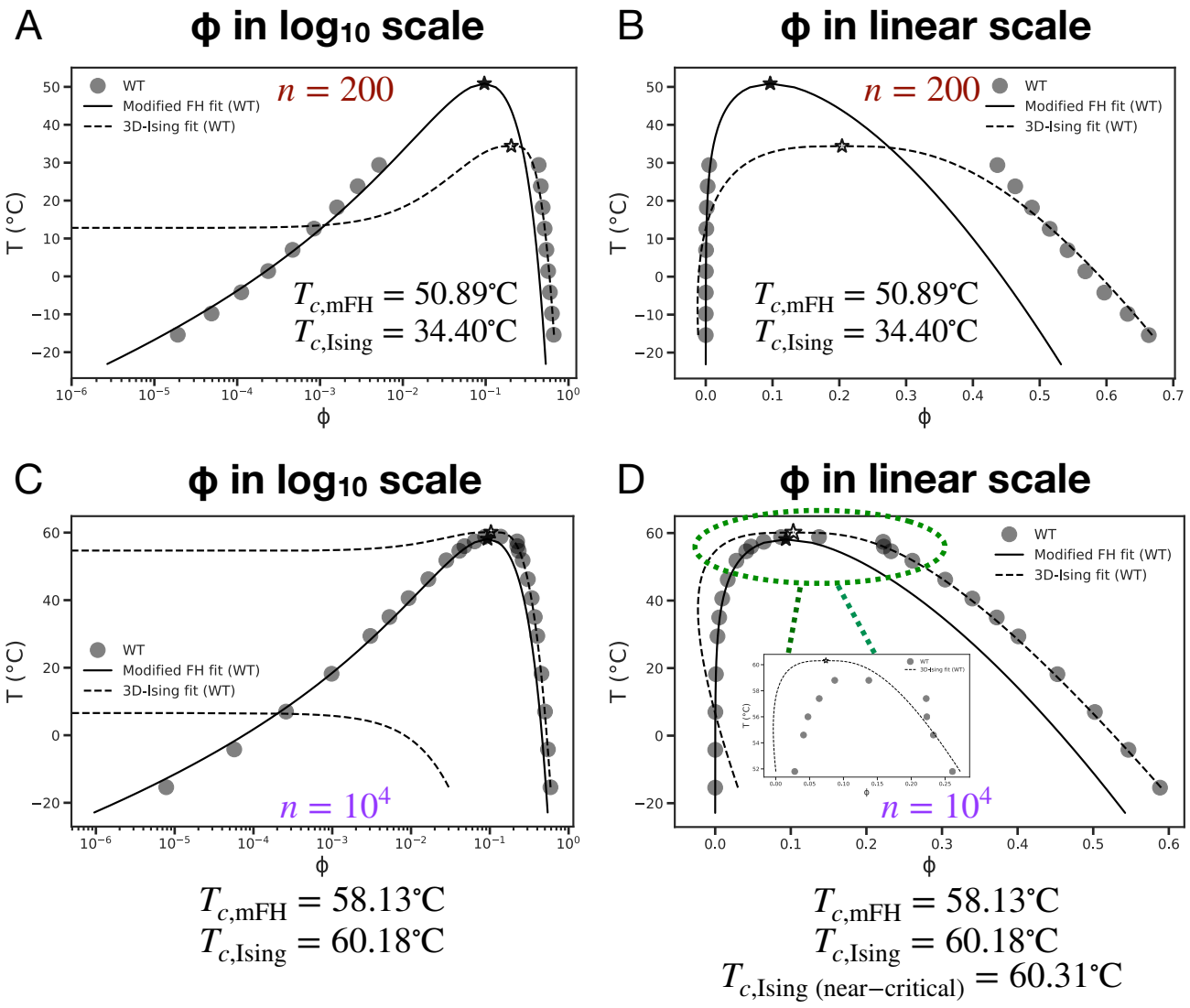

Fig. S6. Comparative analysis of estimates of critical points and fitting binodals based on two different approaches. Fits performed using the modified FH model and the 3D-Ising model to the dense and dilute arms of the binodal for a system containing 200 molecules of WT A1-LCD, in both (A) log scale and (B) linear scale of the volume fraction ϕ. For a system containing 10^4 molecules of WT A1-LCD, the corresponding fits in (C) log scale and (D) linear scale of the volume fraction ϕ are also shown. The inset in (D) shows the fit performed using the 3D-Ising model approach to the points in the near-critical region, i.e., $T > T_{\text{crossover},\Delta}$. Even though the estimate of T_c improves for the $n = 10^4$ system as opposed to the $n = 200$ system, the fit to the dilute arm using the 3D-Ising model approach is poor.

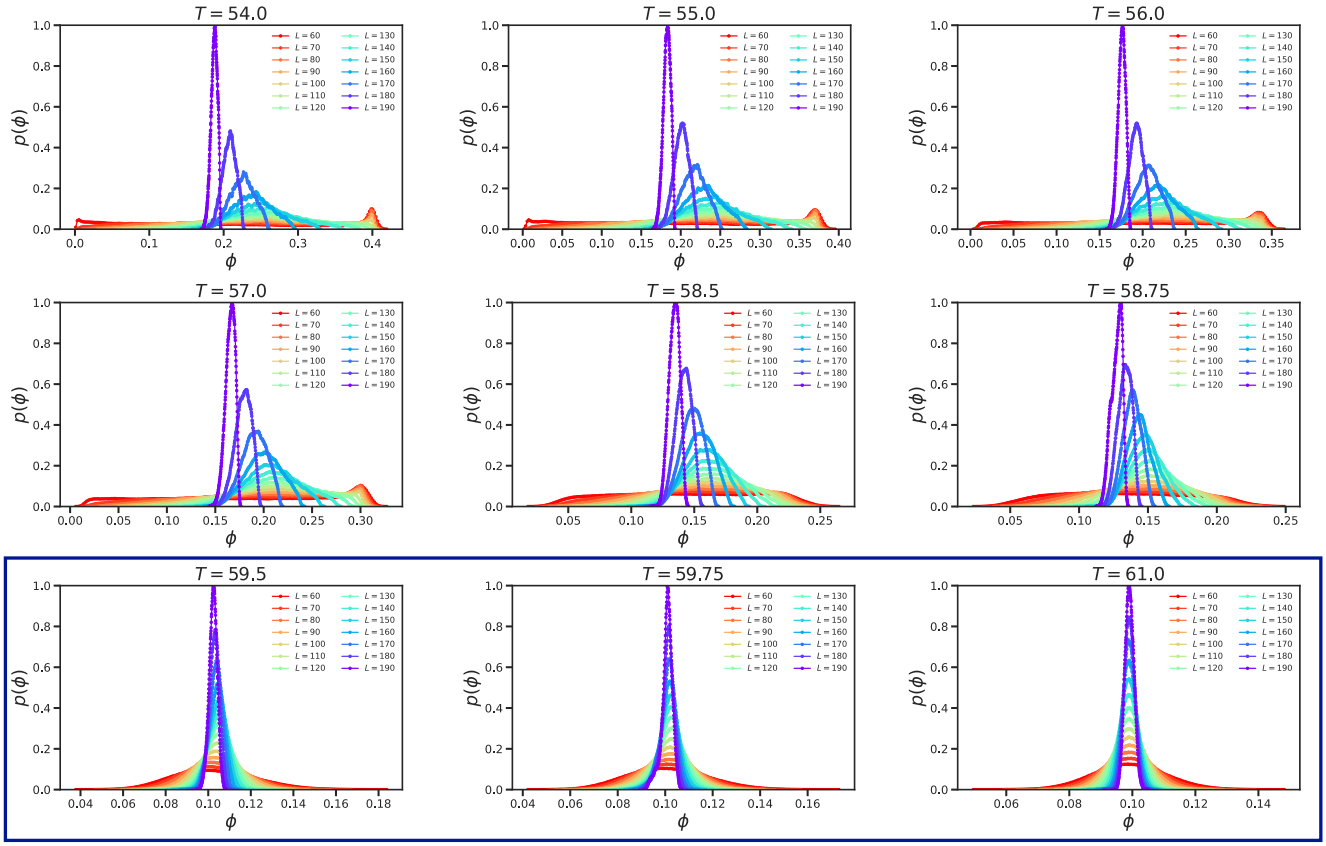

Fig. S7. Distributions $p(\phi)$ as a function of ϕ obtained by sampling sub-boxes of various sizes, ranging from $L = 60$ to $L = 190$, at different representative simulation temperatures. Note that as T approaches $T_c \approx 59.36$, the density distributions show Gaussian behavior for large L. In contrast, at lower values of T (for which the system is in the two-phase regime) and small sub-box sizes, we observe a bimodality in the density distributions. For each of the temperatures shown inside the navy blue box that corresponds to $T > T_c$, all the distributions collapse onto a unimodal Gaussian form across all sub-box sizes, consistent with homogeneous single-phase behavior.

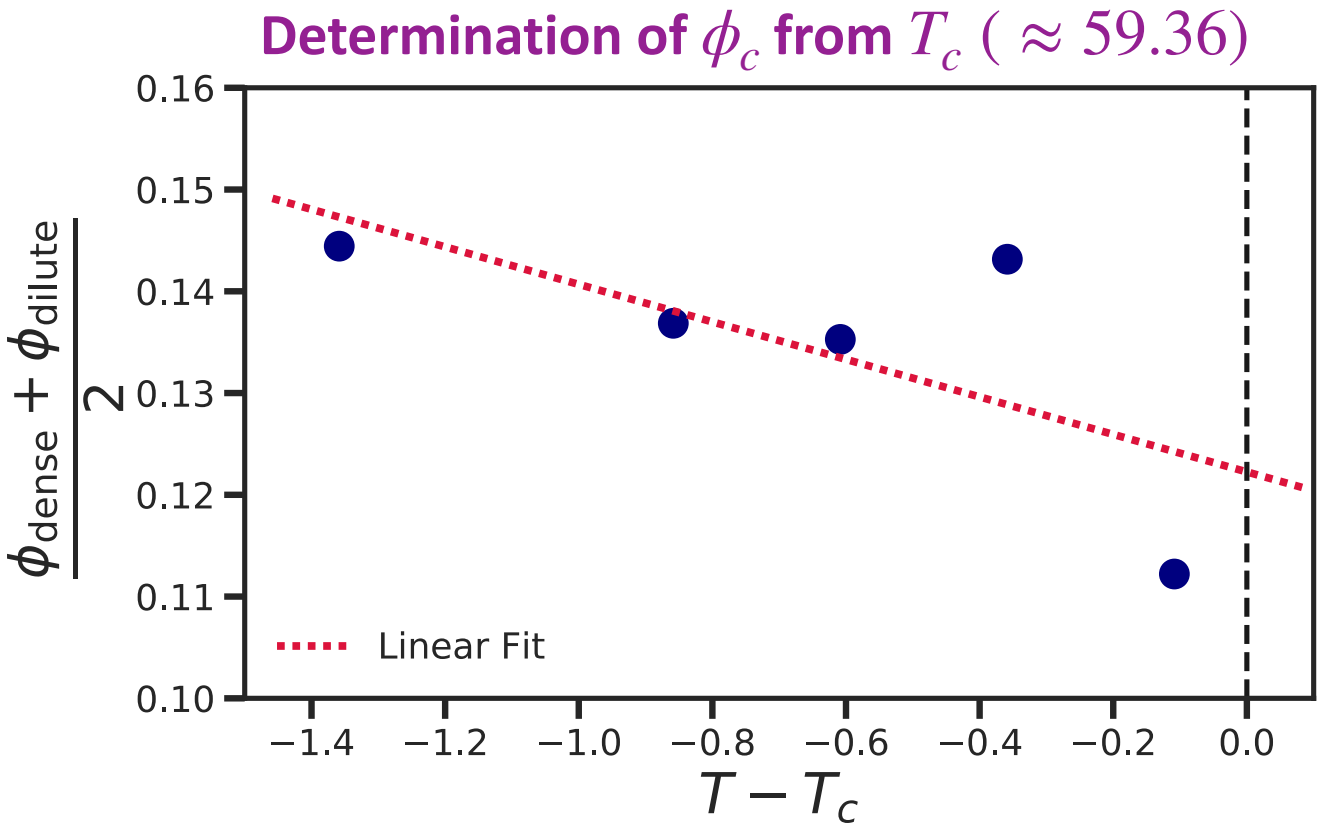

Fig. S8. Determination of the critical volume fraction ϕ_c. For this, we use the value of $T_c \approx 59.36$ estimated from the Binder cumulant analysis. ϕ_c is estimated using the law of rectilinear diameters, to perform a fit using linear regression over the points that are a part of the near-critical regime. The value of ϕ_c is obtained as the intercept of the linear fit is ≈ 0.12. The slope of the linear fit is $m \approx -0.018$ per unit T.

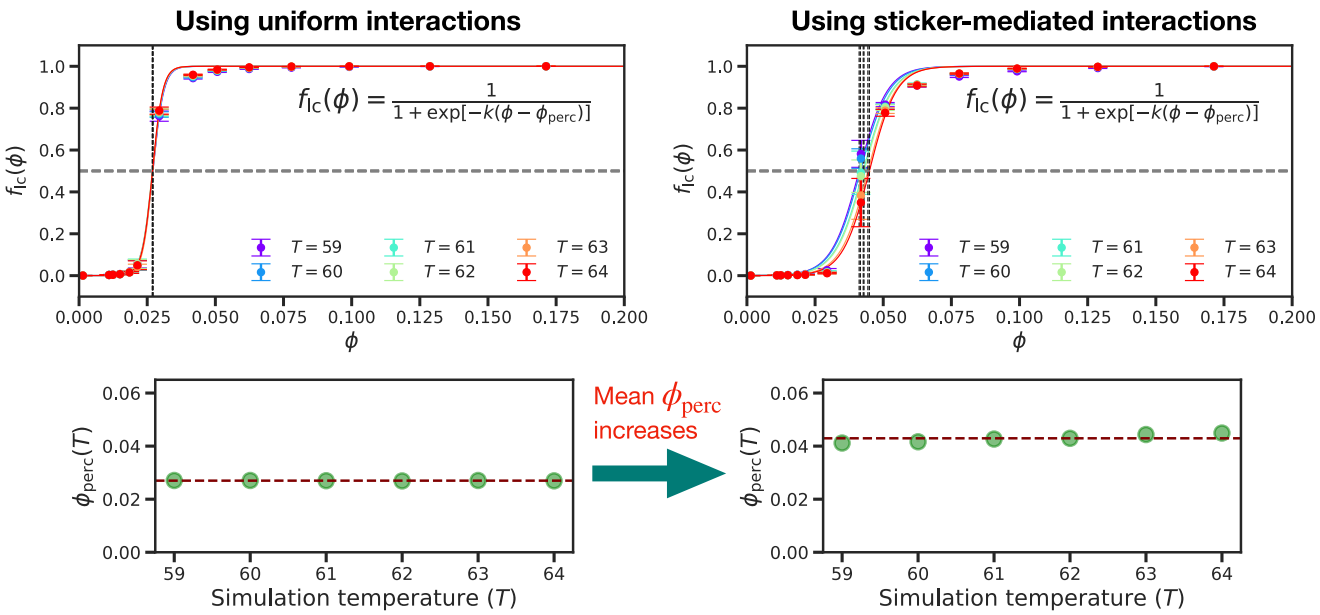

Fig. S9. Calculation of the percolation threshold ($\phi_{\mathrm{perc}}(T)$). This calculation uses information regarding the fraction of molecules in the largest cluster $f_{lc}(\phi) = \frac{n_{\mathrm{largest}}}{n}$, which we compute as a function of the simulation volume fraction ϕ. Here, $n = 10^4$ (the total number of A1-LCD molecules in the system). This analysis is performed for temperatures above and up to the critical temperature T_c for the cases when uniform interactions between all residues are considered, and when sticker-sticker interactions between only the strongest stickers tyrosine (Y), phenylalanine (F) and arginine (R) residues are taken into account. Percolation threshold ($\phi_{\mathrm{perc}}(T)$) as a function of simulation temperature T is also shown for the two cases, and the mean $\phi_{\mathrm{perc}}(T)$ increases for the latter.

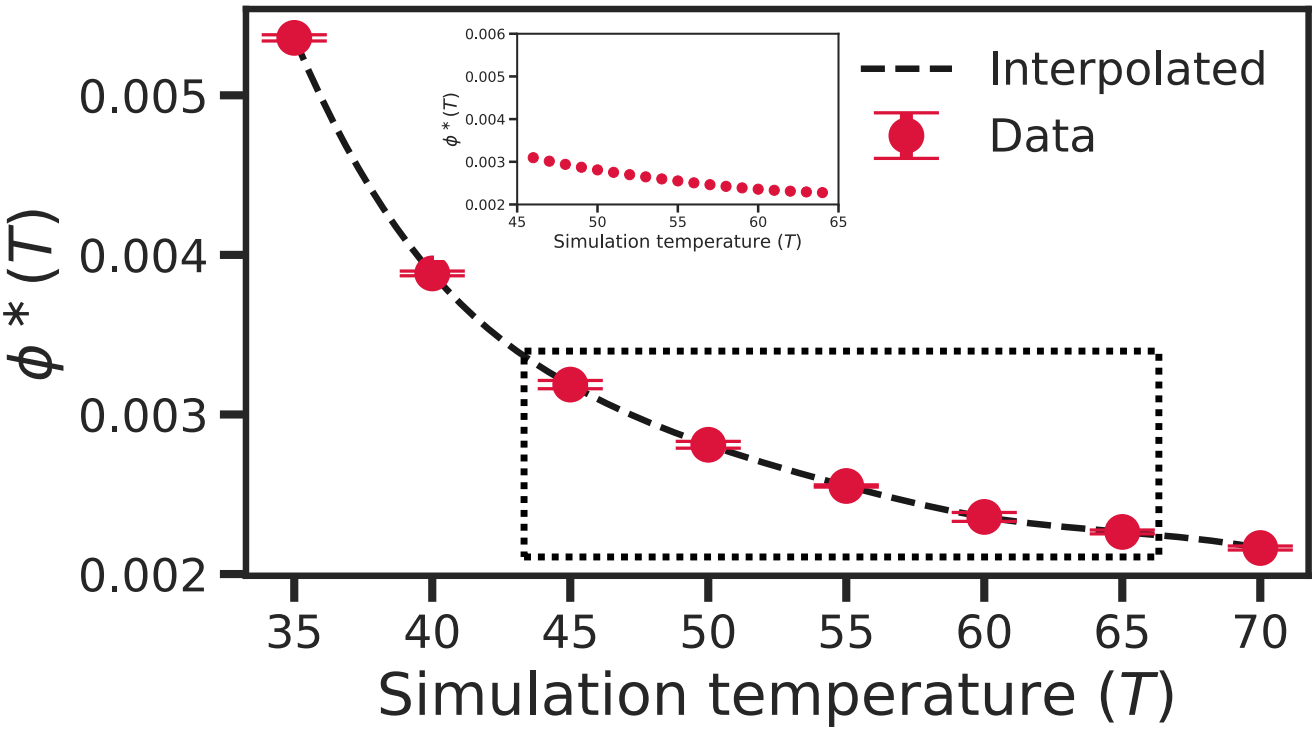

Fig. S10. Overlap volume fraction ϕ^* as a function of simulation temperature T. The data are shown here for the wild-type A1-LCD. The overlap volume fraction was calculated from single chain simulations. The inset shows the data over a zoomed-in range of T.

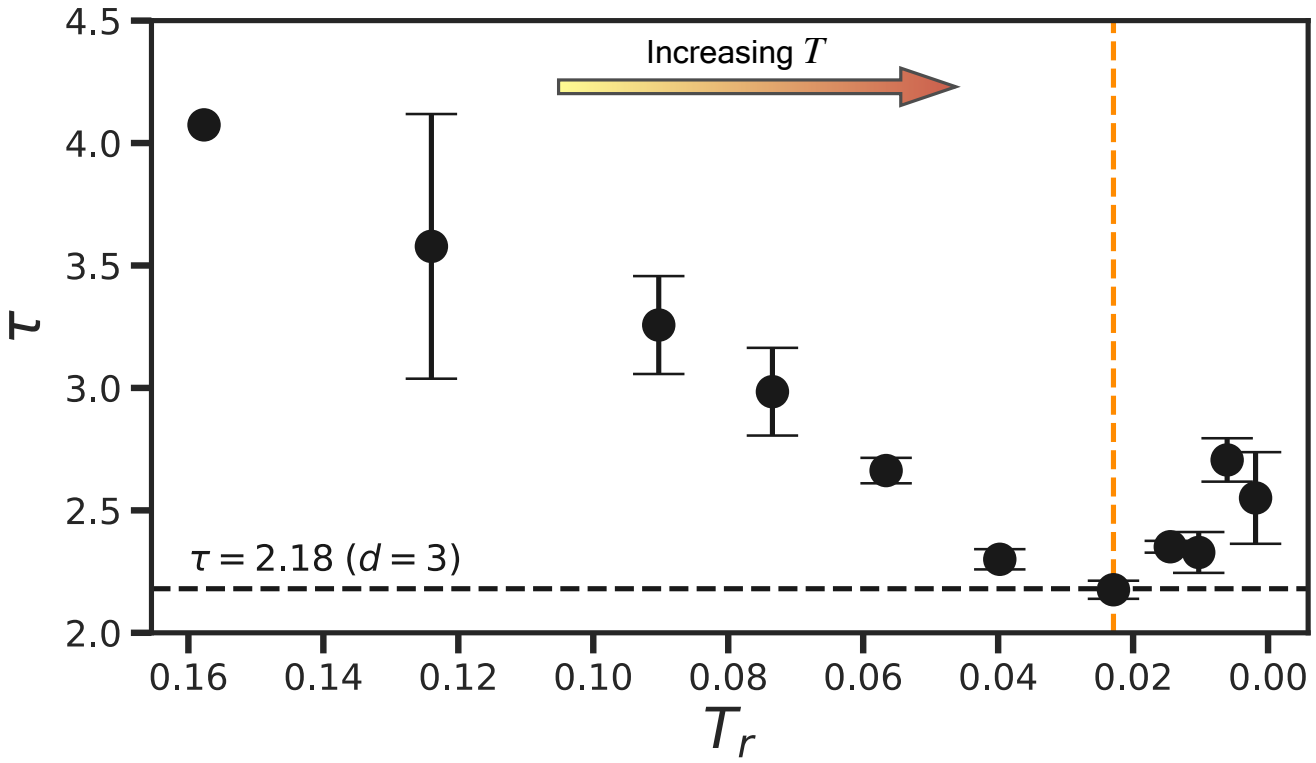

Fig. S11. Temperature dependence of the Fisher exponent for dilute phase cluster-size distributions. The exponent τ extracted from the power law fits of the dilute phase cluster-size distributions as a function of T_r. Each τ value was obtained from a linear regression of $\log_{10} p(n)$ versus $\log_{10} n$, and the associated error bars correspond to the standard errors from the covariance of the fitted slope. The orange vertical line marks $T_r \approx 0.023$ ($T = 58.0$), where $\tau \approx 2.18$, which is the universal critical exponent for three-dimensional percolation ($d = 3$). Above T_p, the cluster-size distributions show large-scale fluctuations. As a result, the delineation of dilute and dense phases becomes problematic because they are interconnected. This makes for noisy estimates of τ beyond the percolation threshold.

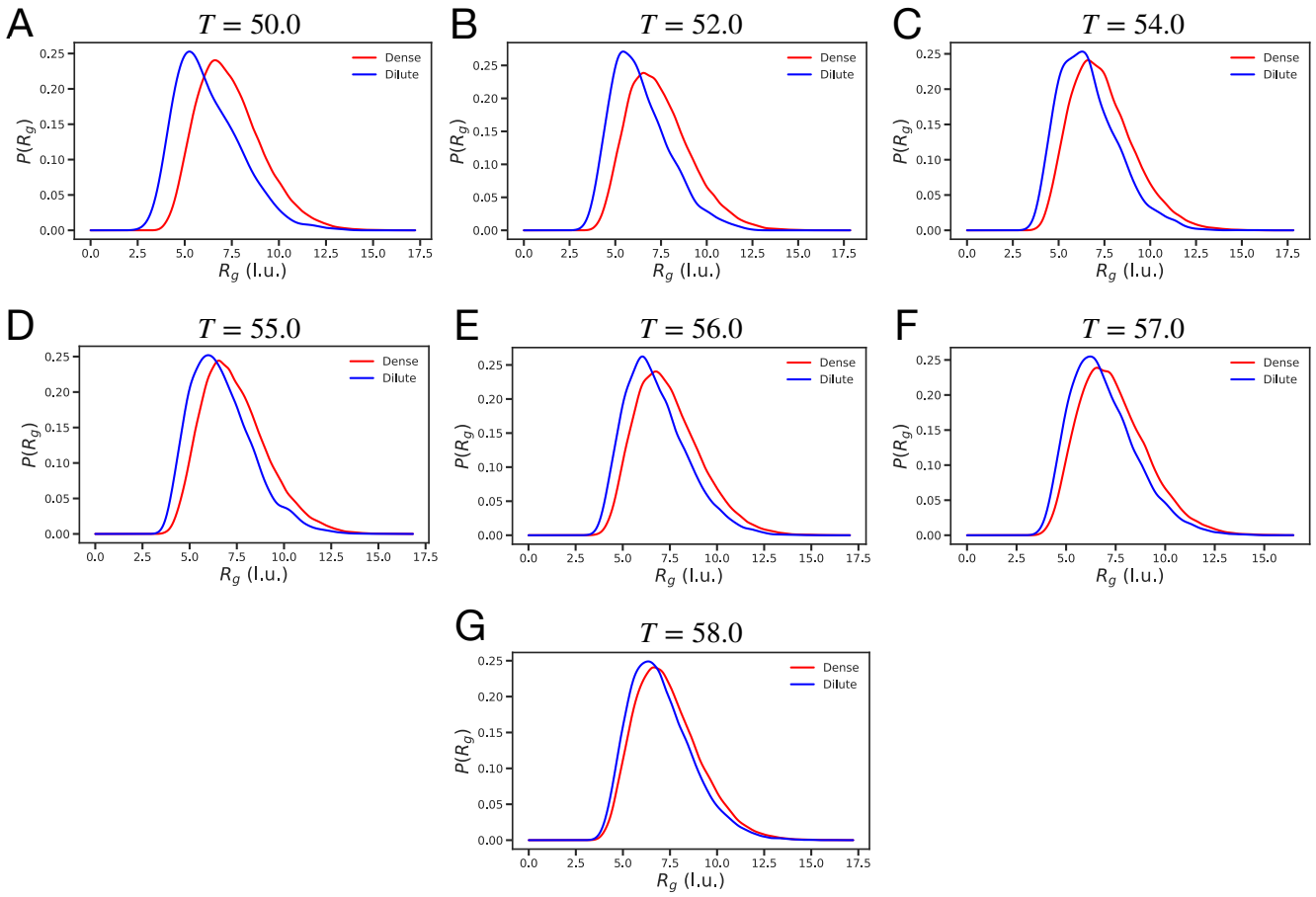

Fig. S12. Distributions of R_g values of individual molecules in the dense ($R_{g,\text{dense}}$) and dilute ($R_{g,\text{dilute}}$) phases. (A-G): Distributions of $R_{g,\text{dense}}$ and $R_{g,\text{dilute}}$ values for various simulation temperatures in Regime I and Regime II.

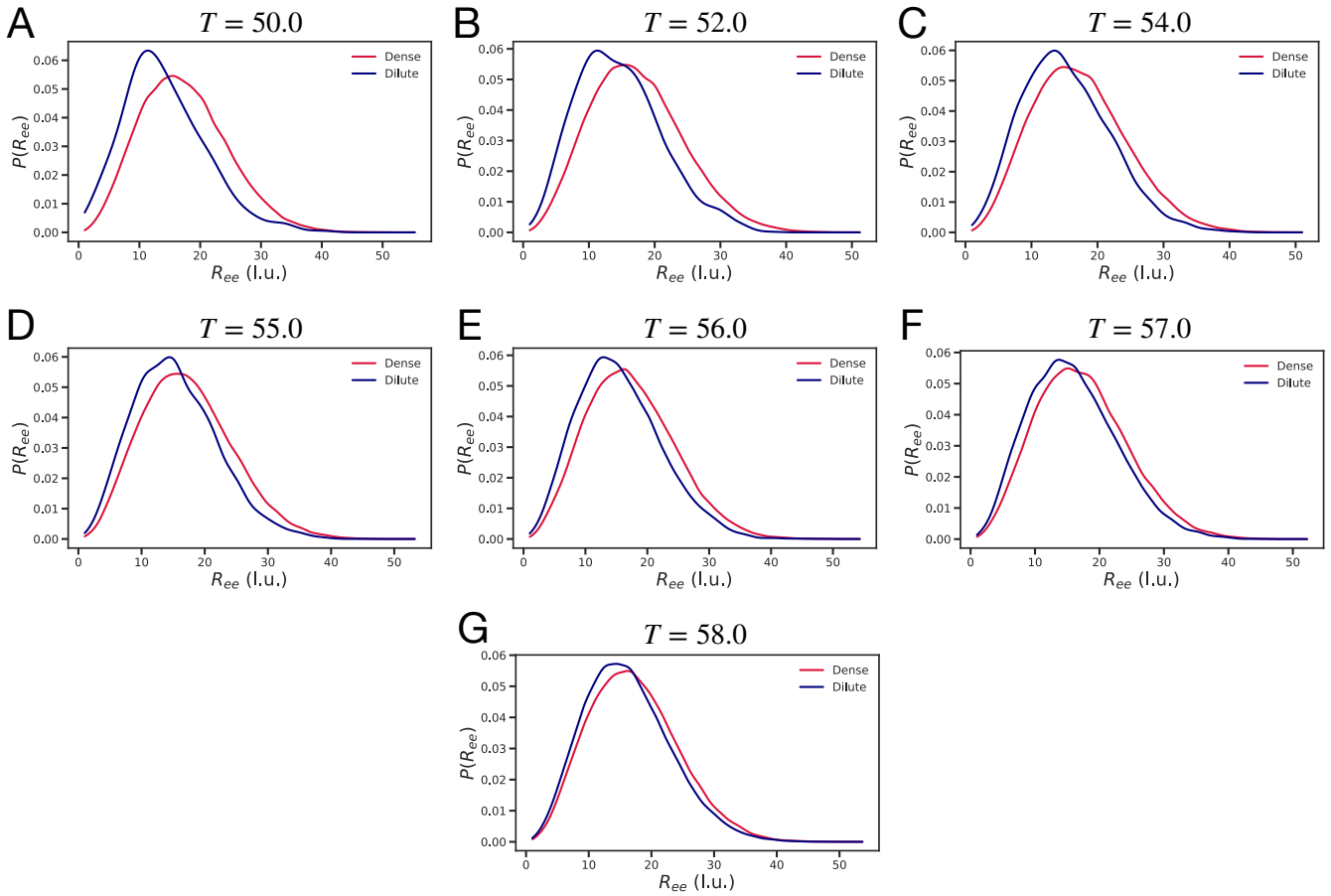

Fig. S13. Distributions of R_{ee} values of individual molecules in the dense ($R_{ee,\text{dense}}$) and dilute ($R_{ee,\text{dilute}}$) phases. (A-G): Distributions of $R_{ee,\text{dense}}$ and $R_{ee,\text{dilute}}$ for various simulation temperatures in Regime I and Regime II.

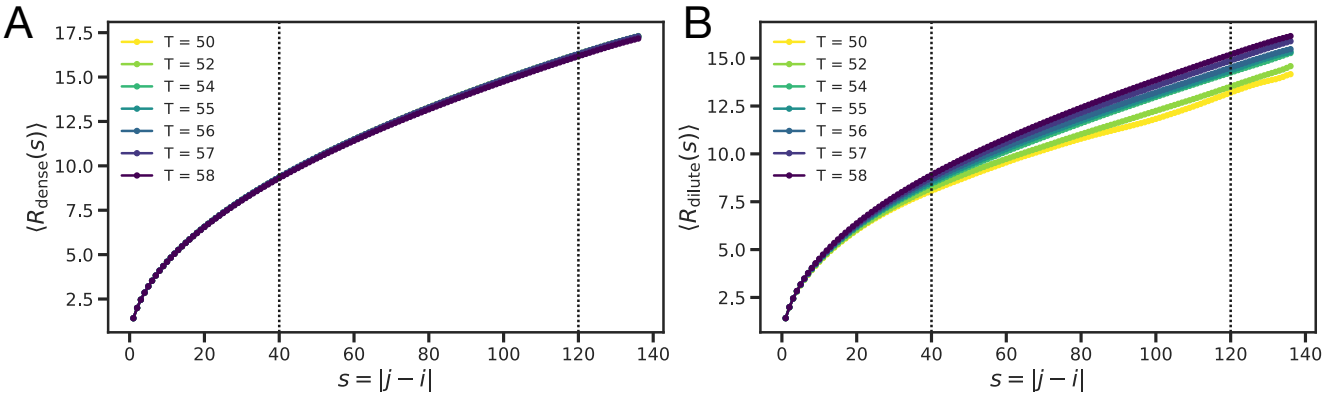

Fig. S14. Scaling of $R(s)$ versus s for individual chains in the dense and dilute phases. Plots showing the internal scaling of (A) $R_{\text{dense}}(s)$ versus s, (B) $R_{\text{dilute}}(s)$ versus s, for various simulation temperatures in Regime I and Regime II. The dotted vertical lines denote $s_l = 40$ and $s_u = 120$, which are bounds that were set for the scaling analysis used to extract ν.

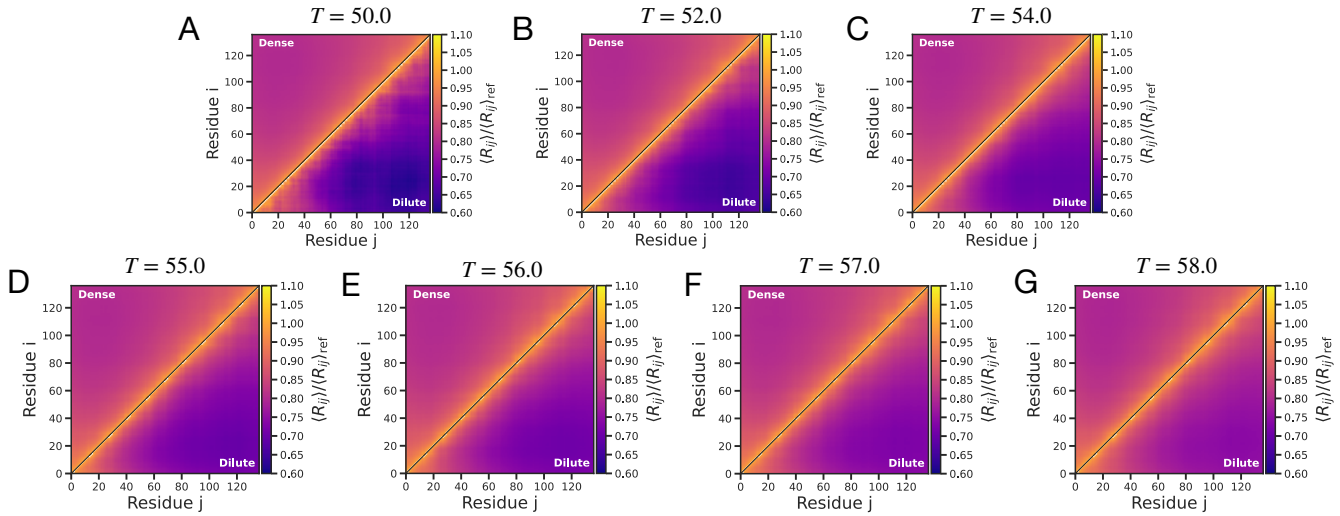

Fig. S15. Normalized inter-residue distance maps for individual chains in the dense and dilute phases. (A-G): Normalized inter-residue distance maps $R_{ij,\mathrm{dense}}$ (top-left triangle) and $R_{ij,\mathrm{dilute}}$ (bottom right triangle) for various simulation temperatures in Regime I and Regime II. The normalization for $\langle R_{ij} \rangle$ was performed with respect to the $\langle R_{ij} \rangle_{\mathrm{ref}}$ extracted from simulations of a single-chain of WT A1-LCD for which all the pairwise interaction energies between residues were set to zero.

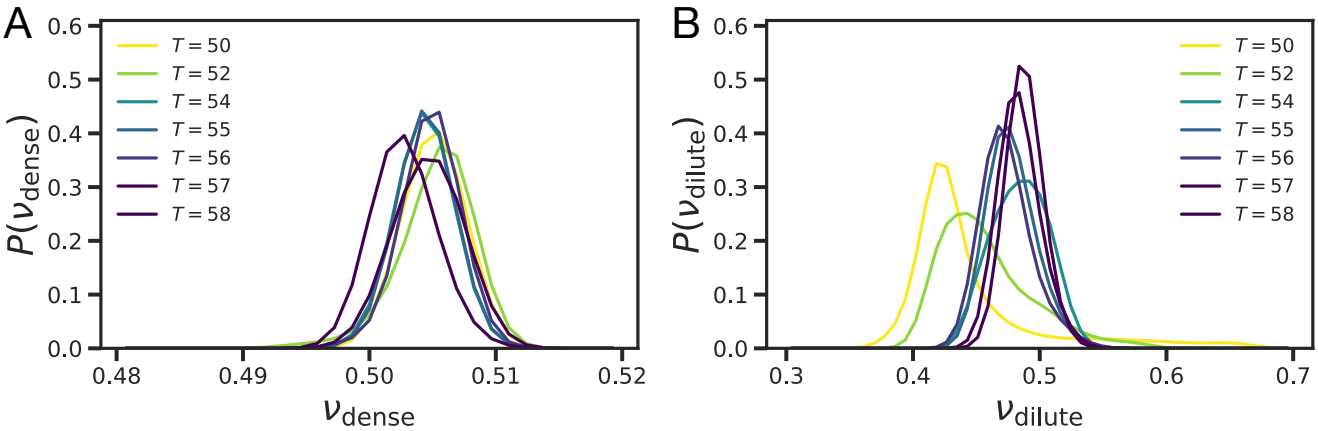

Fig. S16. Distributions of ν values for individual molecules in the dense and dilute phases. Distributions of (A) ν_{dense} and (B) ν_{dilute} for various simulation temperatures in Regime I and Regime II show that individual chains in both dense and dilute phases show behaviors close to Gaussian chain statistics. Note that the range of values along the abscissa for the dense and dilute phase ν distributions in (A) and (B) are different from each other.

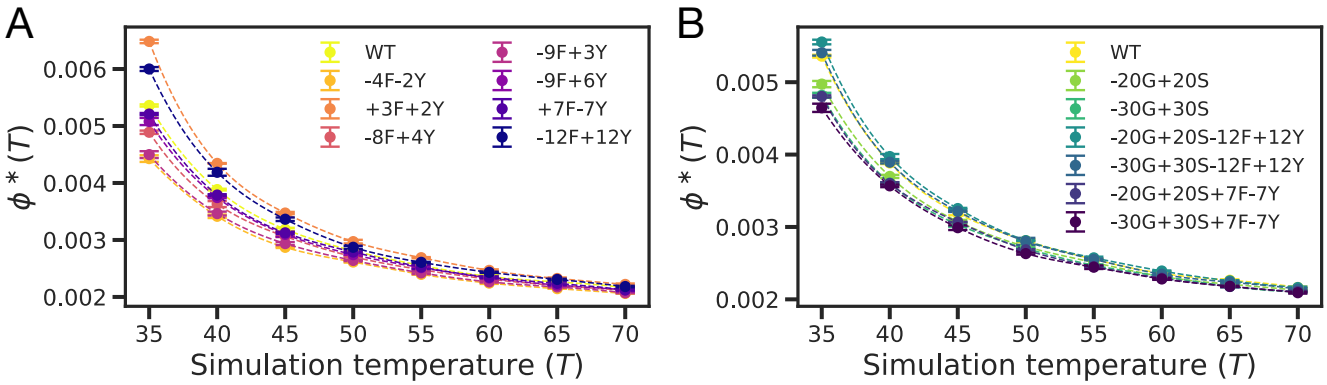

Fig. S17. Overlap volume fractions ϕ^* as a function of T for sequence variants designed by Bremer et al. (67). Overlap volume fraction ϕ^* as a function of simulation temperature T for (A) WT and 7 other aromatic variants of A1-LCD, and (B) WT and 6 other spacer variants.

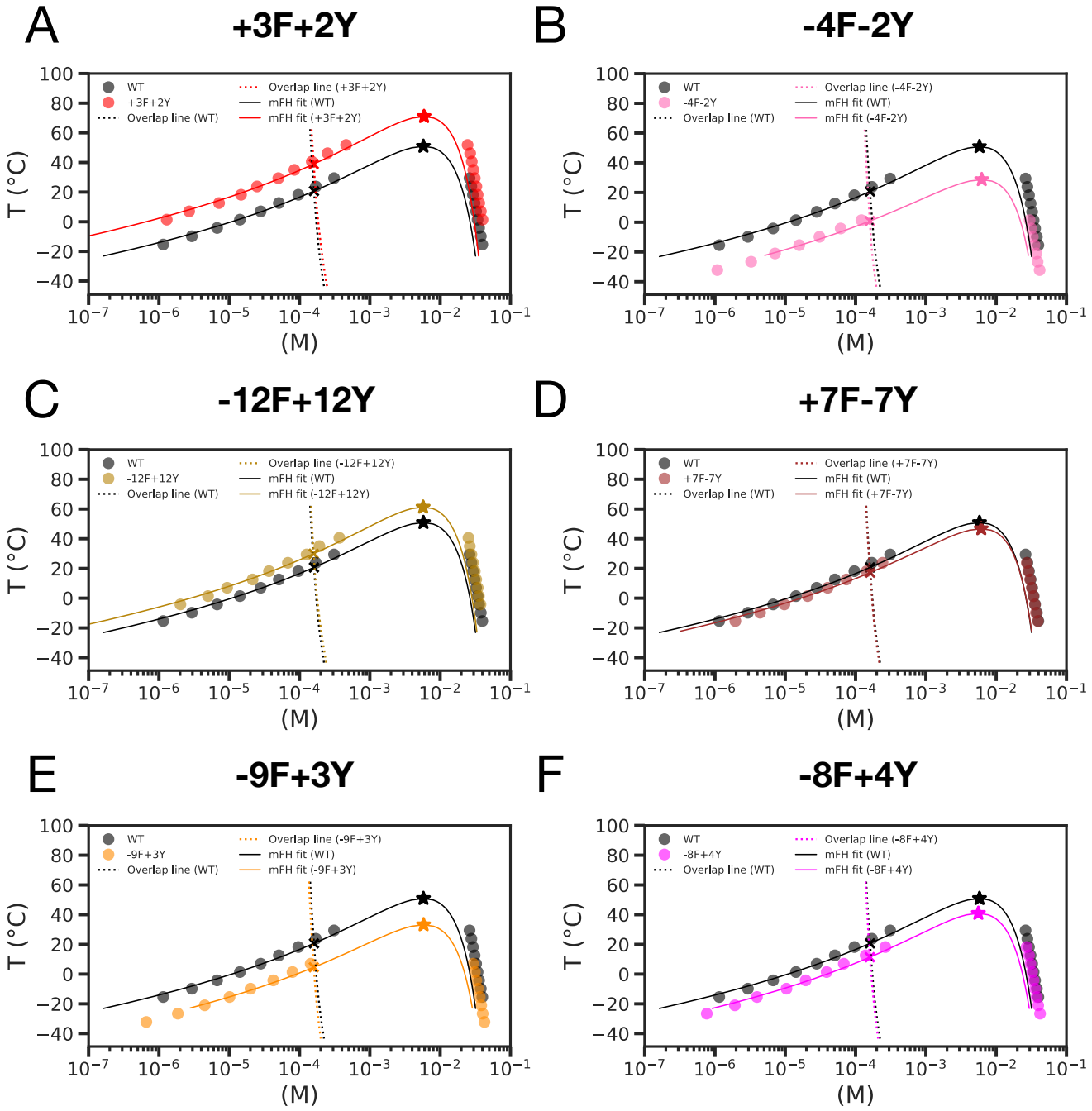

Fig. S18. Fits to the computed binodals for WT and other aromatic variants using modified Flory-Huggins model. (A-F): Binodals from simulations of 200 chains for various sequence-variants of WT A1-LCD. These variants were designed by altering the number of tyrosine (Y) and phenylalanine (F) residues relative to the composition of the WT. Solid lines are modified Flory–Huggins (FH) fits for each variant. For each variant, the binodal for the WT is also shown (black circles) as a reference, along with the modified FH fit for the WT (black solid curve). Also, the temperature-dependent overlap concentrations ($\phi^*(T)$) for the respective variant and the WT are shown as dotted line(s) that intersect the dilute arm of the binodal.

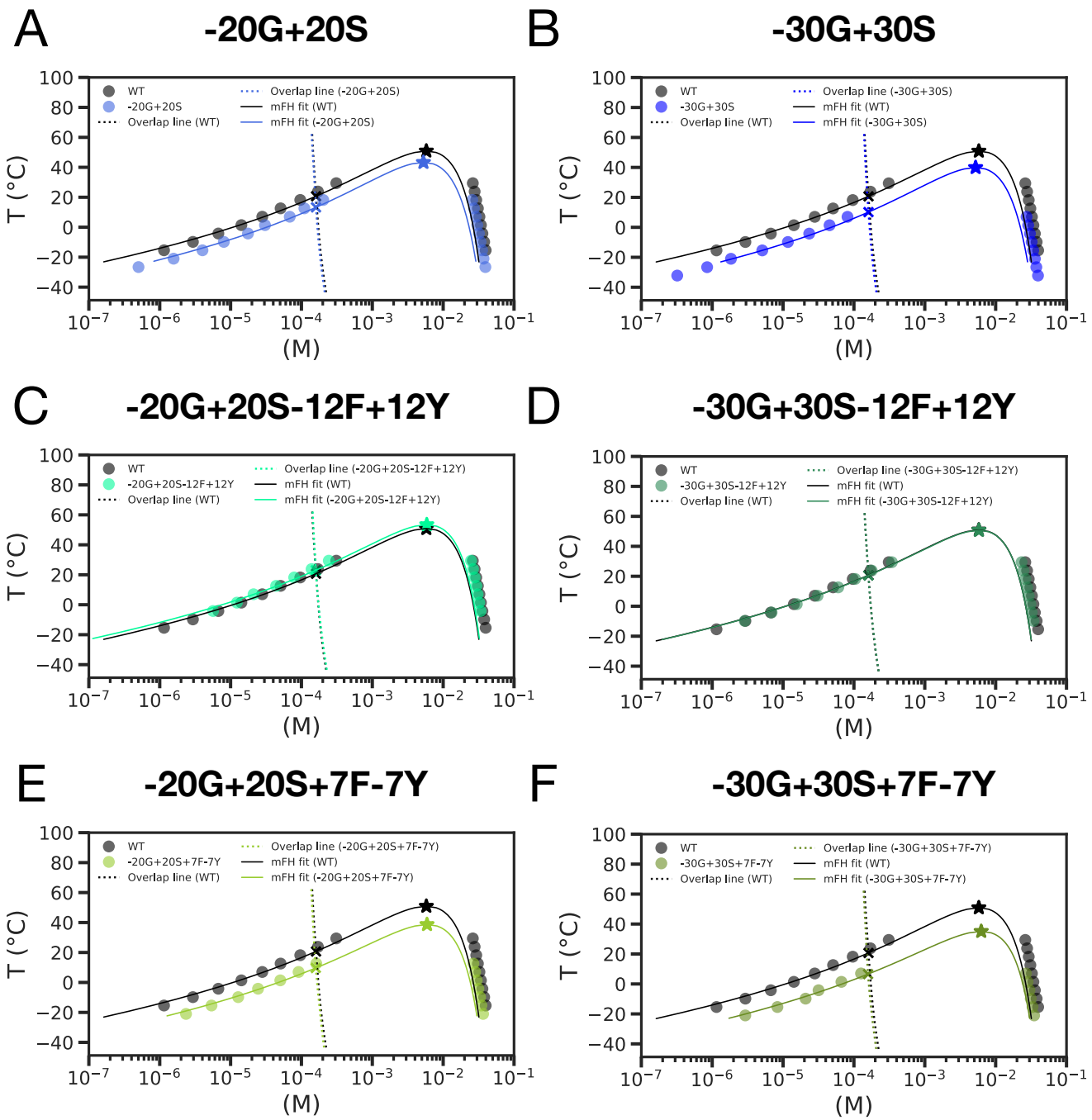

Fig. S19. Fits to the computed binodals for WT and other spacer variants using modified Flory-Huggins model. (A-F): Binodals from simulations of 200 chains for another set of sequence-variants of WT A1-LCD. These variants were designed by altering the number of spacer residues, namely glycine (G) and serine (S), relative to the WT composition. Solid lines are modified Flory–Huggins (FH) fits for each variant. For each variant, the binodal for the WT is also shown (black circles) as a reference, along with the modified FH fit for the WT (black solid curve). Also, the temperature-dependent overlap concentrations ($\phi^*(T)$) for the respective variant and the WT are shown as dotted line(s) that intersect the dilute arm of the binodal.

Table S1. Summary of the apparent critical temperatures ($T_{c,\mathrm{mFH}}$) and the overlap temperatures (T^*_{mFH}) for WT A1-LCD and its sequence variants obtained using the modified Flory–Huggins (FH) fits. These correspond to a system with $n = 200$. The entries in the table are sorted in descending order of the values obtained for $T_{c,\mathrm{mFH}}$ and T^*_{mFH}, with the Aro+ variant positioned at the top of the table, and Aro- at the bottom.
[a] T_c for WT A1-LCD for a system with $n = 10^4$ was determined to be $\approx 59.42°$C.
[b] T^* for WT A1-LCD for a system with $n = 10^4$ was determined to be $\approx 27.1°$C.

Sequence Variant	$T_{c,\mathrm{mFH}}$ **(in °C)**	T^*_{mFH} **(in °C)**
+3F+2Y (Aro+)	70.98	38.94
-12F+12Y (allY)	61.25	29.91
-20G+20S-12F+12Y	53.43	23.00
WT	50.89[a]	20.68[b]
-30G+30S-12F+12Y	50.71	20.52
+7F-7Y (allF)	46.73	17.21
-9F+6Y	45.28	15.62
-20G+20S	43.29	13.25
-8F+4Y	40.90	11.37
-30G+30S	39.95	10.16
-20G+20S+7F-7Y	38.69	9.66
-30G+30S+7F-7Y	35.11	6.57
-9F+3Y	33.20	4.61
-4F-2Y (Aro-)	28.69	0.81