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. The mapping of critical points is essential for generating quantitative assessments of driving forces and for distinguishing phase behaviors in the critical versus off-critical regimes. Computations play an important role in connecting the molecular-scale interactions to mesoscale phase behaviors of PLCDs and other intrinsically disordered proteins. We report results from accurate mapping of the critical regime for an archetypal PLCD. This is achieved by combining large-scale simulations with computations of Binder cumulants and the use of rigorous finite-size scaling approaches. The computed binodal, defined by knowledge of the critical point and intersection of the left arm of the binodal by the overlap and percolation lines, can be demarcated into three distinct regimes. Regime I is farthest from the critical point, with the coexisting 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, coexisting 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. Thus, Regime III comprises two interconnected, system-spanning networks. In addition to accurately mapping the critical point, we also evaluated methods for identifying the theta temperature. We find that conventional scaling approaches lead to erroneously low estimates of the theta temperature. Instead, accurate estimation of the theta temperature requires direct calculation of the temperature dependence of the two-body interaction coefficient. We discuss the broader implications of these findings for inferring solvent quality from scaling analyses.

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 process that explains the complete set of known 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 into coexisting dense and dilute phases via a process known as phase separation (6, 31). 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).

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 and solution components (31) across coexisting phases, and the degree of clustering and networking of molecules (46). Importantly, 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 signatures of near- versus off-critical behaviors of condensates (48, 49). To set up the relevant expectations for a one-component system, we turned to simulations and mapped the complete phase behavior of A1-LCD, which is an archetypal IDP. This is the intrinsically disordered PLCD from the protein hnRNP-A1 (Fig. 1A) (50).

Coarse-grained simulations have been used to map binodals for an assortment of systems (17, 20, 44, 51–57). These off-

lattice molecular dynamics (MD) simulations and on-lattice Monte Carlo (MC) simulations typically include $O(10^2)$ molecules per simulation cell. In the off-lattice MD simulations, the computed dilute and dense phase arms of the binodal are connected by a numerical fitting procedure (58). The system is assumed to belong to the same universality class as the 3-dimensional Ising model. The binodal is then fit to the scaling of the order parameter prescribed by the Ising model, and the critical temperature is estimated through the fitting procedure. Here, we assess the mapping of critical points using the approach established in the literature, and then examine the accuracy of this method by performing large-scale simulations. We find it imperative to deploy rigorous finite size scaling based on analysis of Binder cumulants (59). This analysis yields accurate estimates of the critical point. We use this approach to map the full binodal of A1-LCD and demonstrate how the percolation and overlap lines intersect the coexistence curve (binodal). We identify three distinct regimes, which are characterized by distinct, regime-specific cluster-size distributions in the coexisting dilute phases. We also estimate the theta temperature as the point where the two-body interaction coefficient (B_2) for a pair of chains equals zero. We find that the temperature where this criterion is met is very different from the apparent theta temperature inferred by analysis of the scaling of mean segmental distances $R(s)$ as a function of s, the length of a segment.

Methods

Our coarse-grained lattice-based Monte Carlo (MC) simulations (61–63) employ the LaSSI simulation engine (46, 56), which is based on a generalization of the bond fluctuation model (64, 65). A single A1-LCD molecule comprises 137 beads, one for each of the residues, and each residue is a single bead on a lattice. For the inter-residue potentials (Fig. 1B, Fig. 1D), we used a short-range potential function that was parameterized using a Gaussian Process Bayesian Optimization procedure (56, 66). The parameters are transferable unto other PLCDs but not to other systems (56, 57). To map the critical region, we used finite-size scaling that relies on the quantification of density fluctuations via analysis of Binder cumulants (59, 67). We used a nine-step protocol, which is followed sequentially to arrive at a complete characterization of the binodal of A1-LCD (Fig. 2). 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. 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 is $O(10^{11})$. Based on these considerations, the simulation system includes 10^4 A1-LCD molecules for each simulation temperature T_s. These simulation temperatures T_s are converted to Kelvin by multiplying them by a factor of 5.6 (56) (see Section 2.2). For every value of T_s, we performed at least three independent MC simulations with different random seeds. Each simulation comprises 6×10^{11} MC steps. In the currency of MD simulations, 6×10^{11} MC steps would translate to simulation times of ~ 3 ms assuming an integration time step of 5 fs. Additional simulations of equivalent length (for more temperatures in the critical region) were performed when the Binder cumulant analysis needed to be deployed. For computing binodals as a function of 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.

LaSSI simulations. The lattice-based MC simulations were performed following the approach of Farag et al., (56, 57). Beads occupy positions on a cubic lattice, and the energy of interaction is $-\varepsilon_{ij}$ (Fig. 1B, Fig. 1C) 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., (56) (Fig. 1D). The energies were parameterized to reproduce the sequence specificity of measured radii of gyration (60) and the parameterization scheme was based on the Gaussian Process Bayesian Optimization approach of Ruff et al. (66). We analyzed the results from two sets of simulations, one that includes 200 A1-LCD molecules in a cubic box ($L_{\text{sim}} = 120$ l.u.) and the other that includes 10^4 molecules ($L_{\text{sim}} = 240$ l.u.). LaSSI simulations to study the phase behaviors were performed for $46 \leq T_s \leq 64$ ($257.60\,\text{K} \leq T \leq 358.40\,\text{K}$, where T is the temperature in Kelvin).

System size considerations for the LaSSI simulations. To discern the optimal box size and numbers of A1-LCD molecules (n) that are 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 a 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 large due to 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

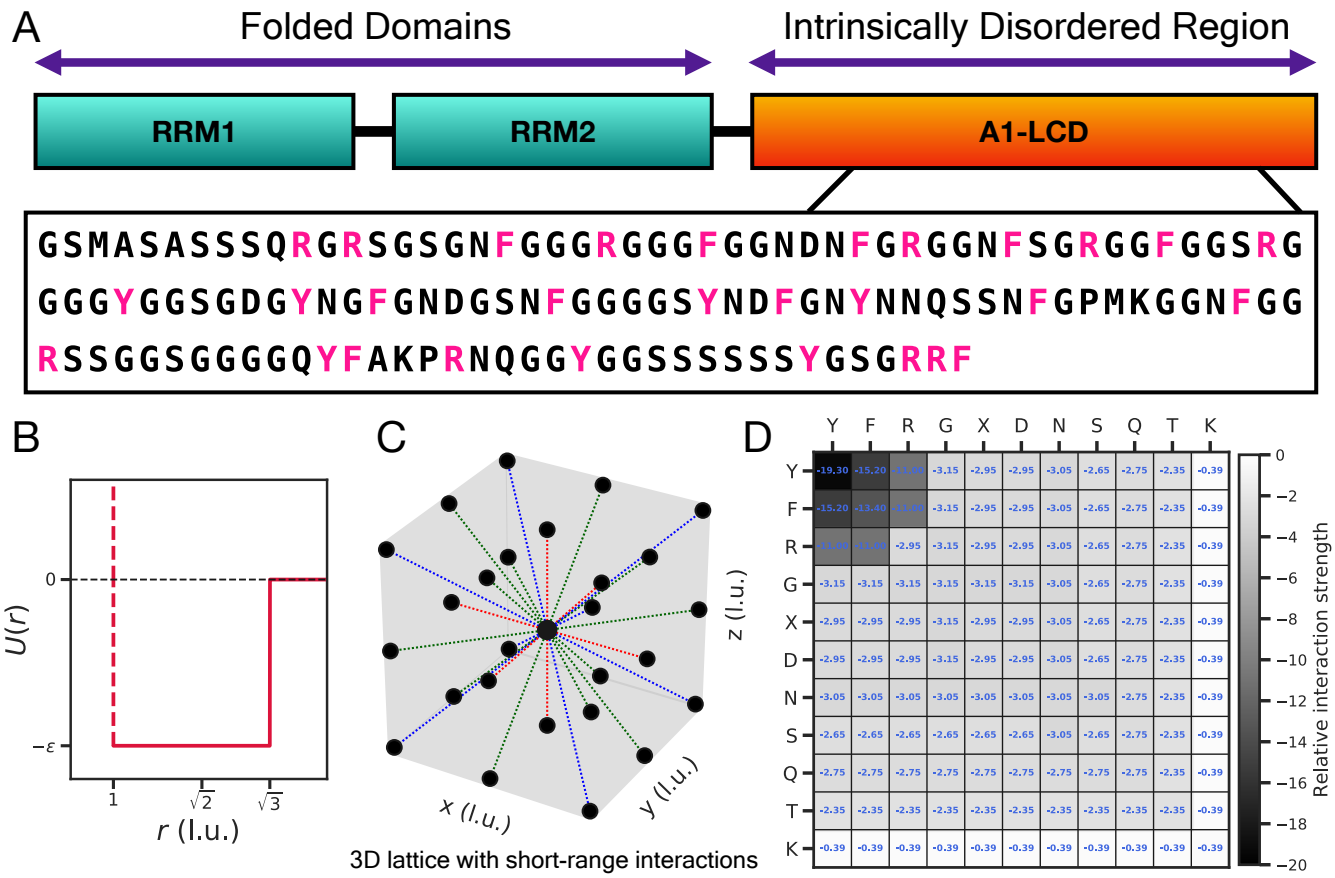

Fig. 1. Details of the coarse-grained lattice-based Monte-Carlo simulations. (A) Architecture of the hnRNP-A1 protein and amino acid sequence of A1-LCD. hnRNP-A1 comprises the folded RNA Recognition Motifs (RRMs) and the PLCD (Prion-Like Low Complexity Domain), which is designated as A1-LCD. The amino acid sequence of A1-LCD, which has 137 residues, is also shown in the box. Sticker residues in the A1-LCD sequence were identified in previous studies (56, 60), and these are tyrosine (Y), phenylalanine (F), and arginine (R), which are colored in magenta to delineate them from the rest of the residues (colored in black). (B) Pairwise energy of interaction shown as a function of the distance between any two beads on the lattice (See Methods section). (C) Schematic of a cubic lattice with short-range interactions where each bead 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). (D) Matrix showing the pairwise interaction strengths used in the LaSSI computational model. "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 $k_B T_s$, where $k_B = 1$ and T_s is the simulation temperature.

sampling of a broader spectrum of density fluctuations.

To estimate the appropriate numbers of molecules and box sizes, we used 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 A1-LCD molecules. Since $n \propto L_{\text{sim}}^3$, to determine the minimum number of molecules that would be required to map phase behaviors close to criticality, the box length $L_{\text{sim, target}}$ at $\phi \approx \phi_c$ should 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

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

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

Simulations with $n \approx 10^4$ lie 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.

Conversion from simulation to experimental units. Scaling factors are required to convert the computed coexistence curves from simulation temperatures into degree Kelvin and volume fractions into molar units. For the latter, we followed Wei et al. (38) and Farag et al. (56). A volume fraction of 1.0 corresponds to a mass concentration of 1310 mg/ml. Following Farag et al. (56), 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. Farag et al. derived this factor 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, 56). 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 (56, 60) and to the LCD of FUS (57). 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, 56). 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 corresponds to the average distance between adjacent C_α atoms ($3.8 - 4.0$ Å) (69, 70). 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.

Protocol for accurately mapping the critical regime and estimating critical parameters. **Step 1** involves setting up simulations of the appropriate size (box size and numbers of molecules) and length (number of MC steps). **Step 2** is the extraction of the interfacial width $\Delta(T)$ and the midpoint of the interface 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 (71). **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 for temperatures closer to the critical region, to estimate T_c. This knowledge can be used to estimate ϕ_c directly using the full binodal. As an additional route to estimating ϕ_c, **Step 6** uses the known T_c to calculate ϕ_c via 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** quantifies the percolation and overlap lines, and estimates the points where these lines intersect the binodal. All steps of this protocol are shown in the flowchart in Fig. 2.

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 profile of $d\phi(r)/dr$ exhibits a peak centered at the location of the interface. We fit this peak to a Gaussian function (72),

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

where r_{mid} is the midpoint of the interface and σ is the width of the Gaussian. The full width at half maximum (FWHM) of the fitted Gaussian is taken to be 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 (71), 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 temperature. As an important aside, we note that Farag et al., (56) 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. 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 (56). 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 adopted the two-step approach introduced by Chauhan et al., (72). We first used the derivatives of the temperature-specific radial density profiles $\phi(r)$ to estimate $\Delta(T)$ and $r_{\text{mid}}(T)$ (see Fig. S1 in *SI Appendix*). 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}.

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. 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:

$$m_L = \phi_L(T) - \langle \phi_L(T) \rangle \quad \textbf{(1)}$$

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

$$U_L(T) = 1 - \frac{\mu_4(L;T)}{3\mu_2(L;T)^2} \quad \textbf{(2)}$$

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 $U_L \to 2/3$ (59, 67, 73). 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

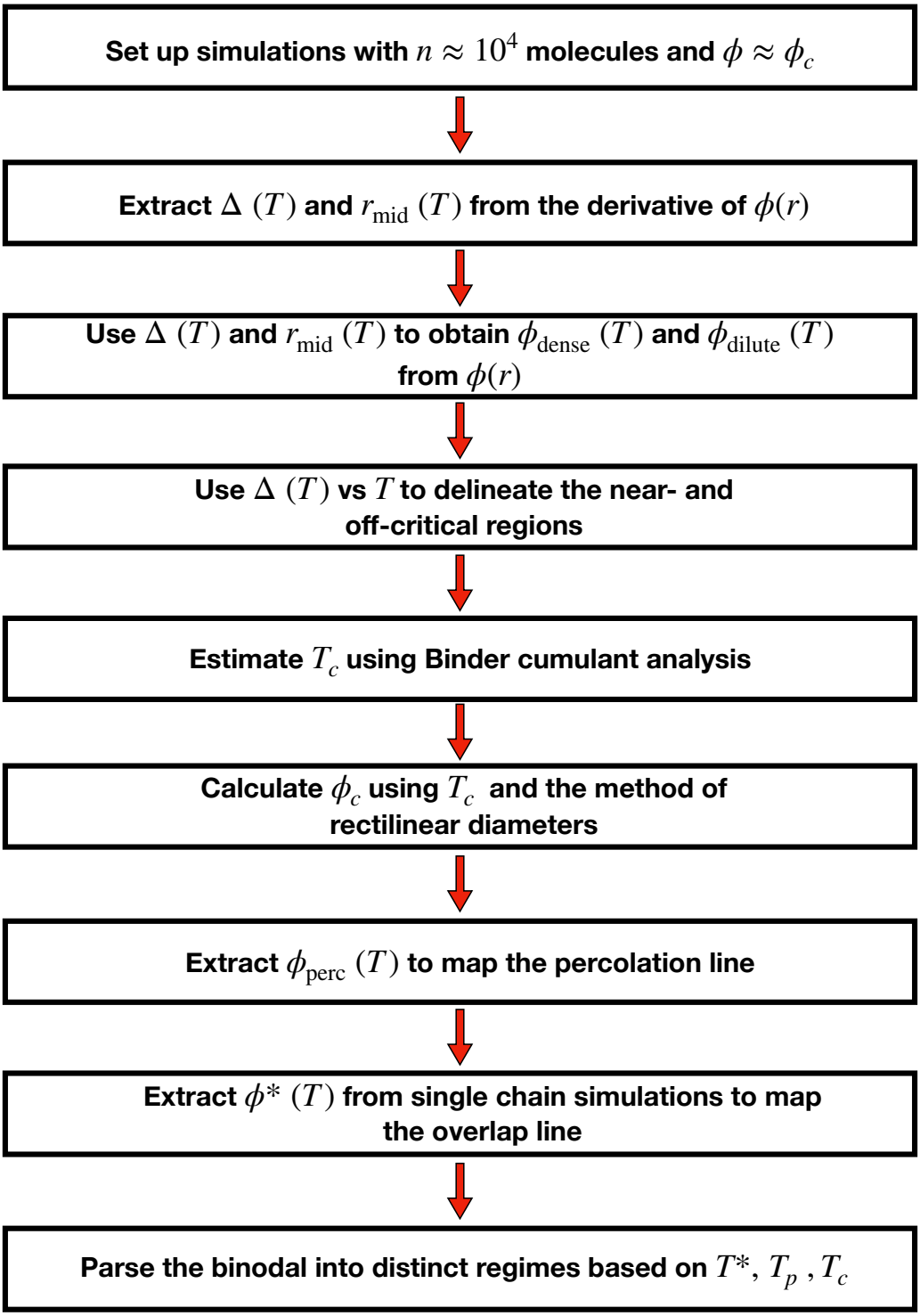

Fig. 2. Flowchart showing the nine-step simulation and analysis protocol to map the full binodal for WT A1-LCD. Details of the protocol are provided in the Methods section of the main text.

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 mean end-to-end distance of a single polymer (R_{ee}), while remaining small enough relative to L_{sim} to provide sufficient statistical sampling of density fluctuations. The upper limit of R_{ee} corresponding to the highest temperature in our simulations is ≈ 17 l.u. as shown in Fig. S2 in *SI Appendix*. Thus, the range of sub-box sizes we used for the Binder cumulant analysis ($L \approx 50 - 200$ l.u.) fall within acceptable limits.

Sampling of sub-boxes and calculation of sub-box densities. 10^4 cubic sub-boxes of size L were placed uniformly at random locations 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 $\phi_i = \frac{1}{L^3} \sum_{\mathbf{r} \in V_i} \phi(\mathbf{r})$, 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^* (59). 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, 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 (74) 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 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).

Percolation threshold as a function of temperature. To determine the percolation threshold ϕ_{perc} at a given temperature, we compute the percolation order parameter defined as the fraction of chains belonging to the largest connected cluster (lc), following the approach used by Choi et al., (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 $f_{\text{lc}}(\phi)$ versus 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 (see Fig. S4 in *SI Appendix*).

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 (see Fig. S2 and Fig. S5 in *SI Appendix*),

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

and then estimating the 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.

Defining clusters based on connectivity of stickers. Clustering of A1-LCD chains is quantified using a sticker-mediated connectivity criterion applied to individual simulation snapshots. "Sticker" residues (17, 29, 60, 75) are defined as tyrosine (Y), phenylalanine (F), and arginine (R), as these residues are known to be primary drivers of intermolecular interactions in PLCDs (17, 60). The choice of these residues as stickers is supported by the hierarchy of interaction strengths encoded in the simulation energy matrix (see Fig. 1D), wherein Y–Y, F–F, Y–F, and R–Y, R–F interactions constitute the most favorable non-bonded contacts. Accordingly, only beads corresponding to Y, F, and R residues are considered when constructing intermolecular contacts. For each snapshot, sticker bead coordinates are wrapped under periodic boundary conditions and inserted into a lattice-based spatial lookup table. Two chains are considered connected if at least one sticker bead from one chain occupies a nearest-neighbor lattice site (including axial and diagonal neighbors, corresponding to distances of 1, $\sqrt{2}$, or $\sqrt{3}$ l.u.) of a sticker bead from another chain. This procedure generates a chain-level adjacency graph in which nodes represent chains and edges represent at least one inter-chain sticker contact (56, 76, 77). Connected components of this graph are identified using a depth-first search (DFS) algorithm, and each connected component is defined as a cluster. Note that even at the highest concentrations, less than 10% of the

lattice volume is taken up by chain molecules. Additionally, the bulk A1-LCD concentrations are identical across all temperatures, implying that the cluster-size distributions that we extract from the simulations are emergent properties of the phase behaviors and not a trivial consequence of including 10^4 molecules in the simulations.

Calculation of the scaling exponent $\nu(T)$ from the ensemble-averaged segmental distances $R(s)$. For each temperature T, the ensemble-averaged internal scaling profile $R(s)$ was first obtained by averaging over equilibrated simulation snapshots and across three independent replicates. The scaling exponent ν was then extracted from this temperature-specific profile using a finite-difference formulation in logarithmic space. Specifically, for segment pairs corresponding to contour separations s_l and s_u, we computed

$$\nu = \frac{\ln R(s_u) - \ln R(s_l)}{\ln s_u - \ln s_l}$$

following the approach described in (78). This expression estimates the local slope of $\ln R(s)$ versus $\ln s$, thereby yielding an effective scaling exponent without imposing a global regression. The analysis was restricted to intermediate segmental separations satisfying $40 \leq s_l, s_u \leq 120$. The lower bound excludes locally stiff segments that do not obey asymptotic scaling behavior, whereas the upper bound avoids artifacts arising from finite chain length and dangling chain ends. Only segment pairs satisfying $R(s_u) > R(s_l)$ were retained to ensure physically meaningful positive slopes. This procedure yields a distribution of ν values at each temperature, from which the mean and standard deviation were computed. The scaling analysis was performed only for temperatures above the "globule" regime where the internal scaling profiles become effectively flat. At lower temperatures, strong chain collapse results in nearly s-independent $R(s)$ profiles, rendering the logarithmic derivative ill-defined and thus preventing the meaningful extraction of a scaling exponent. We therefore, restrict the scaling analysis to all temperatures $T > 140$ K.

Computing the temperature dependence of the two–body interaction coefficient B_2. To determine the temperature dependence of the two-body inter-chain interaction coefficient B_2 for A1-LCD, we performed umbrella sampling simulations of a pair of chains across a range of temperatures T. The reaction coordinate was defined as the separation between the centers-of-mass of the two chains. Harmonic biasing potentials were applied at multiple restraint distances to ensure sufficient overlap between adjacent windows (see Fig. S6 in *SI Appendix*). The Weighted Histogram Analysis Method (WHAM) (79–81) was then used to combine biased sampling data from multiple overlapping umbrella windows to self-consistently reconstruct the unbiased potential of mean force $W(r)$ for each temperature. The Mayer f-function is defined as (82, 83):

$$f(r) = \exp(-\beta W(r)) - 1$$

where $\beta = \frac{1}{k_B T}$. The temperature-dependent two-body interaction coefficient is proportional to the excluded volume $\mathrm{v_{ex}} \propto B_2$:

$$\begin{aligned}\mathrm{v_{ex}}(T) &= -2\pi \int_0^\infty f(r)\, r^2\, dr \\ &= -2\pi \int_0^\infty \Big[\exp(-\beta W(r)) - 1\Big] r^2\, dr,\end{aligned}$$

where the upper limit of integration is chosen such that $W(r) \to 0$. Repeating this procedure at different temperatures yields $\mathrm{v_{ex}}(T) \propto B_2(T)$, thus providing a quantitative measure of how the effective pairwise attractions between A1-LCD chains evolve with temperature. We compute $\mathrm{v_{ex}}(T)$ and then estimate $B_2'(T)$ as the ratio of $B_2(T)$ at a temperature T to the magnitude of B_2 corresponding to the highest temperature in our simulations.

Results

Assessing finite-size artifacts. 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. Here, ϕ_{dense} is the volume fraction of the dense phase, ϕ_{dilute} is the volume fraction of the dilute phase, T is the simulation temperature, and T_c is the critical temperature. The procedure that is typically followed 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, 51, 53–55, 84, 85). Note that the scaling of the order parameter is only valid in the vicinity of the critical point (86). Accordingly, some authors exercise caution in the use of the regression analysis (87), although this will be affected by the paucity of points that are sampled in the vicinity of the critical point from simulations of finite-sized systems.

To assess whether the vicinity of the critical regime is actually being sampled in simulations of finite-sized systems, we examined the T-dependent radial density profiles for a system of 200 A1-LCD molecules (Fig. 3A). A defining feature of the critical regime is that the concentration of the dilute phase should approach that of the dense phase. Concentrations extracted from these profiles using the Fisk-Widom function show that at all temperatures sampled, the dilute and dense phase concentrations differ by over an order of magnitude (circles in Fig. 3B, Fig. 3C). This suggests a lack of data points around the critical point. Additionally, as described in the Methods section, sampling the critical regime for a simulation with 200 molecules of A1-LCD would require a box size of $L_{\mathrm{sim}} \approx 65$ l.u. which imposes finite-size artifacts.

To work around this paucity of points, the common practice is to enforce the scaling of the order parameter across the entirety of the computed binodal. The reduced temperature is defined as

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

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:

$$\frac{\phi_{\text{dense}}(T)+\phi_{\text{dilute}}(T)}{2}=\phi_c-mT_cT_r,$$

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 $\beta \approx 0.33$:

$$\Delta\phi(T)=\phi_{\text{dense}}(T)-\phi_{\text{dilute}}(T)=A_oT_r^{\beta},$$

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

$$\phi_{\text{dense}}(T)=\phi_c-mT_cT_r+\tfrac{1}{2}A_oT_r^{\beta},$$
$$\phi_{\text{dilute}}(T)=\phi_c-mT_cT_r-\tfrac{1}{2}A_oT_r^{\beta}.$$

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 by utilizing the 200 chain A1-LCD system (Fig. 3B, Fig. 3C). 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. This is congruent with the fact that this scaling should only apply in the vicinity of the critical point. Next, we fit the computed binodal to a modified Flory-Huggins model (17, 60, 68). While this method generates a mean field scaling of the order parameter in the vicinity of the critical point (88), its use for estimating the location of the critical regime is appropriate. The binodal obtained by fitting the data to the modified Flory-Huggins model does a better job, especially in terms of fitting the dilute arm. The estimated T_c is considerably higher than what is estimated by imposing scaling of the order parameter based on the 3D Ising model. These discrepant estimates, along with the radial density profiles, suggest that the critical regime is inadequately sampled or not sampled at all in simulations based on 200 A1-LCD molecules, thus emphasizing the challenges posed by finite size artifacts (89) (Fig. 3).

Mapping the binodal for large systems and assessing its accuracy. For the simulations that include 10^4 A1-LCD molecules, the presence of a two-phase system in the simulation cell is evident from the temperature-specific radial density profiles $\phi(r)$ (Fig. 4). A hyperbolic tangent, Fisk-Widom function (56, 71) is fit to the T-dependent profiles for $\phi(r)$ to estimate $\phi_{\text{dense}}(T)$ and $\phi_{\text{dilute}}(T)$ for the system of 10^4 A1-LCD molecules (Fig. 4). The simulations for a system with 10^4 chains show that the envelope of the two-phase regime extends to higher temperatures. As T increases towards the apparent critical regime, we observe nearly equivalent dilute and dense phase volume fractions (Fig. 5A), which could not be observed in simulations of the 200-chain system. To assess the accuracy of the computed binodals, we compared them to one another and to experiments. Note that the experiments used absorbance for measuring off-critical coexistence points, and cloud point measurements to map the critical region (60). The cloud point measurements suggest that the critical region is well above $T_{c,\text{Ising}} = 307.40$ K, as estimated from the 200-chain simulations. However, 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. 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} (90). 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. 5B). These comparisons give us confidence regarding the accuracy of the model used in the LaSSI simulations.

Fitting the entire binodal for the 10^4-chain system to the scaling of the order parameter that is based on the 3D Ising model yields a higher estimate of T_c, when compared to the 200-chain system. However, the reproduction of the binodal is very poor (Fig. 6) and this is expected because the scaling is invalid away from the critical regime. Indeed, when we restrict the analysis to the near-critical regime, i.e., to temperatures above $T_{\text{crossover},\Delta}$ (see Fig. 7B), the estimated T_c does not change much. 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. 6B). 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 goes against expectations of behaviors in the vicinity of the critical point (82, 88).

In summary, the preceding analysis establishes why 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. For this exercise, we recommend the use of fits to the modified Flory-Huggins model (17), as opposed to the use of regression based on assuming scaling of the order parameter based on the 3D Ising model (54, 55). 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.

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 as $T \to T_c$ (44, 91). The derivative of $\phi(r)$ with respect to r (Fig. 7A) helps delineate the location of the interface. We inserted the computed values of $\Delta(T)$ and $r_{\text{mid}}(T)$ into a hyperbolic tangent, Fisk-Widom function (56, 71). This function is then fit to the T-dependent profiles for $\phi(r)$ to

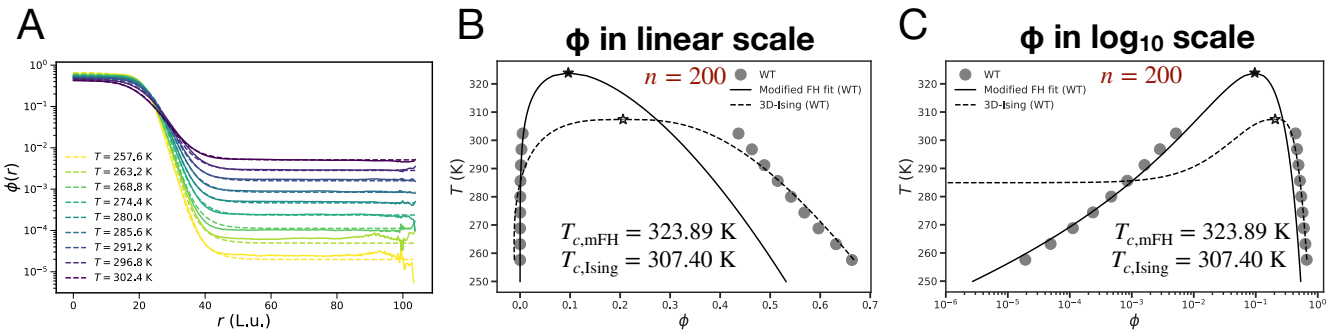

Fig. 3. Finite-sized systems pose limitations in mapping the full binodal. The results are shown here for A1-LCD. (A) T-dependent radial density profiles $\phi_T(r)$ for the system containing 200 molecules of WT A1-LCD. Different profiles were obtained for different temperatures T (in Kelvin). For each temperature, the computed radial density profile was used to extract the dense and dilute coexistence concentrations by fitting them to the Fisk-Widom function. (B-C) Comparative analysis of estimates of critical points and fitting binodals to the modified FH model and the 3D-Ising model. The results are shown in (B) linear scale and (C) log scale of the volume fraction ϕ. These results highlight the weaknesses of simulations based on finite-sized systems and the use of fits to analytical expressions without assessing the contributions from finite size artifacts.

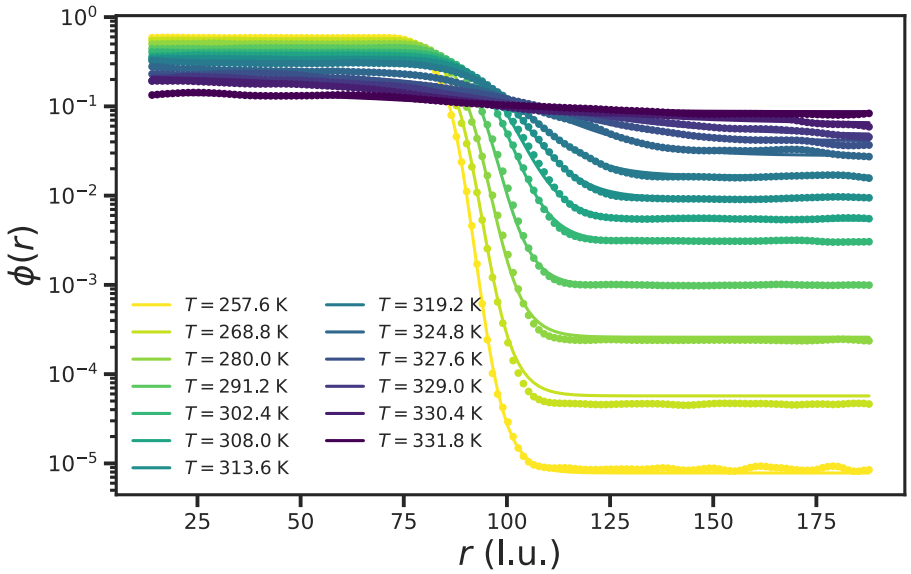

Fig. 4. T-dependent radial density profiles for a system with 10^4 molecules of A1-LCD. Radial density profiles $\phi(r)$ as a function of r (in lattice units) at different temperatures for the 10^4 chain system, which are then fit using the Fisk-Widom function to estimate the coexisting dense and dilute phase concentrations (ϕ_{dense} and ϕ_{dilute}).

estimate $\phi_{\text{dense}}(T)$ and $\phi_{\text{dilute}}(T)$ (Fig. 4, Fig. 7A).

The width $\Delta(T)$ shows the presence of two regimes as a function of temperature, with the crossover temperature being $T_{\text{crossover},\Delta} \approx 322.74$ K (Fig. 7B). 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, 91). Thus, we used $T_{\text{crossover},\Delta}$ to demarcate the critical and off-critical regimes, with $T > T_{\text{crossover},\Delta}$ designated as 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. 7C) (92). Here, $l = 4$ is a factor that converts 1 l.u. to 4 Å (see Methods section). 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 $T \to T_c$ (44). The calculated values of $\gamma(T)$ are consistent with measured interfacial tension values for condensates (4, 40, 93–95) and for colloid-polymer systems (92).

Analysis of Binder cumulants to estimate T_c. To avoid the problems outlined in the preceding sections and take advantage of the accuracy afforded by large-scale simulations, we used an approach based on Binder cumulants, which rests on computing density cumulants and the use of rigorous finite-size scaling (59, 67, 89, 96–98). The Binder 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. 8.

Probability distributions, $p(\phi)$, calculated within sub-boxes of different size L, are shown in Fig. 9(A-D) for four representative temperatures (also see Fig. S3 in the *SI Appendix*). Below T_c, and for small sub-boxes, the density distributions are bimodal Fig. 9(A-C), reflecting the coexistence of dense and dilute phases. The density distributions become unimodal

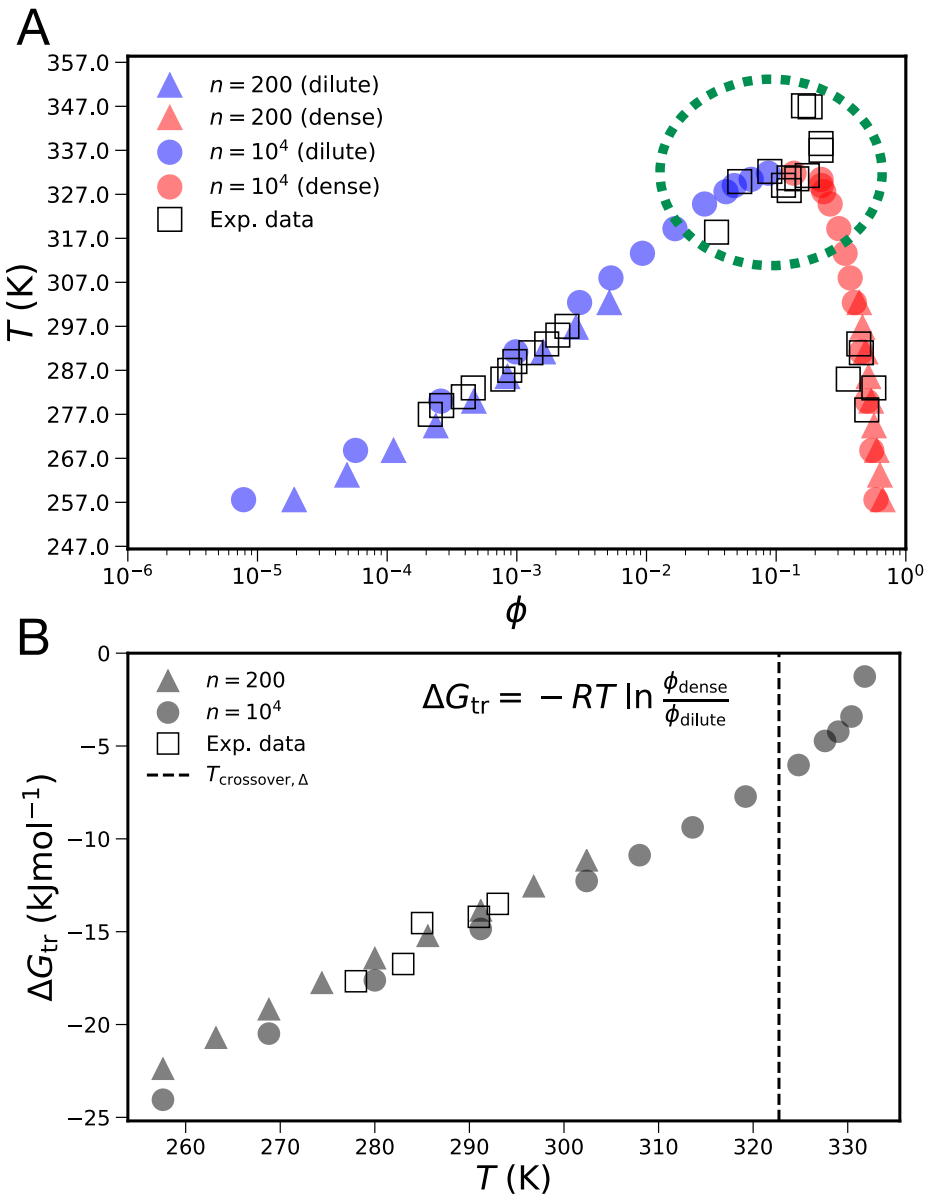

Fig. 5. Comparisons between computed and measured binodals. (A) Binodal of A1-LCD, showing coexistence concentrations of the dilute phase (blue) and dense phase (red). Results from simulations with 10^4 chains are shown as circles, and those from simulations with 200 chains are shown as triangles. The measured coexistence points (60) are shown as open squares. The critical regime is delineated by the green dotted oval. In this regime, the experiments reveal large density fluctuations. 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. Conversely, simulations based on the 200-chain system cannot access the critical regime. This emphasizes the erroneous nature of the estimates of T_c and ϕ_c that we obtain from binodals for the 200-chain system. (B) To assess the accuracy of the binodal, we computed transfer free energies, calculated as $\Delta G_{\text{tr}} = -RT\ln(\phi_{\text{dense}}/\phi_{\text{dilute}})$. These are plotted as a function of temperature. Here, $R = 8.314\,\text{Jmol}^{-1}\text{K}^{-1}$. 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 $T_{\text{crossover},\Delta} \approx 322.74$ K between the mean field and critical regimes (see text and Fig. 7B).

and Gaussian for all sub-box sizes L as T_c is approached and crossed (Fig. 9D).

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 (Fig. 9E, Fig. 9F). Note that this analysis was performed for cumulants computed in sub-boxes of size $L \geq 160$ l.u. This choice ensures that all density fluctuations are fully accounted for in estimating T_c. The mean and median values of T_c were found to be 332.14 K and 332.42 K, respectively (Fig. 9E, Fig. 9F). At the median value, we find that $U_L(T) \approx 0.466$. This is congruent with the value of $U^* = 0.466$ that is expected at $T = T_c$ for the 3D Ising model (99). The statistical error in the estimation of T_c is obtained from the lower and upper bounds of the set of all pairwise crossings of the Binder cumulant curves (shown in the inset of Fig. 9F). The values of these lower and upper bounds are 331.35 K and 332.63 K respectively.

From the point at which the horizontal line that intersects the binodal and the ordinate at T_c one can drop a vertical

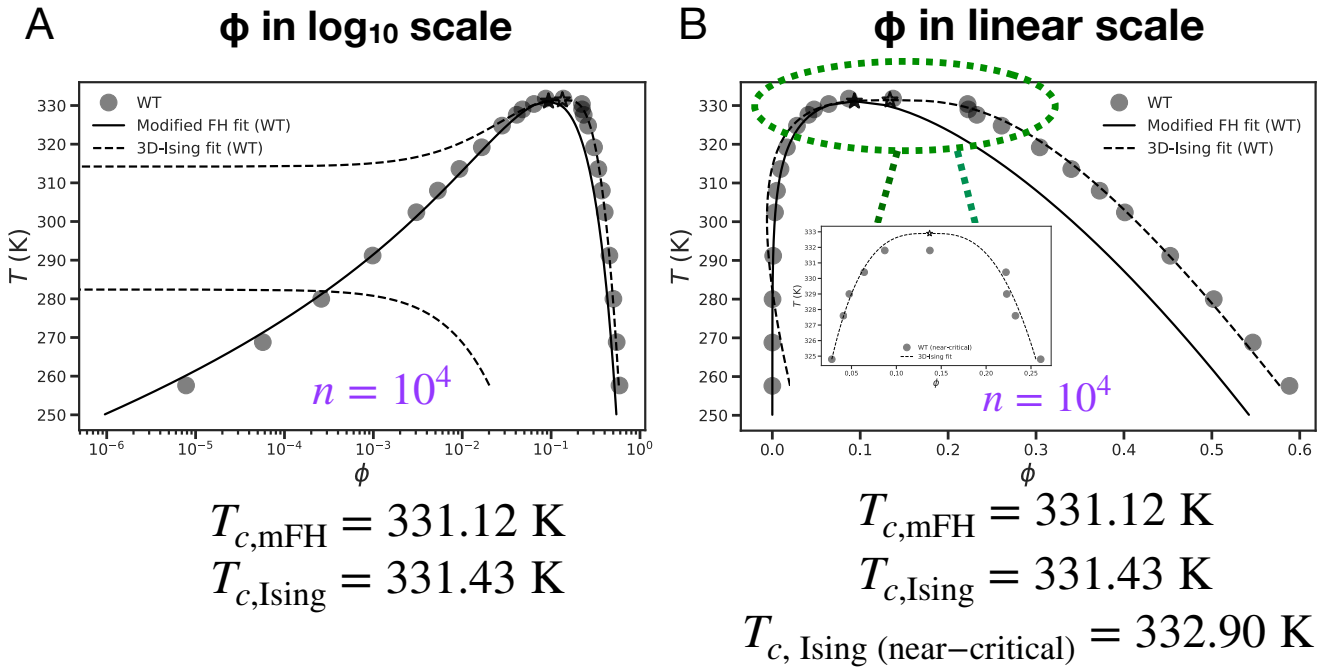

Fig. 6. Illustration of the error of using a single model to fit the entire binodal. Comparative analysis of estimates of critical points based on two different approaches. Fits performed using the modified FH model and the 3D-Ising model to the entirety of the binodal for a system comprising 10^4 molecules of WT A1-LCD. The results are shown in both (A) log scale and (B) linear scale of the volume fraction ϕ. The inset in (B) shows results obtained by fitting the 3D-Ising model to points in the critical regime, i.e., $T > T_{\text{crossover},\Delta}$ (see Fig. 7B). Even though the analysis has access to concentrations in the critical regime, the fit to the dilute arm using the 3D-Ising model approach is poor.

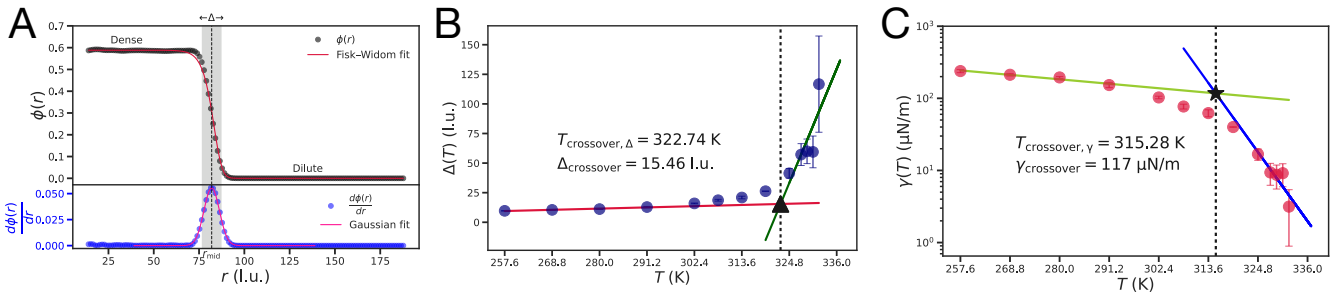

Fig. 7. Demarcation of the off-critical (mean field) and near-critical regimes. (A) For a given temperature, the derivative of the radial density profile $\phi(r)$ is fitted to a Gaussian function, from which the interfacial width Δ is extracted as the full-width-at-half-maximum (FWHM). The midpoint 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 (56, 71). 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) Temperature dependence of $\Delta(T)$ showing a crossover from the off-critical to the near-critical regime at $T_{\text{crossover},\Delta} \approx 322.74$ K, with a corresponding interface width $\Delta_{\text{crossover}} \approx 15.5$ lattice units. (C) A similar crossover is observed for the interfacial energy per unit area $\gamma(T)$, with $T_{\text{crossover},\gamma} \approx 315.28$ K and $\gamma_{\text{crossover}} \approx 117\,\mu\text{N/m}$.

line and geometrically estimate ϕ_c as the value where this line intersects the abscissa. This approach yields a value of $\phi_c \approx 0.099$ (Fig. 10A). Using this method and the bounds in our estimate of T_c, we obtain values of ≈ 0.094 and ≈ 0.122 as the lower and upper bounds on the estimate of ϕ_c, respectively. As an independent approach to estimate ϕ_c, we used the law of rectilinear diameters (Fig. 10B). Usage of this law is valid because the interactions are short-range and the system is in 3-dimensions (100, 101). The temperature-dependent average of the dense and dilute phase volume fraction $\phi_{\text{diam}}(T)$, computed at different temperatures in the near-critical regime, is fit using linear regression and the knowledge of T_c. A limitation of the law of rectilinear diameters is its sensitivity to the number of data points used in the regression near T_c. To quantify the uncertainty in the estimate obtained for ϕ_c, we used bootstrap sampling with replacement of the points in the near-critical regime. For each of the 1000 resampled datasets, we fixed $T_c = 332.42$ K and fit $\phi_{\text{diam}}(T) = \frac{1}{2}\left[\phi_{\text{dense}}(T) + \phi_{\text{dilute}}(T)\right] = \phi_c + m\,(T - T_c)$, and obtained ϕ_c as the intercept. The bootstrap distribution yields mean and median estimates of $\phi_c \approx 0.124$ and ≈ 0.122, respectively, with a standard deviation of ≈ 0.012 (Fig. 10B). The reported slope, $m \approx -0.003$ K^{-1}, corresponds to the average value obtained across all the bootstrap samples.

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 (58) becomes problematic because of the unavoidable sparseness of the number of points in the critical regime. Additionally, the order pa-

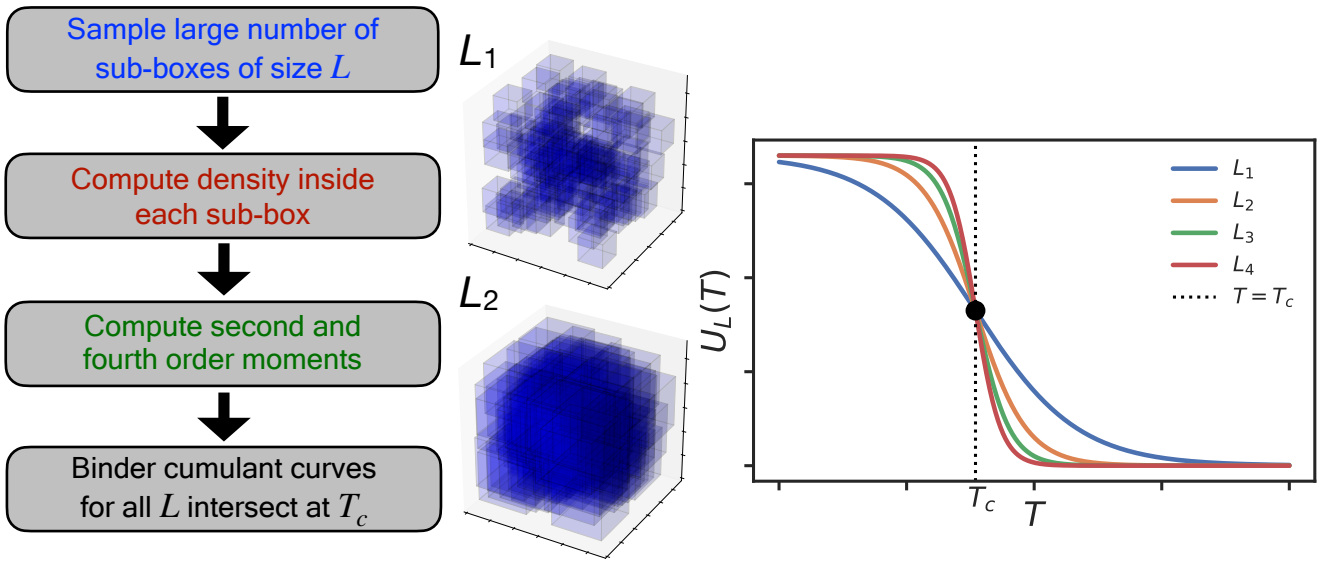

Fig. 8. Workflow for Binder Cumulant Analysis. Schematic illustrating the procedure used to obtain an accurate and independent estimate of the critical temperature T_c. This is based on computations of higher order moments of the order parameter $m_L = (\phi_L(T) - \langle\phi_L(T)\rangle)$, which quantifies 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.

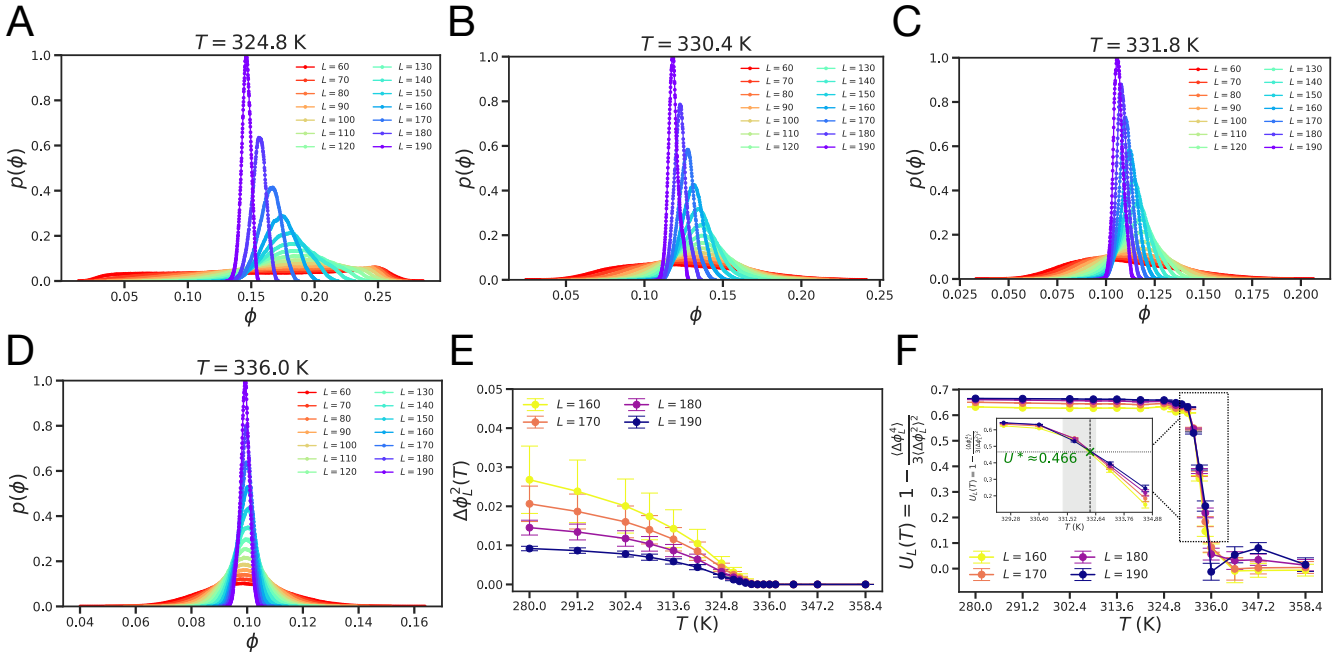

Fig. 9. Distributions of the order parameter and results from Binder cumulant analyses. (A–D) Probability distributions of the sub-box densities $p(\phi)$ at various temperatures T and for different sub-box sizes L (in lattice units). These distributions are used to calculate the second- and fourth-order 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. 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 temperature T for various sub-box sizes L. (F) Binder cumulant $U_L(T)$ as a function of T for various sub-box sizes L. The curves intersect at a median critical temperature of $T_c \approx 332.42$ K, as shown in the inset. The shaded gray region (in the inset) indicates the range between the lower and upper 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 (99).

rameter 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 $\Delta T_i = (T_c - T_i)$. This approach was inspired by the method of Meng et al., that was used to quantify scaling exponents for denatured states of proteins (78). 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. For the 3D Ising model, $\beta = 0.3264$ (102). Thus, 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, 51, 53–55, 58, 84, 85, 87).

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 (see Fig. S4 in *SI Appendix*, Fig. 10A). 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, 103), we find that the actual percolation line, computed via rigorous clustering analysis (46, 104), intersects the binodal at a temperature $T_p = 327.58\,\text{K} < T_c$.

The overlap line. Next, we quantified the overlap concentration ϕ^* as a function of T. Above the overlap concentration, the likelihood of polymers interacting via intermolecular interactions is higher than the likelihood of them interacting via intramolecular interactions (38, 82). The overlap line is the locus of ϕ^* values computed as a function of T (see Fig. S5 in *SI Appendix*, Fig. 10A). 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, 82, 88) with the chain concentration being above the overlap value. For WT A1-LCD, $T^* \approx 300.05\,\text{K}$.

The binodal has three regimes. The binodal of A1-LCD (Fig. 10A) 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. 10A).

The dense phase is a percolated network. Across the three regimes of the binodal, the concentration of the dense phase changes by less than a factor of four. 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 $T < T_c$ clearly indicates that the phase transition of A1-LCD is consistent with PSCP (56, 60).

Clustering in dilute phases is different across the regimes. In contrast to the dense phase, the concentration of coexisting dilute phase decreases by more than four orders of magnitude between T_c and the lowest temperature for which phase behavior was investigated. 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 the temperature T.

The probability $p(n)$ of forming clusters containing n molecules can be described using a power-law relation of the form: $p(n) = n^{-\tau}$ (105). 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 in the coexisting dilute phases as a function of temperature. Using linear regression analysis, which required the fitting of $\log_{10} p(n)$ to $-\tau \log_{10} n$, we extracted the temperature dependence of τ (Fig. 11D).

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 $3.6 \leq \tau \leq 4.1$, with the lower values corresponding to higher temperatures (Fig. 11A). The coexisting dilute phase in Regime I is akin to a "gas" of A1-LCD molecules. In Regime II, the coexisting dilute phase is a semidilute solution (38, 82, 88). 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 $\tau \approx 2.18$ for $T \approx T_p$ (Fig. 11B, Fig. 11D). 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. 11D). 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. These flat tails imply that the probabilities of realizing clusters of different sizes become equivalent for clusters comprising ten to hundred molecules. Similar observations were made by Lan et al., in live cells (49). 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 $\approx 100\,\mu\text{N/m}$ at $T \approx T^*$ to less than $10\,\mu\text{N/m}$ as T_p is approached.

In Regimes I and II, the dense phases can be described as "confined physical gels" (29). This is because $\phi_{\text{dense}} > \phi_{\text{perc}}$, and the single largest cluster localizes to the dense phase (Fig. 11A, Fig. 11B) across the range of temperatures corresponding to the two regimes. Note that the interfacial tension γ is highest in Regime I and decreases to $\approx 10\mu\text{N/m}$ at the highest temperature within Regime II. Thus, capillary fluctuations are expected to increase as the interfacial tension decreases (92). This in turn will weaken the confinement of the percolated network.

Percolation is a continuous networking transition that is different from density transitions (9, 26, 83, 104, 106). 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 (104). $\phi_p = \phi_{\text{perc}}$ is the value of ϕ where the percolation line intersects the binodal, and hence

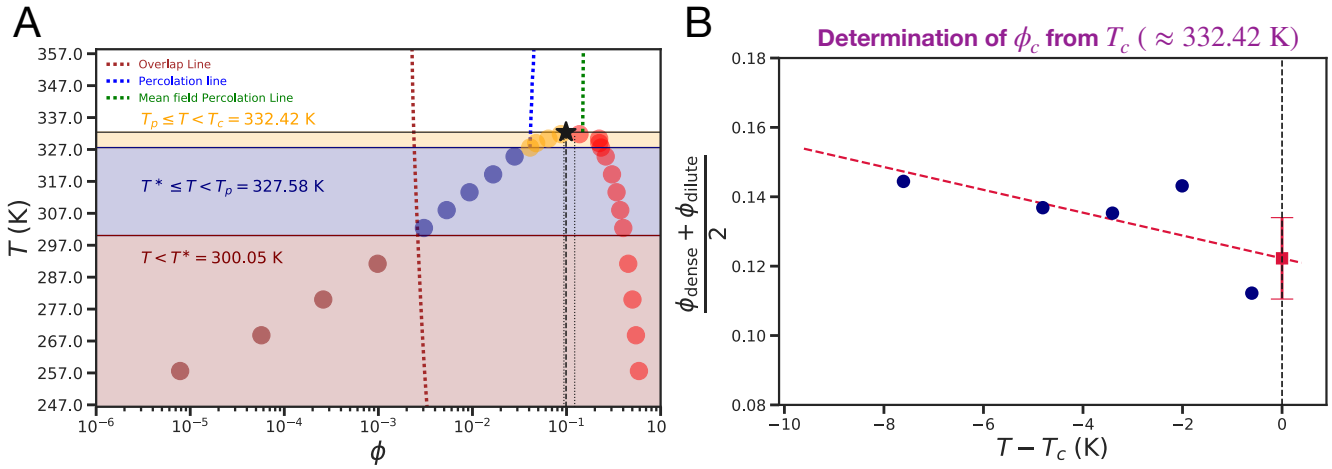

Fig. 10. Fully annotated binodal for WT A1-LCD and determination of ϕ_c. (A) Fully annotated binodal for WT A1-LCD. The overlap and the percolation lines are indicated by maroon and blue dashed lines, respectively. The mean field percolation line, mapped using the generalized bond percolation model (21) (see *SI Appendix*), is shown as a green dashed line. The overlap and the percolation lines intersect the left arm of the binodal and divide it into three distinct regimes. The critical point ($\phi_c \approx 0.099, T_c \approx 332.42$ K) is marked by an asterisk. Here, ϕ_c is estimated by a vertical linear extrapolation and identified as the value where the vertical drop from T_c intersects the abscissa (shown by the dashed black line). We also show the upper and lower bounds in the estimation of ϕ_c using two dotted black lines that correspond to vertical extrapolations done from the upper and lower bounds in the estimation of T_c (shown in Fig. 9F as the bounds of the shaded gray region). Red circles denote 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), for $T < T^*$; (ii) the intermediate regime (navy blue), for $T^* \le T < T_p$; (iii) the high-temperature regime closest to criticality (orange), for $T_p \le T < T_c$. (B) An independent estimate of the critical volume fraction ϕ_c was obtained using the law of rectilinear diameters, with $T_c = 332.42$ K fixed from Binder cumulant analysis. Bootstrap analysis with 1000 resampled datasets yields mean and median estimates of ϕ_c equal to 0.124 and 0.122, respectively, with a standard deviation of 0.012.

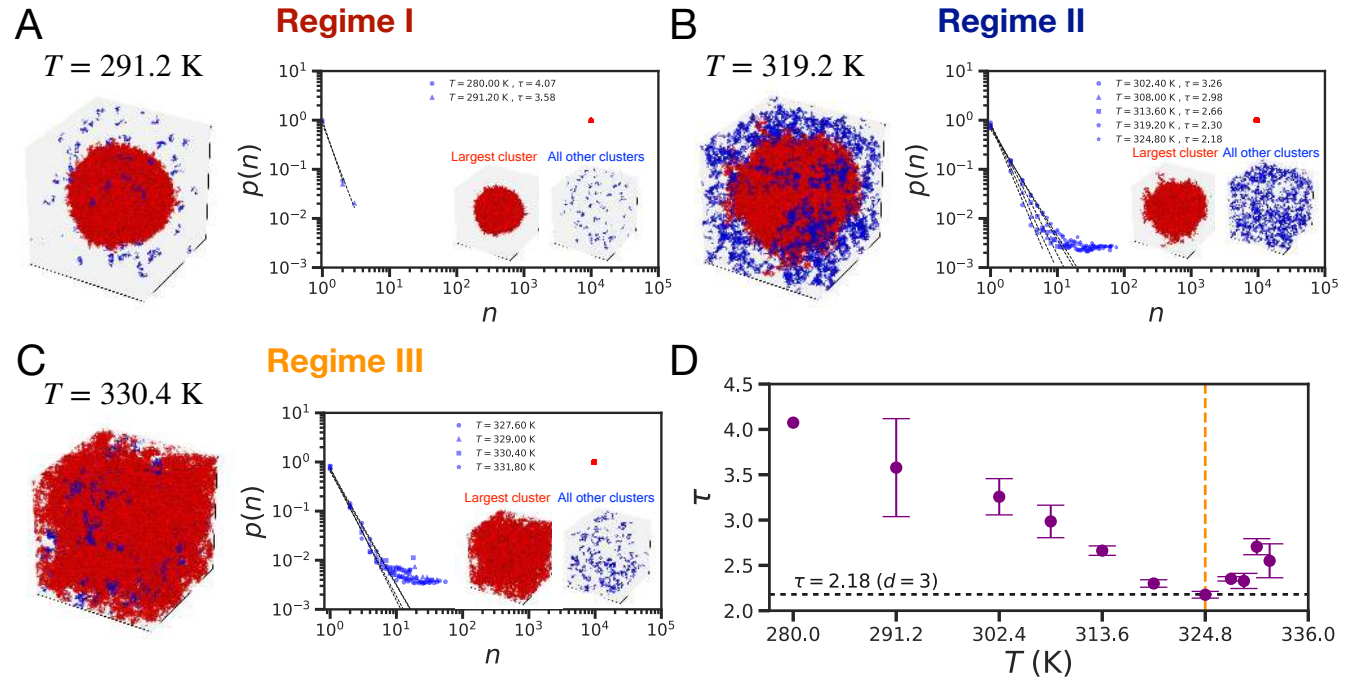

Fig. 11. Cluster-size distributions in coexisting dilute phases for each of the three distinct regimes of the binodal. (A-C) Snapshots showing the largest connected cluster (red) and all other clusters (blue), and the cluster-size distributions for the dilute phase (shown by blue symbols) and dense phase (shown by red symbols). The legends show the dilute phase τ values obtained at different temperatures T, for Regimes I and II. (A) Regime I of the binodal defined by $T < T^*$. The dense phase that is representative of the largest connected cluster is colored in red, while the dilute phase is shown in blue. (B) Regime II of the binodal defined by $T^* \le T < T_p$. Here also, the dense phase that is representative of the largest connected cluster is colored in red, while the dilute phase is shown in blue. (C) Regime III of the binodal defined by $T_p \le T < T_c$. Here, the snapshot shows the largest connected cluster (red) and smaller clusters (blue) that form. (D) Temperature dependence of the Fisher exponent τ for dilute phase cluster-size distributions. The exponent τ extracted from the power law fits of the dilute phase cluster-size distributions is plotted as a function of the temperature T. The orange vertical line marks $T = 324.80$ K, 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.

(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. 11D). 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 ($T_p \leq T < T_c$) 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. 11C). 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 percolated networks corresponding to the dense and dilute phases swell and become interconnected (Fig. 11C).

Estimating the theta temperature. The critical point (ϕ_c, T_c) marks the end of the two-phase regime, implying that the system enters the one-phase regime for $T > T_c$. Thus, for $T < T_c$, the two-body interaction coefficient (B_2) is expected to be negative. B_2 being negative would imply that the effective interactions between pairs of chains or pairs of residues within a chain are attractive for $T < T_c$. The theta temperature (T_θ) is the temperature at which $B_2 = 0$, and therefore for a system such as A1-LCD, which has an upper critical solution temperature (UCST), it follows that T_θ should be greater than T_c (82, 88). Accordingly, we estimated T_θ and asked if it follows canonical expectations.

In Flory-style two-parameter theories (82), T_θ is the temperature where $B_2 = 0$ for a pair of residues in a chain, and this is taken to be congruent with the temperature at which an individual polymer in a dilute solution behaves like a Gaussian chain (51, 69, 107). To identify the lowest temperature at which Gaussian chain statistics are observed, we performed simulations of individual A1-LCD molecules as a function of the temperature T (Fig. 12A). To extract the values of $\nu(T)$, we analyzed the scaling of ensemble-averaged segmental distances $R(s)$ plotted against the segment length s (Fig. 12B). These plots are referred to as internal scaling profiles (108). For a given profile, $R(s)$ quantifies the ensemble-averaged spatial separation $R(s)$ for all linear segments of length s. If the scaling of $R(s)$ against s shows power law behavior, then the implication is that $R(s) \sim s^\nu$. A value of $\nu = 0.5$ defines a chain that follows Gaussan statistics.

From the analysis of the internal scaling profiles extracted at different temperatures, we can identify the lowest temperature at which $\nu = 0.5$ (Fig. 12C). In Flory-style mean field theories, this would be taken as the inferred (apparent) theta temperature $T_{\theta,\mathrm{app}}$. We find that the lowest temperature ($T_{\theta,\mathrm{app}}$) at which $\nu = 0.5$ is 269.22 K. Other approaches (109, 110) introduced in the literature yield similarly low estimates of T_θ. These estimates of T_θ are incorrect because they are well below the T_c value of 332.42 K. Our findings negate the applicability of the assumptions of Flory-style two parameter theories, which mandate that $\nu = 0.5$ at temperatures where $B_2 = 0$ for a pair of chain molecules.

Given the errors of, and weaknesses inherent to the scaling analysis, we computed T_θ directly by setting it to be the lowest temperature at which $B_2 = 0$ for a pair of A1-LCD molecules. For this, we computed potentials of mean force (PMFs) $W(r)$ for a pair of A1-LCD molecules at different temperatures (Fig. 13A). Each PMF quantifies the work done at a given temperature to bring a pair of chains from infinite separation to a distance r between their centers-of-mass. The Mayer f-function can be written as $f(r) = \exp(-\beta W(r)) - 1$ and the integral of $-f(r)$ yields the excluded volume, which is used to compute B_2 (see Methods section). The normalized B_2 crosses zero at a temperature of 361.73 K. This value is greater than $T_c = 332.42$ K and appears to be a more reasonable estimate of T_θ (Fig. 13B). That the scaling analysis underestimates T_θ and that the temperature where $B_2 = 0$ will be higher than the $T_{\theta,\mathrm{app}}$ derived from scaling analysis was established by Kholodenko and Freed via conformational space renormalization group calculations (111). Our results are in accord with their calculations, and have implications for drawing inferences regarding solvent quality for disordered proteins, as discussed below.

Discussion

Mapping of critical points is essential for a complete physical description of the phase behaviors of intrinsically disordered proteins such as PLCDs. Knowledge of T_c provides a direct route to obtaining comparative assessments of sequence-specific driving forces for phase transitions (54, 55). In this work, we demonstrate that methods used in the computational literature do a poor job of mapping the critical regime. The main challenge comes from finite size artifacts. Furthermore, incorrect delineation of the mean field versus critical regimes leads to imposition of a single form for the scaling of the order parameter across the entirety of the binodal. Even the analytical form used for the scaling of the order parameter assumes a symmetry that is not present in computed or measured binodals, with there being larger variations in the dilute arm as opposed to the dense arm. We show that these challenges can be overcome by using a systematic and rigorous approach: First, we deploy simulation cells that are large enough, encompassing the appropriate numbers of molecules that allow for collecting statistics regarding densities and fluctuations near the critical point. Second, we delineate the off-critical (mean field) and near-critical regimes by mapping the temperature dependent changes to interfacial widths, from which we derive the interfacial tension. Third, we compute the critical temperature using rigorous finite size scaling approaches based on the analysis of Binder cumulants. This automatically yields an estimate of ϕ_c, which we verify independently by applying the law of rectilinear diameters. Finally, with the full binodal in hand, we asked where the overlap and percolation lines intersect the binodal. We find that these lines intersect the left arm of the binodal thereby setting up three distinct regimes for the coexisting dilute phases (36).

Beyond establishing a protocol for accurate mapping of critical points, we uncovered distinct features regarding the cluster-size distributions in coexisting dilute phases that are characteristic to each of the three regimes of the binodal (Fig. 10A, Fig. 11). In a recent study, Varma et al. (112) proposed a

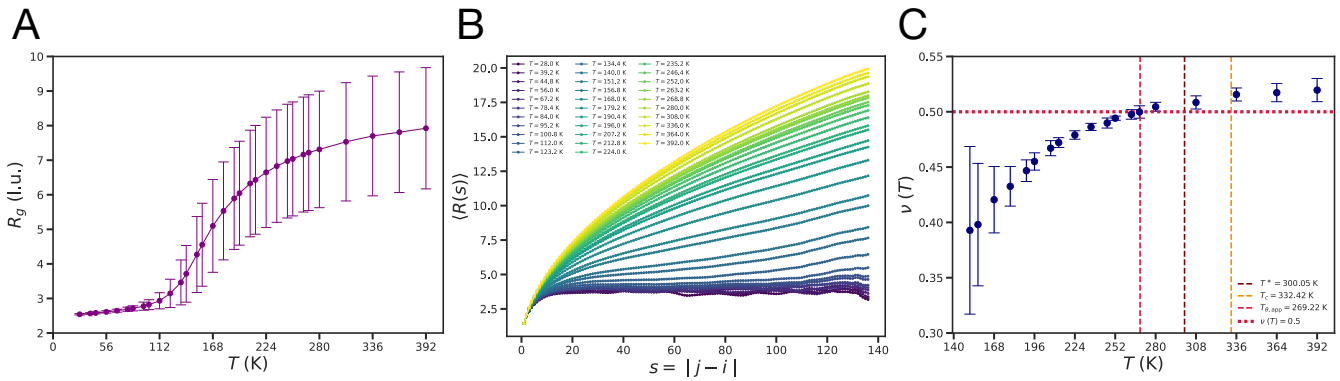

Fig. 12. Conventional scaling analysis severely underestimates the theta temperature (T_θ) for A1-LCD. (A) Radius of gyration (R_g) values from single chain simulations of A1-LCD as a function of the temperature T. These were computed from the last 25% of the snapshots from the equilibrated simulations and across three replicates for each temperature. Error bars denote the standard deviation in the computed R_g for each temperature. (B) $R(s)$ versus s for single A1-LCD chains at various temperatures T. (C) ν values obtained from analysis of the scaling of $R(s)$ versus s as a function of T. The ν values were extracted as the slopes of plots of $\ln(R(s))$ against $\ln(s)$. This analysis was restricted to temperatures where fractal behavior was evident and not for profiles where $R(s)$ is essentially flat because this corresponds to the globule regime. The inferred (apparent) theta temperature $T_{\theta,\mathrm{app}}$, corresponding to where $\nu(T) = 0.5$, is approximately 269.22 K (annotated by the red vertical dashed line). This is well below the critical temperature $T_c \approx 332.42$ K. The overlap temperature T^* and the critical temperature T_c are annotated by maroon and orange vertical dashed lines, respectively. The scaling analysis based on the segmental distances $R(s)$ as a function of the segment length s results in a severely erroneous estimate of the theta temperature for A1-LCD.

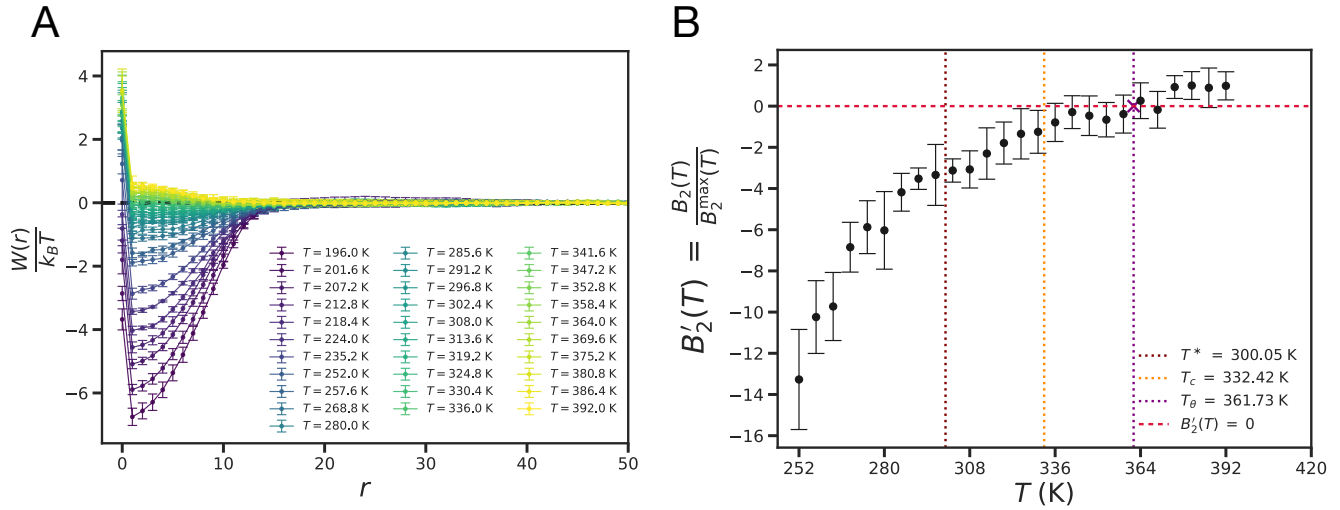

Fig. 13. Estimation of the theta temperature (T_θ) for A1-LCD. (A) The potential of mean force $W(r)$ as a function of the separation between the centers-of-mass of two A1-LCD chains, for various temperatures T, extracted from umbrella sampling simulations (see Methods section and Fig. S6 in *SI Appendix*) using the Weighted Histogram Analysis Method (WHAM) (79–81). For each temperature, umbrella sampling simulations were performed across 6 replicates. (B) The normalized two-body interaction coefficient $B_2'(T)$ as a function of T, obtained from the 2-chain PMF values. The temperature at which $B_2'(T) = 0$ is the theta temperature (T_θ). From the temperature dependence of $B_2'(T)$, we obtain the theta temperature T_θ using linear interpolation to be ≈ 361.73 K. The overlap temperature T^*, the critical temperature T_c and the theta temperature T_θ are annotated by maroon, orange and purple dotted lines, respectively.

connection between the partition coefficient (K_P) and a normalized temperature distance (t) from the critical point. We propose that a more direct assessment of the vicinity to the critical point comes from measurements of cluster-size distributions in coexisting dilute phases. This is feasible in cells (48, 49, 113–115) and in vitro (116–123). When combined with titrations above and below the overlap concentration, cluster-size distributions can be used as diagnostics of proximity to the critical point. For the measurements to be effective, they will need to be able to access the entirety of the cluster-size distribution. While this is challenging for methods that are based on scattering or fluorescence (117, 123), it should be possible using methods such as microfluidic resistive pulse sensing (MRPS) that provide a detection range from 70 nm - 15 μm (116, 121, 122). MRPS has been used to quantify the cluster-size distributions in coexisting dilute phases of different systems (116, 121, 124)

Temperatures are useful proxies for the normalized inter-chain interaction coefficients (B_2) because $T \propto \frac{1}{|B_2|}$ (6, 38, 55). This is relevant for understanding the effects of changes to B_2 on clustering in dilute phases. 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 $|B_2| < |B_2^*|$ $(T > T^*)$ 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 increasing the magnitude of B_2 (making T lower), and the coexisting dilute phase is akin to what we observe for Regime II.

Our calculations show that the percolation line intersects the binodal to the left of the critical point. Thus, the dense phase is a confined physical gel defined by an intra-condensate percolated network (9, 24, 25). For A1-LCD and related systems, the networks have been shown to have small-world like topology with smaller-scale density inhomogeneities (56). Quick-freeze deep-etch cryo-electron microscopy of a variant of A1-LCD has provided direct evidence of percolated network-like organization within condensate interiors (116). Additionally, single-fluorogen imaging of A1-LCD condensates (125) has revealed the inhomogeneous organization of molecules within A1-LCD condensate interiors (125).These properties give condensates their tunable viscoelasticity, leading to their designation as viscoelastic network fluids (76, 126–130).

In Regime III, the dense phase becomes unconfined and this network swelling has been inferred via measurements of volumes and volume fractions by Smokers and Spruijt (45). 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 (131). The observed behavior was referred to as spinodal decomposition. Indeed, Regime III is expected to coincide with the spinodal. Thus, the dynamics of phase separation in Regime III will likely follow that of viscoelastic phase separation (132). System-spanning networks that are consistent with behaviors in Regime III have also been observed in signaling systems that form at membranes (133).

Availability of the full binodal allowed us to assess the accuracy of using internal scaling profiles to identify the theta temperature. We find that this yields an erroneously low estimate for T_θ and the false result of T_c being higher than T_θ. Direct computation of B_2 provides a more reliable estimate of T_θ that is in accord with canonical expectations (82). The central tenet of Flory-style two parameter theories is that $\nu = 0.5$ for a single chain when $B_2 = 0$ for pairs of residues (beads) (82). It is known that this is not formally valid for $d = 3$, where d is the dimensionality of the system (134). The Flory model comes about by assuming an equivalence between an individual chemical or Kuhn monomer - residue (bead) in our case - and the entire chain. However, the effective two-body interaction coefficient for pairs of chains will be influenced by correlations that stem from chain connectivity, chain flexibility, and three-body as well as higher-order interactions (110). The impact of these correlations leads to a decoupling of the impact of solvent quality on chain size versus shape, which will be pronounced for finite-sized heteropolymers *viz.*, IDPs (109, 135).

Our results have broad implications for studies of IDPs. It is customary to use either simulations or measurements to derive internal scaling profiles and quantify the apparent value of ν (17, 69, 108, 136–140). The derived ν value is then used to judge the quality of the solvent for the IDP of interest. Our results suggest that these inferences of solvent quality must be treated with caution. This is because it is quite likely that the temperature or equivalent control parameter at which $\nu = 0.5$ will lie below T_c or the critical value of control parameter. Therefore, any rigorous assessment of solvent quality will require direct computation or measurement of the two-body interaction coefficient B_2, or alternatively, computation or measurement of phase boundaries (60, 141) as a function of temperature or changes to solvent via chemical perturbations such as changes in pH, denaturant, or salt concentration.

Finally, we note that the findings reported here should also be accessible via off-lattice MD simulations. For each temperature, the simulations will need to be on the order of milliseconds, and the simulation cells will need to encompass $O(10^4)$ molecules. A recent study (77) showed that structures within simulated dense phases of A1-LCD molecules share similarities across three different simulation engines that include LaSSI, MPIPI-recharged (142), and CALVADOS (90, 143). 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 (59, 89). It might also be essential to revisit the use of slab geometries that are in vogue across many MD simulation protocols (20, 51, 53–55, 84, 90, 144). Slabs cause a smoothening of density profiles and this is achieved by suppressing fluctuations that are the defining hallmarks of critical regions. These geometries are helpful deep inside the mean field regime, but are likely to be problematic in the critical regime. 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.

Acknowledgements

We are grateful to Tanja Mittag for collaborations, and to Michael Rubinstein for pedagogy and insights.

Funding

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), the St. Jude Children's Research Hospital Collaborative on the BIology and Biophysics of RNP granules, and the Center for Biomolecular Condensates in the James McKelvey School of Engineering at Washington University in St. Louis.

Author contributions

Conceptualization: GM, SG, GC, RVP; Methodology: GM, GC, SG, RZ, RVP; Formal Analysis: GM, GC, SG, KMR, RVP; Investigation: GM, SG, RVP; Software: GM, GC, RZ, KMR; Project Administration: RVP; Data Curation: GM, SG, KMR; Visualization: GM, SG, KMR, RVP; Writing - Original Draft: RVP; Writing - Editing and Revising: GM, SG, KMR, RVP; Funding Acquisition: RVP.

Bibliography

1. Salman F Banani, Hyun O Lee, Anthony A Hyman, and Michael K Rosen. Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular and Cell Biology*, 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. *Chemical Reviews*, 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*, 6:30294, 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):042403, 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. 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):5237, June 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-60345-9.
25. 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:e81907, June 2023. ISSN 2050-084X. doi: 10.7554/elife.81907.
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:e67176, 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. *Nature Chemistry*, 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. *Nature Chemistry*, 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 Computational Biology*, 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. 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.
52. Gregory L. Dignon, Robert B. Best, and Jeetain Mittal. Biomolecular phase separation: From molecular driving forces to macroscopic properties. *Annual Review of Physical Chemistry*, 71(1):53–75, April 2020. ISSN 1545-1593. doi: 10.1146/annurev-physchem-071819-113553.
53. 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.
54. 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. 13:RP99068, January 2025. doi: 10.7554/elife.99068.2.
55. 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.
56. 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. *Nature Communications*, 13 (1):7722, December 2022. doi: 10.1038/s41467-022-35370-7.
57. 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. *Nature Communications*, 14(1):5527, September 2023.
58. 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.
59. Kurt Binder. Finite size scaling analysis of ising model block distribution functions. *Zeitschrift für Physik B Condensed Matter*, 43(2):119–140, 1981.
60. 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.
61. 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.
62. 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.
63. 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.
64. 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.
65. 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.
66. 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.
67. Kurt Binder. Computer simulations of critical phenomena and phase behaviour of fluids. *Molecular Physics*, 108(14):1797–1815, 2010.
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. 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.
70. 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.
71. 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.
72. 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.
73. K Binder. Applications of monte carlo methods to statistical physics. *Reports on Progress in Physics*, 60(5):487, may 1997. doi: 10.1088/0034-4885/60/5/001.
74. 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.
75. Steffen Frey, Ralf P. Richter, and Dirk Gorlich. Fg-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science*, 314(5800):815–817, November 2006. ISSN 1095-9203. doi: 10.1126/science.1132516.
76. Samuel R Cohen, Priya R Banerjee, and Rohit V Pappu. Direct computations of viscoelastic moduli of biomolecular condensates. *The Journal of Chemical Physics*, 161(9):095103, September 2024.
77. 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.
78. 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.
79. Shankar Kumar, John M Rosenberg, Djamal Bouzida, Robert H Swendsen, and Peter A Kollman. Multidimensional free-energy calculations using the weighted histogram analysis method. *Journal of Computational Chemistry*, 16(11):1339–1350, 1995.
80. Benoît Roux. The calculation of the potential of mean force using computer simulations. *Computer Physics Communications*, 91(1–3):275–282, September 1995. ISSN 0010-4655. doi: 10.1016/0010-4655(95)00053-i.
81. Alan Grossfield. Wham: the weighted histogram analysis method. 2.1.0, 2014.
82. M. Rubinstein and R.H. Colby. *Polymer Physics*. OUP Oxford, 2003. ISBN 9780198520597.
83. 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):106, April 2025. ISSN 2399-3669. doi: 10.1038/s42004-025-01502-5.
84. 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.
85. 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.
86. B. Widom. Theory of phase equilibrium. *The Journal of Physical Chemistry*, 100(31): 13190–13199, January 1996. ISSN 1541-5740. doi: 10.1021/jp9536460.
87. 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.
88. 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.
89. 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.
90. 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.
91. B. Widom. Intermolecular forces and the nature of the liquid state. *Science*, 157(3787): 375–382, 1967. doi: 10.1126/science.157.3787.375.
92. 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.
93. 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.
94. 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.
95. 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):e2407423122, May 2025. ISSN 1091-6490. doi: 10.1073/pnas.2407423122.
96. Michael E. Fisher and Michael N. Barber. Scaling theory for finite-size effects in the critical region. *Physical Review Letters*, 28:1516–1519, Jun 1972. doi: 10.1103/PhysRevLett.28.1516.
97. Brézin, E. An investigation of finite size scaling. *Journale de Physique, France*, 43(1):15–22, 1982. doi: 10.1051/jphys:0198200430101500.
98. Anders Malthe-Sørenssen. *Finite Size Scaling*, page 85–99. Springer International Publishing, 2024. ISBN 9783031599002. doi: 10.1007/978-3-031-59900-2_6.
99. Alan M. Ferrenberg, Jiahao Xu, and David P. Landau. Pushing the limits of monte carlo simulations for the three-dimensional ising model. *Physical Review E*, 97:043301, Apr 2018. doi: 10.1103/PhysRevE.97.043301.

100. N. D. Mermin. Lattice gas with short-range pair interactions and a singular coexistence-curve diameter. *Physical Review Letters*, 26:957–959, Apr 1971. doi: 10.1103/PhysRevLett.26.957.
101. 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.
102. 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. *Physical Review B*, 29:4030–4040, Apr 1984. doi: 10.1103/PhysRevB.29.4030.
103. 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.
104. 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.
105. 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.
106. 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.
107. Andreas Vitalis and Rohit V Pappu. ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions. *Journal of Computational Chemistry*, 30 (5):673–699, April 2009.
108. Albert H. Mao, Nicholas Lyle, and Rohit V. Pappu. Describing sequence–ensemble relationships for intrinsically disordered proteins. *Biochemical Journal*, 449(2):307–318, December 2012. ISSN 1470-8728. doi: 10.1042/bj20121346.
109. Martin Oliver Steinhauser. A molecular dynamics study on universal properties of polymer chains in different solvent qualities. part i. a review of linear chain properties. *The Journal of Chemical Physics*, 122(9), February 2005. ISSN 1089-7690. doi: 10.1063/1.1846651.
110. 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.
111. 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.
112. Kaarthik Varma, Dana Matthias, Chase B. Shapiro, Sullivan Bailey-Darland, Takumi Matsuzawa, Charlotta Lorenz, Teagan Bate, Stephen J. Thornton, Senthilkumar Duraivel, Robert W. Style, James P. Sethna, and Eric R. Dufresne. Partition coefficients reveal changes in properties of low-contrast biomolecular condensates. *bioRxiv*, 2026. doi: 10.64898/2026.02.20.707107.
113. Won-Ki Cho, Jan-Hendrik Spille, Micca Hecht, Choongman Lee, Charles Li, Valentin Grube, and Ibrahim I. Cisse. Mediator and rna polymerase ii clusters associate in transcription-dependent condensates. *Science*, 361(6400):412–415, 2018. doi: 10.1126/science.aar4199.
114. Thomas R. Mumford, Diarmid Rae, Emily Brackhahn, Abbas Idris, David Gonzalez-Martinez, Ayush Aditya Pal, Michael C. Chung, Juan Guan, Elizabeth Rhoades, and Lukasz J. Bugaj. Simple visualization of submicroscopic protein clusters with a phase-separation-based fluorescent reporter. *Cell Systems*, 15(2):166–179.e7, February 2024. ISSN 2405-4712. doi: 10.1016/j.cels.2024.01.005.
115. Nikolay Zolotarev, Yuanting Wang, Manyu Du, Marc Bayer, Anna Grosschedl, Ibrahim Cisse, and Rudolf Grosschedl. Regularly spaced tyrosines in ebf1 mediate brg1 recruitment and formation of nuclear subdiffractive clusters. *Genes and Development*, 38(1–2):4–10, January 2024. ISSN 1549-5477. doi: 10.1101/gad.350828.123.
116. 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, Kannan-ganattu V. Prasanth, and Rohit V. Pappu. Nuclear speckle proteins form intrinsic and malat1-dependent microphases. *Cell*, 189:832–852.E24, 2026. doi: 10.1101/2025.02.26.640430.
117. Mrityunjoy Kar, Furqan Dar, Timothy J. Welsh, Laura T. Vogel, Ralf Kühnemuth, Anupa Majumdar, Georg Krainer, Titus M. Franzmann, Simon Alberti, Claus A. M. Seidel, Tuomas P. J. Knowles, Anthony A. Hyman, and Rohit V. Pappu. Phase-separating rna-binding proteins form heterogeneous distributions of clusters in subsaturated solutions. *Proceedings of the National Academy of Sciences*, 119(28):e2202222119, 2022. doi: 10.1073/pnas.2202222119.
118. Ian Seim, Ammon E. Posey, Wilton T. Snead, Benjamin M. Stormo, Daphne Klotsa, Rohit V. Pappu, and Amy S. Gladfelter. Dilute phase oligomerization can oppose phase separation and modulate material properties of a ribonucleoprotein condensate. *Proceedings of the National Academy of Sciences*, 119(13), March 2022. ISSN 1091-6490. doi: 10.1073/pnas.2120799119.
119. Mrityunjoy Kar, Laura T. Vogel, Gaurav Chauhan, Suren Felekyan, Hannes Ausserwöger, Timothy J. Welsh, Furqan Dar, Anjana R. Kamath, Tuomas P. J. Knowles, Anthony A. Hyman, Claus A. M. Seidel, and Rohit V. Pappu. Solutes unmask differences in clustering versus phase separation of fet proteins. *Nature Communications*, 15(1), May 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-48775-3.
120. Amber Yanas, Him Shweta, Michael C. Owens, Kathy Fange Liu, and Yale E. Goldman. Rna helicases ddx3x and ddx3y form nanometer-scale rna-protein clusters that support catalytic activity. *Current Biology*, 34(24):5714–5727.e6, December 2024. ISSN 0960-9822. doi: 10.1016/j.cub.2024.10.055.
121. Christian Hoffmann, Kiersten M. Ruff, Irina A. Edu, Min Kyung Shinn, Johannes V. Tromm, Matthew R. King, Avnika Pant, Hannes Ausserwöger, Jennifer R. Morgan, Tuomas P.J. Knowles, Rohit V. Pappu, and Dragomir Milovanovic. Synapsin condensation is governed by sequence-encoded molecular grammars. *Journal of Molecular Biology*, 437(8):168987, April 2025. ISSN 0022-2836. doi: 10.1016/j.jmb.2025.168987.
122. Macy L. Sprunger, Min Kyung Shinn, Sabrina K. Talir, Ken Lee, Rohit V. Pappu, and Meredith E. Jackrel. Matrin-3 forms spherical and wormlike assemblies that are modulated by rna binding and als/ftd-associated mutations. *Molecular Cell*, 85(19):3640–3660.E9, 2025. ISSN 1097-2765. doi: https://doi.org/10.1016/j.molcel.2025.08.034.
123. Yingda Ge, Tapas Paul, Margarita Gordiychuk, Nilimesh Das, Yaojun Zhang, and Sua Myong. Fus nanoclusters are a distinct state within the dilute phase. *Nature Communications*, 16 (1):9956, November 2025. ISSN 2041-1723. doi: 10.1038/s41467-025-64909-7.
124. Miriam Linsenmeier, Min Kyung Shinn, Thomas R. Mumford, Vicky Liu, Lukasz J. Bugaj, Rohit V. Pappu, and James Shorter. Nuclear-import receptors remodel the dilute phase to suppress phase transitions of rna-binding proteins with prion-like domains. *bioRxiv*, 2025. doi: 10.1101/2025.11.14.688546.
125. 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.
126. Archishman Ghosh and Huan-Xiang Zhou. Determinants for fusion speed of biomolecular droplets. *Angewandte Chemie International Edition*, 59(47):20837–20840, September 2020. ISSN 1521-3773. doi: 10.1002/anie.202006711.
127. Archishman Ghosh, Divya Kota, and Huan-Xiang Zhou. Shear relaxation governs fusion dynamics of biomolecular condensates. *Nature Communications*, 12(1):5995, October 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-26274-z.
128. 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. *Nature Physics*, 20(9):1482–1491, September 2024.
129. Furqan Dar, Samuel R Cohen, Diana M Mitrea, Aaron H Phillips, Gergely Nagy, Wellington C Leite, Christopher B Stanley, Jeong-Mo Choi, Richard W Kriwacki, and Rohit V Pappu. Biomolecular condensates form spatially inhomogeneous network fluids. *Nature Communications*, 15(1):3413, April 2024.
130. Rachel S. Fisher and Allie C. Obermeyer. Viscoelasticity of globular protein-based biomolecular condensates. *Chemical Science*, 15(47):19795–19804, 2024. ISSN 2041-6539. doi: 10.1039/d4sc03564j.
131. 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.
132. Hajime Tanaka. Viscoelastic phase separation. *Journal of Physics: Condensed Matter*, 12 (15):R207, 2000.
133. Jenny J. Y. Lin, Shalini T. Low-Nam, Katherine N. Alfieri, Darren B. McAffee, Nicole C. Fay, and Jay T. Groves. Mapping the stochastic sequence of individual ligand-receptor binding events to cellular activation: T cells act on the rare events. *Science Signaling*, 12(564): eaat8715, 2019. doi: 10.1126/scisignal.aat8715.
134. Lothar Schäfer. *Excluded Volume Effects in Polymer Solutions*, pages 437–477. Springer Berlin, 1999. ISBN 3-540-65513-1.
135. Gustavo Fuertes, Niccolò Banterle, Kiersten M. Ruff, Aritra Chowdhury, Davide Mercadante, Christine Koehler, Michael Kachala, Gemma Estrada Girona, Sigrid Milles, Ankur Mishra, Patrick R. Onck, Frauke Gräter, Santiago Esteban-Martín, Rohit V. Pappu, Dmitri I. Svergun, and Edward A. Lemke. Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in saxs vs. fret measurements. *Proceedings of the National Academy of Sciences*, 114(31), July 2017. ISSN 1091-6490. doi: 10.1073/pnas.1704692114.
136. Joshua A. Riback, Micayla A. Bowman, Adam M. Zmyslowski, Catherine R. Knoverek, John M. Jumper, James R. Hinshaw, Emily B. Kaye, Karl F. Freed, Patricia L. Clark, and Tobin R. Sosnick. Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. *Science*, 358(6360):238–241, October 2017. ISSN 1095-9203. doi: 10.1126/science.aan5774.
137. Andrea Holla, Erik W Martin, Thomas Dannenhoffer-Lafage, Kiersten M Ruff, Sebastian L B König, Mark F Nüesch, Aritra Chowdhury, John M Louis, Andrea Soranno, Daniel Nettels, Rohit V Pappu, Robert B Best, Tanja Mittag, and Benjamin Schuler. Identifying sequence effects on chain dimensions of disordered proteins by integrating experiments and simulations. *JACS Au*, 4(12):4729–4743, December 2024.
138. Giulio Tesei, Anna Ida Trolle, Nicolas Jonsson, Johannes Betz, Frederik E. Knudsen, Francesco Pesce, Kristoffer E. Johansson, and Kresten Lindorff-Larsen. Conformational ensembles of the human intrinsically disordered proteome. *Nature*, 626(8000):897–904, January 2024. ISSN 1476-4687. doi: 10.1038/s41586-023-07004-5.
139. Jeffrey M. Lotthammer, Garrett M. Ginell, Daniel Griffith, Ryan J. Emenecker, and Alex S. Holehouse. Direct prediction of intrinsically disordered protein conformational properties from sequence. *Nature Methods*, 21(3):465–476, January 2024. ISSN 1548-7105. doi: 10.1038/s41592-023-02159-5.
140. Michael C. Baxa, Xiaoxuan Lin, Cedrick D. Mukinay, Srinivas Chakravarthy, Joseph R. Sachleben, Sarah Antilla, Nina Hartrampf, Joshua A. Riback, Isabelle A. Gagnon, Bradley L. Pentelute, Patricia L. Clark, and Tobin R. Sosnick. How hydrophobicity, side chains, and salt affect the dimensions of disordered proteins. *Protein Science*, 33(5):e4986, April 2024. ISSN 1469-896X. doi: 10.1002/pro.4986.
141. Scott L. Crick, Kiersten M. Ruff, Kanchan Garai, Carl Frieden, and Rohit V. Pappu. Unmasking the roles of n- and c-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. *Proceedings of the National Academy of Sciences*, 110(50): 20075–20080, November 2013. ISSN 1091-6490. doi: 10.1073/pnas.1320626110.
142. 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.
143. 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):e5172, October 2024. ISSN 1469-896X. doi: 10.1002/pro.5172.
144. 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):e2417920122, March 2025.
ISSN 1091-6490. doi: 10.1073/pnas.2417920122.

Supporting Information (Appendix) for:

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

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

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:

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

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),

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

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

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

For WT A1-LCD, the relevant sticker types are Y (tyrosine), F (phenylalanine), and R (arginine) (17, 29, 60, 75), 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,

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

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

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

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

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

Now,

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

Since $k_B = 1$ in LaSSI (46, 56), the percolation threshold explicitly depends on temperature as

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

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

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

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

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

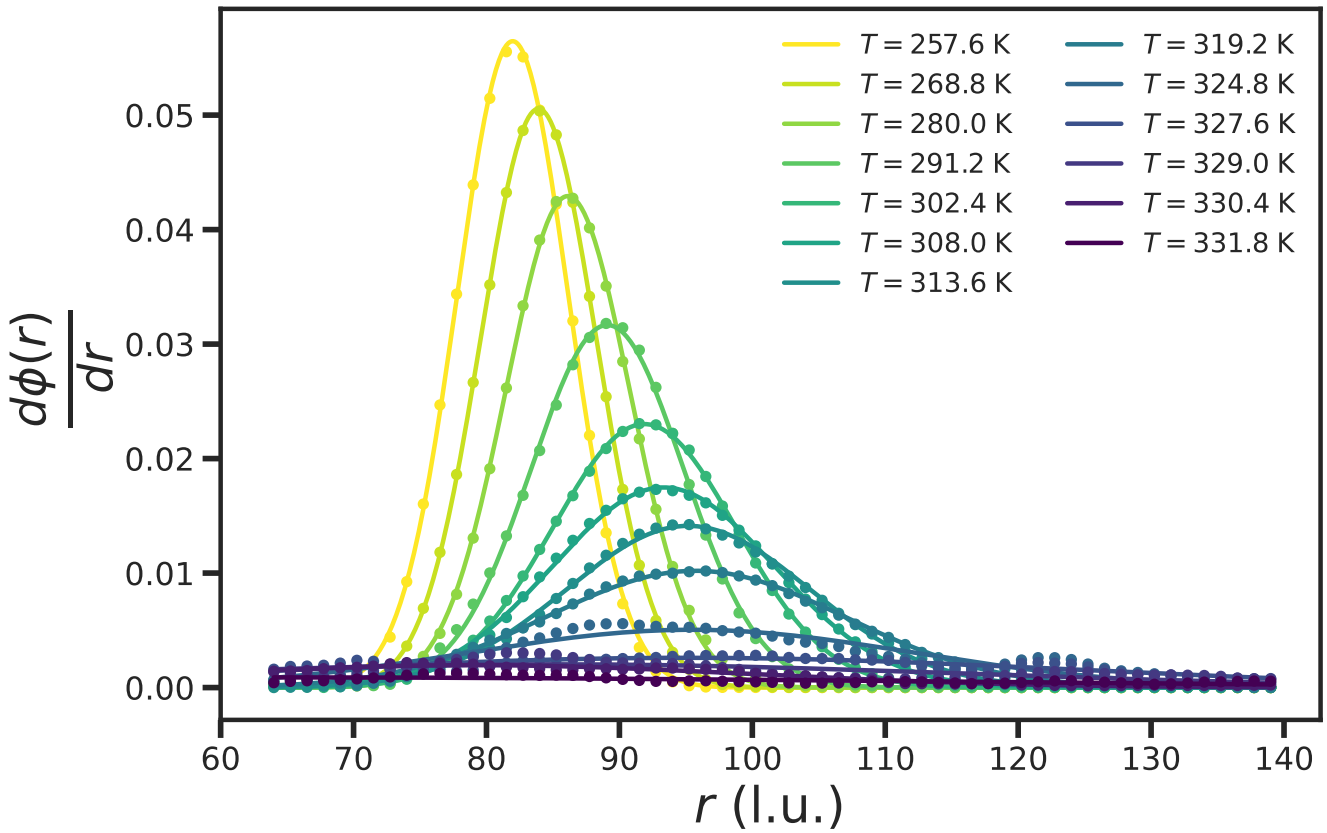

Fig. 1. Derivatives of the radial density profiles, $d\phi(r)/dr$, at various temperatures for the 10^4-chain system of WT A1-LCD. The derivative of the radial density profile at each given temperature is fit to a Gaussian function (72), from which the interfacial width $\Delta(T)$ is extracted as the 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.

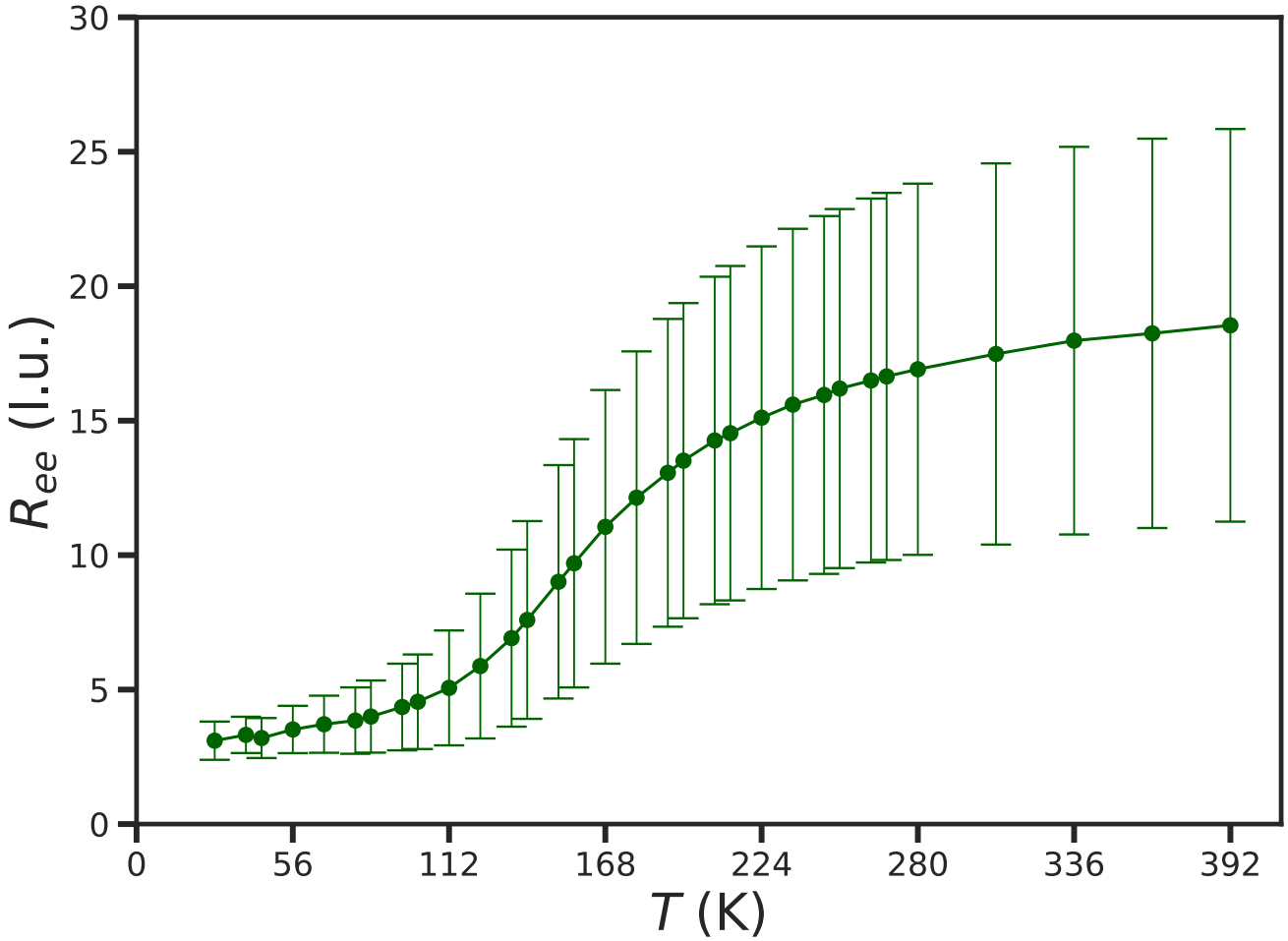

Fig. 2. End-to-end distance R_{ee} of a single A1-LCD chain ($N = 137$) as a function of the temperature T. These were computed from the last 25% of the snapshots from the equilibrated single chain simulations and across three replicates for each temperature. The mean R_{ee} values across the set of temperatures studied range from $\approx 3\,\mathrm{l.u.}$ to $\approx 17\,\mathrm{l.u.}$. Error bars denote the standard deviation in the computed R_{ee} for each temperature.

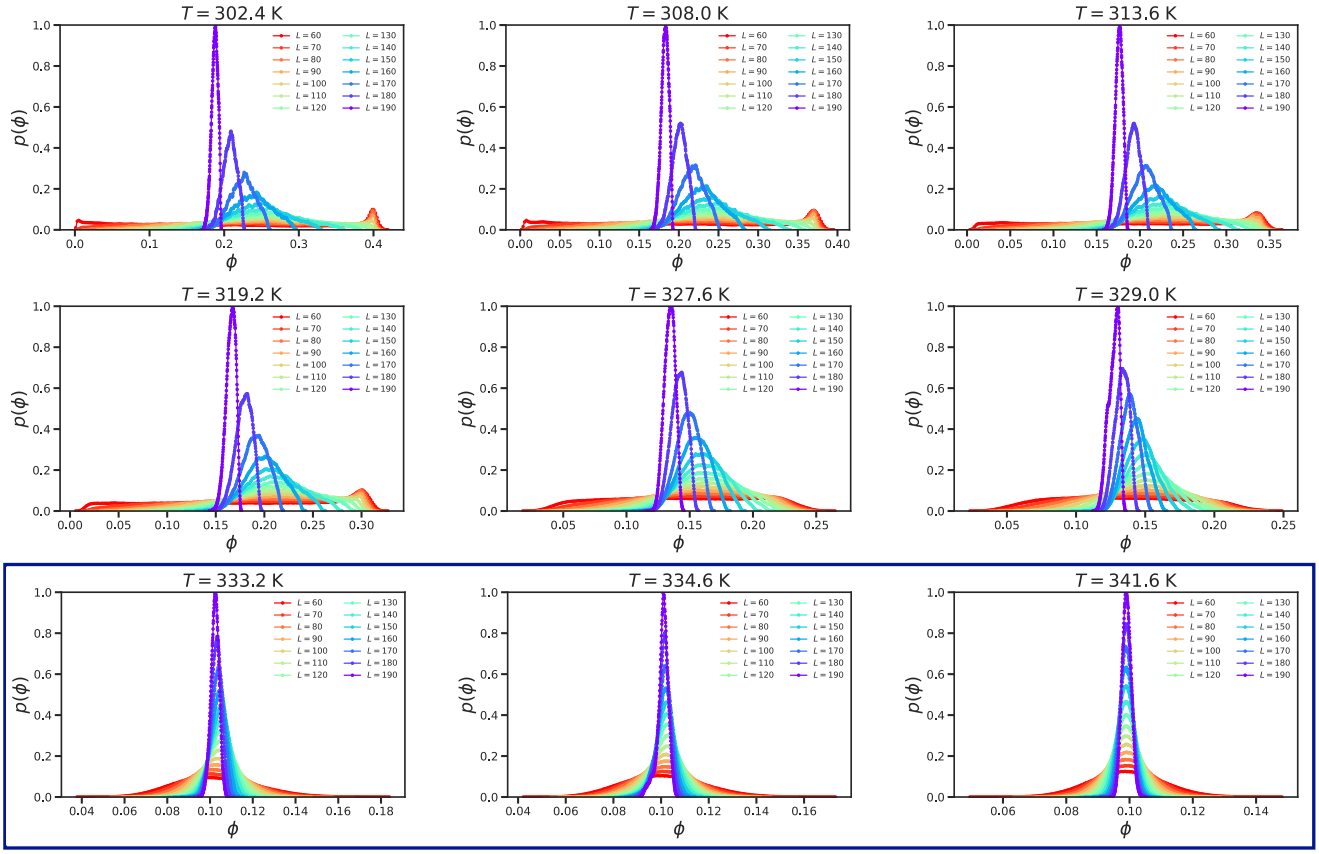

Fig. 3. Distributions $p(\phi)$ as a function of ϕ obtained by sampling sub-boxes of various sizes, at different temperatures. Note that as T approaches $T_c \approx 332.42$ K, the density distributions show Gaussian behavior for large L. In contrast, at lower values of T (for which the system is in 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.

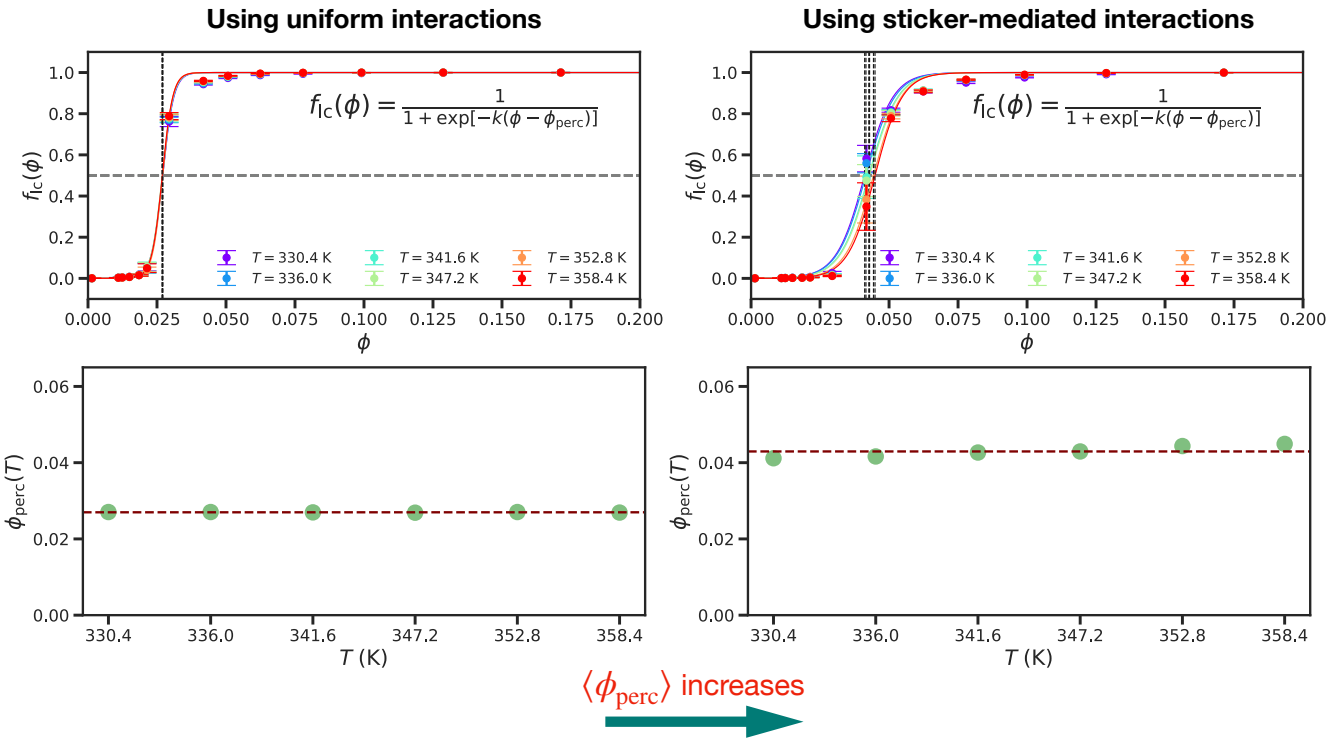

Fig. 4. Calculation of the percolation threshold ($\phi_{\mathrm{perc}}(T)$) using different ways of defining clusters. This calculation uses information regarding the fraction of molecules in the largest cluster $f_{lc}(\phi) = \frac{n_{\mathrm{largest}}}{n}$, which we compute as a function of the simulation 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 the temperature T (in Kelvin) is also shown for the two cases, and the average $\phi_{\mathrm{perc}}(T)$ increases for the latter.

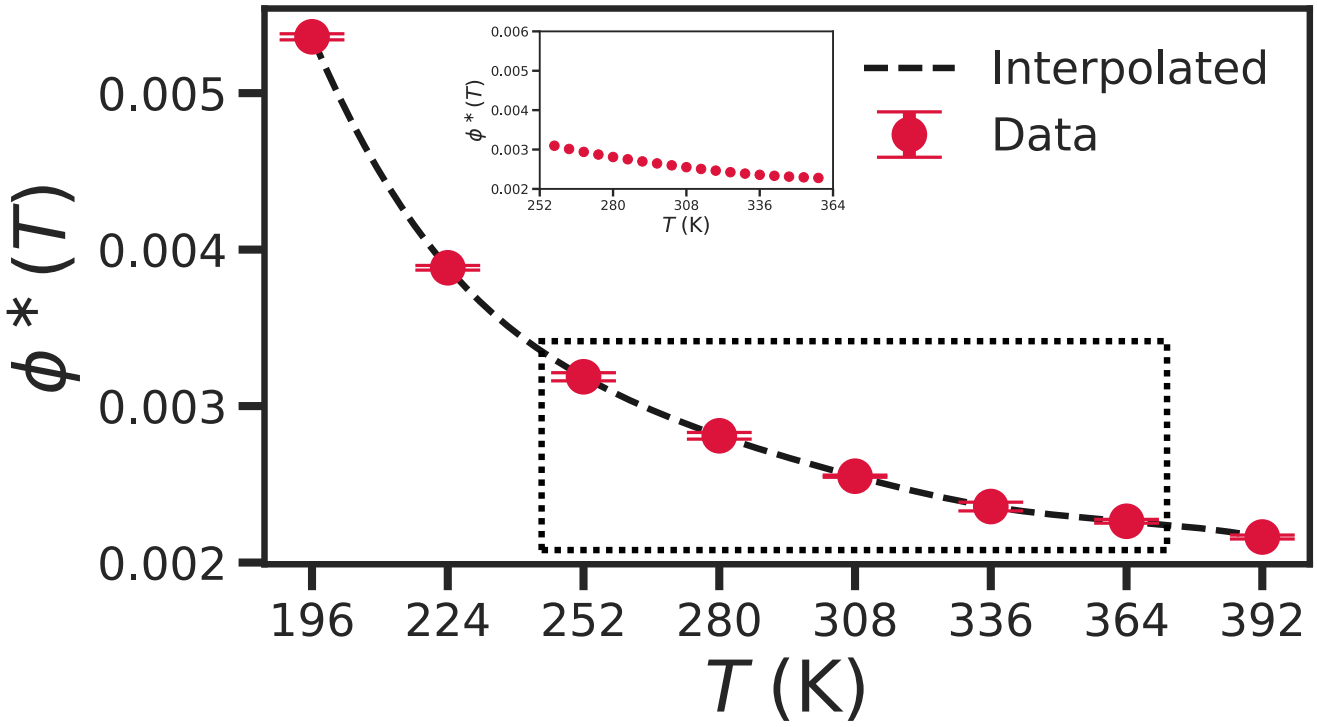

Fig. 5. Overlap volume fraction ϕ^* as a function of the temperature T. The data are shown here for the wild-type A1-LCD. The overlap volume fraction was calculated from single chain simulations. The inset shows the data over a zoomed-in range of T.

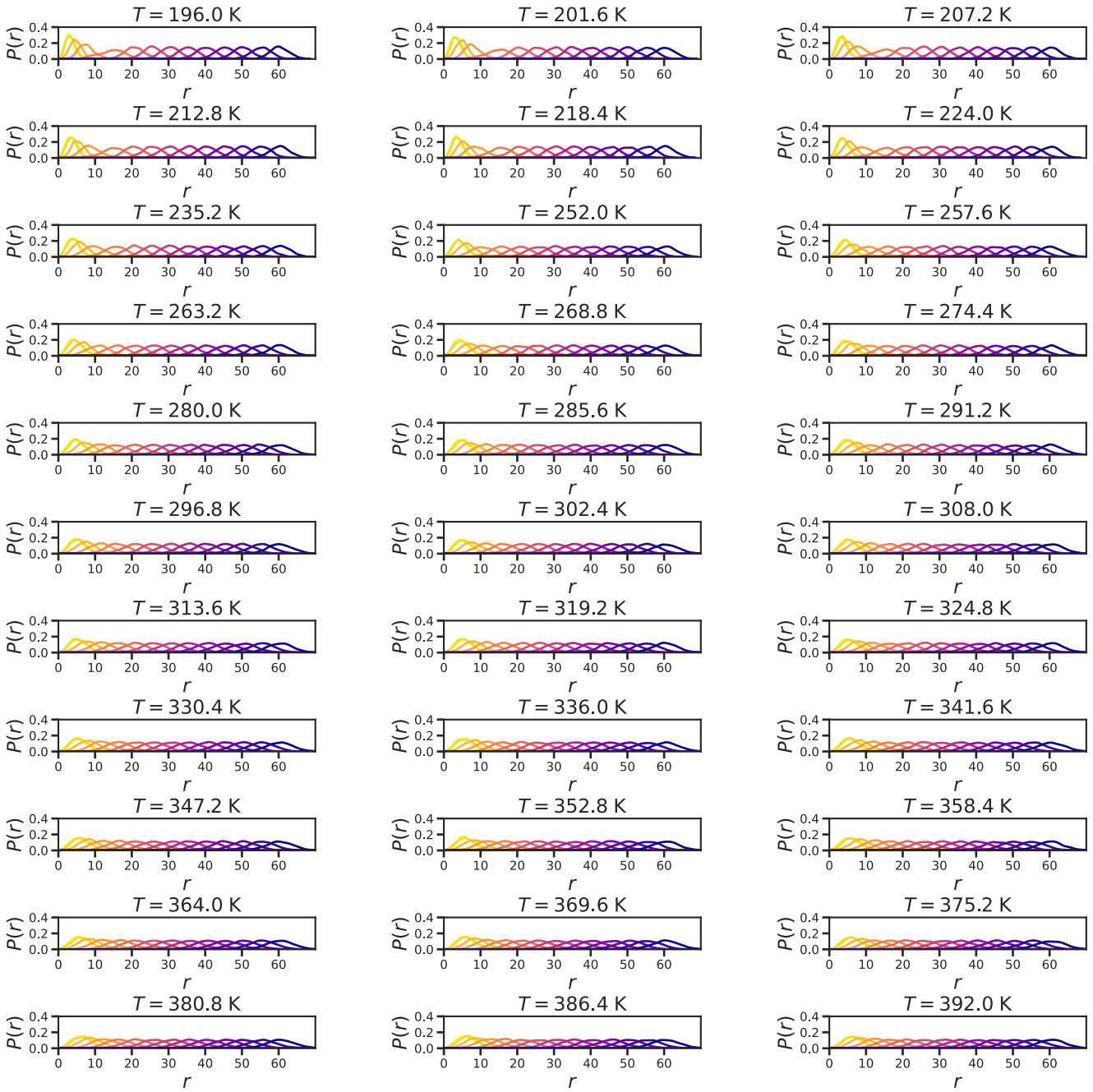

Fig. 6. Probability distributions for A1-LCD obtained from two-chain umbrella sampling simulations at different temperatures. Probability distributions $P(r)$ of the center–of–mass separation r between two A1-LCD chains obtained from umbrella sampling simulations performed at different temperatures. Each panel corresponds to a distinct temperature, and colored curves denote individual umbrella windows centered at increasing restraint distances. The spring constant k_{spring} was chosen to ensure sufficient overlap between neighboring windows thereby ensuring reliable reconstruction of the unbiased potential of mean force (PMF) using WHAM (79–81). At low temperatures, the distributions are concentrated at small separations, consistent with strong intermolecular association, whereas increasing temperature shifts probability weight toward larger separations, reflecting progressive weakening of attractive interactions.