PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY OPEN ACCESS

Dominance analysis to assess solute contributions to multicomponent phase equilibria

Daoyuan Qian[a], Hannes Ausserwoger[a], Tomas Sneideris[a], Mina Farag[b], Rohit V. Pappu[b,1], and Tuomas P. J. Knowles[a,c,1]

Affiliations are included on p. 9.

Edited by Anthony Hyman, Max-Planck-Institut fur molekulare Zellbiologie und Genetik, Dresden, Germany; received April 13, 2024; accepted July 5, 2024

Phase separation in aqueous solutions of macromolecules underlies the generation of biomolecular condensates in cells. Condensates are membraneless bodies, representing dense, macromolecule-rich phases that coexist with the dilute, macromolecule-deficient phases. In cells, condensates comprise hundreds of different macromolecular and small molecule solutes. How do different solutes contribute to the driving forces for phase separation? To answer this question, we introduce a formalism we term energy dominance analysis. This approach rests on analysis of shapes of the dilute phase boundaries, slopes of tie lines, and changes to dilute phase concentrations in response to perturbations of concentrations of different solutes. The framework is based solely on conditions for phase equilibria in systems with arbitrary numbers of macromolecules and solution components. Its practical application relies on being able to measure dilute phase concentrations of the components of interest. The dominance framework is both theoretically facile and experimentally applicable. We present the formalism that underlies dominance analysis and establish its accuracy and flexibility by deploying it to analyze phase diagrams probed in simulations and in experiments.

biomolecular condensates | phase separation | multicomponent systems | dominance analysis | tie lines

Significance

Conventional assessment of driving forces for biomolecular phase separation requires the mapping of phase boundaries and tie lines, through concentration titrations and measurements of solution components, either one at a time or a small number at a time. However, the total number of components in solution is significantly greater than the number of titratable and measurable components, leaving us blind to energetic contributions made by hidden species. Here, we establish a framework to quantify the relative energetic contribution, or dominance, of a component to phase separation with an experimentally accessible approach. In turn, this approach enables us to uncover contributions of hidden components to the driving forces for condensate formation.

Phase separation in aqueous solutions of proteins and nucleic acids is thought to be a major driver of spatial organization in cells that gives rise to mesoscale membraneless bodies known as biomolecular condensates (1, 2). In the simplest picture, a binary mixture comprising a protein or nucleic acid in a complex aqueous solvent separates into two coexisting phases above a threshold macromolecular concentration, thereby minimizing the overall free energy of the system (3). The concentrations of macromolecules and all components of the complex solvent in the coexisting phases are prescribed by conditions of phase equilibrium, viz., the equalization of species-specific chemical potentials and osmotic pressure across phases (3, 4).

In most theories of phase separation and approaches used to analyze experimental data, the contributions of solvent components, including salt, pH, buffer, and crowders, are folded into effective two- and three-body interaction parameters (5–10). Such analysis ignores the possibility of differential interphase partitioning of solvent components, which has recently been documented (11–13). We illustrate this point using the work of Bremer et al. (14, 15). They mapped temperature-dependent phase diagrams for over thirty different variants of A1-LCD, a prion-like low-complexity domain from the protein hnRNPA1 (14). We shall consider two sequences they studied, which they designated as WT (for wild-type) and −12F+12Y. In the latter, all Phe residues (F) were replaced with Tyr (Y). Everything else about the two sequences is identical. At a given temperature and for identical solution conditions, Bremer et al. quantified the threshold concentrations, designated as c_{sat}, above which each of the protein solutions undergoes phase separation. They found that the c_{sat} value of −12F+12Y is lower than that of the WT sequence. Clearly, the driving forces for phase separation are stronger for −12F+12Y when compared to WT. However, what remains unresolved is whether the differences in driving forces are due exclusively to differences in effective protein–protein interactions, with the contributions of solvent components being equivalent, or if the replacement of Phe by Tyr alters the interplay of protein–protein, protein–solvent, and solvent–solvent interactions. For example, the lowering of c_{sat} for −12F+12Y, which would be implicitly attributed to enhanced protein–protein interactions, may be the result of solvent contributions that alter the nature and strengths of three-body interactions without influencing protein–protein associations (16). Likewise, weakened protein–solvent interactions and enhanced solvent–solvent interactions, without any substantive changes to protein–protein interactions, could also account for changes in c_{sat}.

Author contributions: D.Q., H.A., R.V.P., and T.P.J.K. designed research; D.Q., H.A., T.S., and M.F. performed research; D.Q., H.A., T.S., and M.F. contributed new reagents/analytic tools; D.Q., R.V.P., and T.P.J.K. analyzed data; R.V.P. and T.P.J.K. secured funding; D.Q., H.A., T.S., M.F., R.V.P., and T.P.J.K. reviewed and edited the paper; and D.Q. and R.V.P. wrote the paper.

Competing interest statement: R.V.P. is a member of the scientific advisory board and shareholder in Dewpoint Therapeutics Inc. T.P.J.K. is a cofounder and member of the advisory board of Transition Bio Inc. All other authors have no competing interests.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: pappu@wustl.edu or tpjk2@cam.ac.uk.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2407453121/-/DCSupplemental.

Published August 5, 2024.

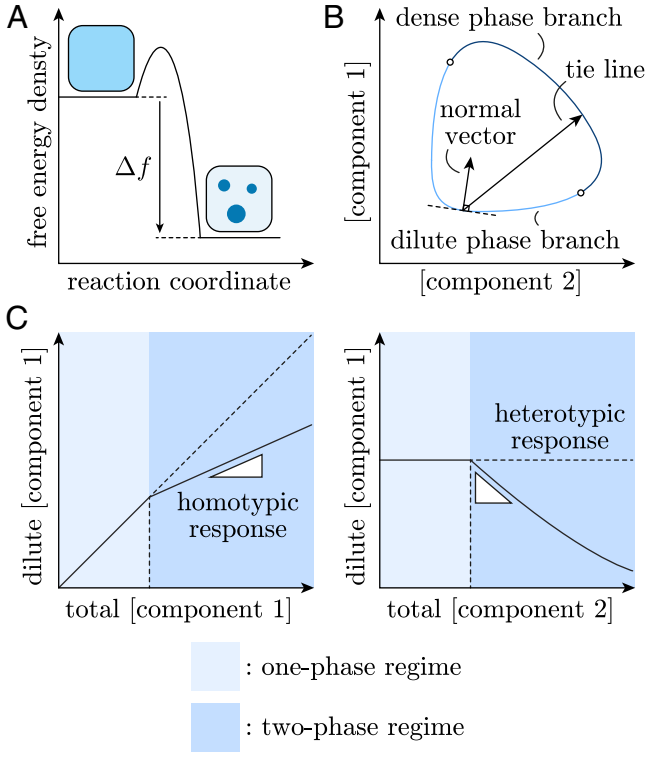

Fig. 1. Illustration of the dominance framework. (*A*) The dominance framework is developed around the free energy change Δf from the homogeneous, well-mixed state (*Left*) to the phase-separated state (*Right*), which quantifies the thermodynamic driving force of phase separation. (*B*) The output of a specific computational model is the full coexistence curve, comprising the dilute and dense phase branches of the binodal and tie lines in concentration space. A vector normal to the phase boundary characterizes its slope in a general dimension. (*C*) Experimentally accessible measurements are dilute phase responses of a component (index 1) as its total concentration or that of another component (index 2) is varied, yielding a homotypic (*Left*) or heterotypic (*Right*) response, respectively.

Being able to resolve the origins of changes to driving forces for phase separation requires knowledge of the extent to which different components contribute toward the total free energy change from a well-mixed state to the phase-separated state. We denote this stabilization free energy as Δf (Fig. 1*A*), and our goal is to dissect the contributions of different components in a multicomponent mixture to Δf.

The problem of dissecting driving forces for phase separation becomes even more important in the context of living cells. Here, condensates comprise hundreds of different proteins and nucleic acids (1, 17, 18). Further, the cellular milieu is a highly complex, nonideal, osmotic solution, made up of an assortment of ions, osmolytes, and metabolites (19, 20). Additional complexities arise from the effects of macromolecular crowding and active processes within cells (21). How does one deduce which components are important for phase separation? Riback et al. probed the effects of titrating the concentration of nucleophosmin (NPM1) on the biogenesis of nucleoli in live cells (22). They estimated the apparent saturation concentration $c_{sat,app}$ for NPM1, above which phase separation occurs. However, as the total NPM1 level increases beyond $c_{sat,app}$, the concentration of NPM1 in the nucleoplasm, which is the dilute phase that coexists with nucleoli and other nuclear condensates, does not stay fixed at $c_{sat,app}$ but increases monotonically as the total concentration of NPM1 increases. The absence of saturating or plateauing behavior suggests that components other than and in addition to NPM1 contribute to the assembly of the granular components of nucleoli. Riback et al. analyzed this feature to show how a blend of homotypic (NPM1–NPM1) and heterotypic interactions (NPM1 interacting with other molecules) contribute to the assembly of facsimiles of granular components of nucleoli. Indeed, a recent study emphasized the importance of heterotypic interactions in nucleolar organization by uncovering the contributions of asymmetrical complex coacervation (11).

To arrive at a mechanistic understanding of the interactions among different components, the typical route is to first propose a specific composition-dependent free energy functional for the macromolecular solution. Examples include the Flory–Huggins free energy (5, 6) and the sticker-spacer model (23, 24). A constrained free energy minimization is then performed to compute the dilute and dense phase branches of the phase boundary as well as tie lines that connect distinct pairs of dilute and dense phase points (25, 26) (Fig. 1*B*). One can then compare the computationally generated phase boundary and tie lines to those that are measured experimentally. This approach has been successfully employed to model valency-dependent phase separation for an artificial protein system in cells (27). However, the complexity of the model increases as the number of solution components increases. To map out the phase boundary of a system, one has to perform concentration titrations of all relevant solution components. The number of samples one has to prepare increases exponentially with the number of components in the system. Furthermore, determination of tie lines requires concentration measurements of solution components in dilute and dense phases, which is challenging due to the small dense phase volume in experiments as well as the large number of species such as ions and other buffer components that would need to be quantified in each phase. The type of measurement accessible in most laboratories is the quantification of the dilute phase concentration of just one or a small number of solutes. These solutes are usually proteins tagged with fluorescent dyes, and this dilute phase concentration can be determined as the total concentration of either the solute in question or another component is varied. When the titrated component is the same as the measured component, the response is a diagonal line in the absence of phase separation and it deviates from the diagonal once phase separation sets in; we term this a "homotypic response" (Fig. 1 *C*, *Left*). Conversely, if the titrated component is different from the one that is measured, the response is "heterotypic" and is a horizontal line in the one-phase regime (Fig. 1 *C*, *Right*).

What is currently lacking is a generally applicable approach for assessing dominant energetic contributions in multicomponent phase separation, one that leverages the principles of phase equilibria without other assumptions and relies only on experimentally accessible measurements outlined above. Here, we introduce such a framework. We consider a system comprising an arbitrary number N of solute components, where two coexisting macrophases viz., a macromolecule-rich and macromolecule-poor phase, result from phase separation. We do not assume any specific form of the free energy. Instead, we assess the information that can be gained by analyzing experimental data obtained by titrations of one or a few macromolecules. We find that near the dilute phase boundary, the stabilization free energy associated with one solute relative to the whole system is closely linked to the shape of the phase boundary, slopes of tie lines, and dilute phase solute concentrations. Accordingly, we define the dominance of a solute as the fractional contribution, quantified in terms of a dominance parameter, that the solute of interest makes to the stabilization free energy. We first demonstrate the flexibility and utility of the dominance parameter by analyzing

data from published simulations of two-component systems. By comparing the dominance trend with those generated from a Flory–Huggins model we deduce the hierarchy of interaction strengths that are consistent with the simulation results. In the second half of the work, we show that a simple dominance measurement approach exists for multicomponent systems, and this can be adapted to any system of interest. We conclude by outlining practical interpretations of dominance values in the context of interactions that modulate the driving forces for phase separation.

Defining Dominance as a Measure of Energetics

In a closed system, we denote the total concentrations of solutes in a sample as ϕ^α, where $\alpha = 1, 2, \ldots, N$. The dilute and dense phase boundaries can be represented by two $(N-1)$-dimensional surfaces in the N-dimensional concentration space, and we use ψ_-^α and ψ_+^α to denote solute concentrations on the dilute and dense surfaces, respectively. Mass balance requires that for each sample ϕ^α, the line connecting ψ_-^α and ψ_+^α at equilibrium passes through ϕ^α. We define the tie line vector $k^\alpha \equiv \psi_+^\alpha - \psi_-^\alpha$ and use v to denote the volume of the dense phase relative to the whole system. The mass balance equation can then be written as $\phi^\alpha = \psi_-^\alpha + vk^\alpha$. Furthermore, we use n_α to denote the vector normal to the dilute phase boundary, defined by $\sum_{\alpha=1}^N n_\alpha \delta\psi_-^\alpha = 0$, where $\delta\psi_-^\alpha$ is a vector lying on the dilute branch of the phase boundary. This construction of the phase space is entirely geometrical and the only assumption is that two phases coexist at equilibrium (Fig. 2*A*). Typically in differential geometry, we need to assign a metric $g^{\alpha\beta}$ that connects the units of different solute concentrations to fully define the space. It however does not play a role in the practical application of the analysis framework so we do not specify a $g^{\alpha\beta}$ here.

In a multicomponent mixture comprising incompatible components, phase separation helps minimize the total free energy of the system. We write the free energy density $f(\phi)$ of a system as a function of volume fractions of solutes ϕ^α, and since the volume is constrained in defining ϕ^α, $f(\phi)$ corresponds to the Helmholtz free energy. Under the assumption that the volume occupied by a molecule is fixed, it can be shown that $f(\phi)$ also plays the role of the Gibbs free energy density, so the results we obtain here can be applied to the practical situation where the pressure, instead of volume, is held constant (*SI Appendix*, section 1). We compute

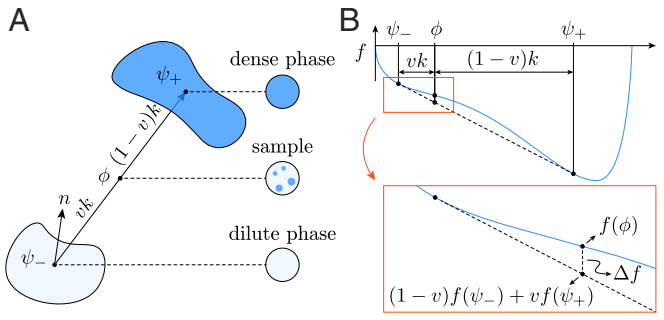

Fig. 2. Geometry and thermodynamics of a general phase-separating system. (*A*) The full N-dimensional phase space consists of dilute (light blue) and dense (dark blue) phase binodal boundaries, tie lines k^α connecting them, and dilute boundary surface normal n_α. (*B*) The free energy before and after phase separation is $f(\phi)$ and $(1-v)f(\psi_-) + vf(\psi_+)$, respectively, and a one-dimensional representation is used here but this is generalizable to any number of dimensions.

Δf in the limit of small v, as the free energy difference between a homogeneous state and a phase-separated state (Fig. 2*B*). Δf is given by (*SI Appendix*, section 2)

$$\Delta f = -\frac{1}{2}\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N} vk^\alpha vk^\beta \partial_\beta \mu_\alpha(\psi_-) + \mathcal{O}(v^3), \qquad [1]$$

where the chemical potential is $\mu_\alpha(\phi) \equiv (\partial_\alpha f)|_\phi$. $vk^\alpha = \phi^\alpha - \psi_-^\alpha$ represents the displacement of the total composition from the dilute composition, so the sum over β can be viewed as computing the change in chemical potential of solute α upon phase separation. Accordingly, we define the individual term

$$\Delta f^\alpha \equiv -\frac{1}{2} vk^\alpha \sum_{\beta=1}^{N} vk^\beta \partial_\beta \mu_\alpha(\psi_-), \qquad [2]$$

and it quantifies the free energy change associated with the partitioning of solute α. We are interested in $\lim_{v\to 0} \frac{\Delta f^\alpha}{\Delta f}$, which quantifies the relative energetic contribution of solute α at the onset of phase separation. This leads to a natural definition of the dominance of α as

$$D^\alpha \equiv \lim_{v\to 0} \frac{\Delta f^\alpha}{\Delta f} = \frac{\sum_{\beta=1}^N k^\alpha k^\beta (\partial_\alpha \partial_\beta f)|_{\psi_-}}{\sum_{\gamma=1}^N \sum_{\sigma=1}^N k^\gamma k^\sigma (\partial_\gamma \partial_\sigma f)|_{\psi_-}}, \qquad [3]$$

such that for each point ψ_-^α on the dilute phase boundary, one can calculate a specific value of D^α. Notice that D^α is dimensionless and that the sum over all solutes is constrained such that $\sum_\alpha^N D^\alpha = 1$. We are now in a position to deduce the energetic importance, from a thermodynamic viewpoint, of a solute compared to others. Hence, the name dominance. If D^α for a solute α is found to be close to 1, this implies that $D^\beta \approx 0$ for $\beta \neq \alpha$ and the phase behavior is dominated by interactions involving the component α even though other components are present in the system. In general, for a multicomponent system, we expect D^α to be nonzero for more than one solute. If we consider two components, α and β, then the relative importance of the two components as drivers of phase separation will be given by the ratio $\frac{D^\alpha}{D^\beta}$. The values of D^α, which typically lie in the interval (0,1), thus serve as a way to identify and quantify the number of components that contribute to the dominant interactions that drive phase separation in multicomponent systems.

Computing Dominance from Phase Diagrams

Having formulated the dominance D^α, we next explore how it can be computed from experimental observables. We start by analyzing simulation results for an exact two-dimensional system consisting of low-complexity domains of the pair of proteins fused in sarcoma (FUS-LCD) and hnRNPA1 (A1-LCD) (28), where the full coexistence curve is available. There are no "hidden" components in the computations since the solvent is implicit in the lattice simulations, and the only components in the system are coarse-grained representations of the two proteins. Both FUS-LCD and A1-LCD undergo phase separation on their own, and by mixing them at different ratios while keeping their total concentration fixed, a two-dimensional phase diagram with tie lines was produced previously (28), and this is reproduced here (Fig. 3*A*). The simulations were performed with a constant total concentration of A1-LCD and FUS-LCD,

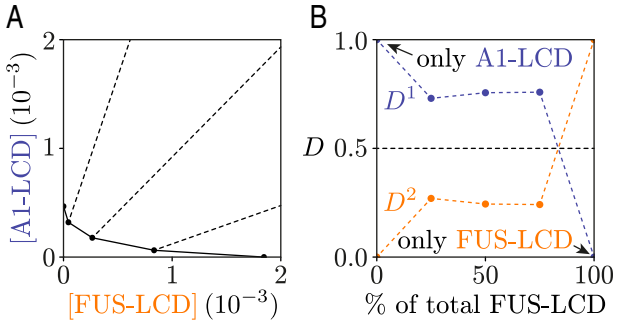

Fig. 3. Dominance can be computed from a full phase diagram. (*A*) Simulation data from ref. 28. A1-LCD and FUS-LCD concentrations are measured in terms of volume fractions, and the dilute phase binodal boundary (solid line) as well as tie lines (dashed lines) are computed at five different A1-to-FUS volume ratios. (*B*) The dilute phase boundary normal and tie line slope at each simulation point are used to compute dominance values.

and five samples were simulated with varying percentages of each protein. The simulation results revealed strong heterotypic attraction between the two proteins, and this was validated by in vitro experiments (28). Ratios between the n_α's and k^α's are computed using the slope of the phase boundary and tie lines. To obtain estimates for the dominance values for the polymers, we note that the tie line vector k^α already appears in Eq. **1**. Accordingly, we focus on the phase boundary normal n_α. By perturbing the equilibrium compositions, we arrive at $n_\alpha \propto \sum_{\beta=1}^{N} k^\beta (\partial_\alpha \partial_\beta f)|_{\psi_-}$ (*SI Appendix*, section 3) and direct substitution and comparison with Eq. **3** establishes the relationship between D^α and the geometrical quantities

$$D^\alpha = \frac{n_\alpha k^\alpha}{\sum_{\gamma=1}^{N} n_\gamma k^\gamma}. \qquad [4]$$

In the case of the two-dimensional system (28), we have $\alpha = 1, 2$, and we use index 1 to denote A1-LCD and 2 for FUS-LCD. The tie line slope k^1/k^2 and the slope of the phase boundary $-n_2/n_1$ can be combined to give the ratio $D^1/D^2 = (n_1/n_2)(k^1/k^2)$. Using $D^1 + D^2 = 1$, we obtain individual estimates of D^1 and $D^2 = 1 - D^1$ (Fig. 3*B*). Plotting the dominance of A1-LCD as a function of the fraction of FUS-LCD shows that when both proteins are present in the system, the dominance of A1-LCD is consistently larger than that of FUS-LCD, and the dominance values are by definition 0 or 1 when either protein is absent from the system. A large D^1 at intermediate amounts of FUS-LCD indicates that A1-LCD is, consistently, the dominant contributor to the minimization of the overall free energy Δf, suggesting that A1-LCD interactions are in general stronger than FUS-LCD interactions. This is corroborated by the fact that when A1-LCD and FUS-LCD are simulated on their own, A1-LCD has a lower saturation concentration than FUS-LCD (28). Can similar insights be gleaned by analyzing a two-component system described using a Flory–Huggins (5, 6) free energy? Answering this question is useful because it provides a general road-map for interpreting dominance in specific systems.

We generate a complete coexistence curve for the two-dimensional Flory–Huggins model using the convex hull algorithm (26, 29, 30) (*SI Appendix*, section 4). In our computations, we assume that the two solutes are of unit length, and we fix $\phi^1 + \phi^2 = 0.5$. Practically, we first compute the dominance value at each point on the dilute branch of the binodal. Next, we calculate the intersection of the tie line and the line $\phi^1 + \phi^2 = 0.5$. Finally, we plot the dominance as a function of $\frac{\phi^2}{\phi^1+\phi^2}$. Both solutes interact favorably with each other. We first fix the homotypic interactions χ_{11} and χ_{22} to be the same and vary the heterotypic interaction χ_{12} to be weaker than, the same as, or stronger than homotypic interactions (Fig. 4 *A*–*C*). The dominance values for the former two cases vary rapidly as the percentage of ϕ^2 increases. This trend is different from that in the data, so we deduce that the heterotypic interaction is stronger than homotypic interactions, which is in accord with published results (28). To further estimate the relative strengths of homotypic interactions, we set χ_{11} to be stronger or weaker than χ_{22} (Fig. 4 *D* and *E*). The dominance trend produced in the former case is qualitatively similar to the simulation results. We conclude that not only are the heterotypic attractions the strongest, the homotypic attraction among A1-LCD molecules is stronger than the homotypic interactions among FUS-LCD molecules. It is worth noting that the dilute phase boundary from simulations is orders of magnitude smaller than in the Flory–Huggins computation. The dimensionless nature of D^α means that this difference can be ignored and only important features of the phase space, including the shape of the phase boundary, and the slopes of tie lines, are relevant.

The A1-LCD and FUS-LCD polymers studied above both exhibit strong homotypic attractions. It is also known that phase separation in a two-dimensional system can occur with strong heterotypic interactions that are attractive or repulsive, in the absence of homotypic interactions (31). What are qualitative features of dominance values in such cases? We again use the Flory–Huggins free energy as a convenient basis to generate phase diagrams and compute dominance values. We set $\chi_{11} = \chi_{22} = 0$, and map coexistence for curves two-phase systems driven by associative interactions (Fig. 5*A*) or segregative effects (Fig. 5*B*). In presenting dominance values, we use a different parameterization: For each point on the dilute branch of the phase boundary, we compute the angle θ made between the vector connecting the origin to that point and the ϕ^1 axis, and plot D^α as a function of θ. This is qualitatively the same as plotting D^α as a function of the percentage of ϕ^2 except in the case of repulsive interactions, where the negative tie line slope means D^α stays constant if the total solute concentration is fixed. With strong heterotypic attraction (Fig. 5*A*), the computed D^1 starts off at a small value and increases as ϕ^2 is increased. This effect is of entropic origin: at low ϕ^2, the entropy of mixing of component 2 opposes phase separation. As ϕ^2 increases, this opposition diminishes due to the logarithmic scaling of the entropic contribution to the Flory–Huggins free energy. As a result, the interaction free energy, previously counterbalanced by the entropic contribution, is released. The free energy release by ϕ^2 is more significant compared to that by ϕ^1 and thus D^2 is greater than D^1. Conceptually, the dominance also reflects which solute is "limiting" in the phase separation process. For a two-component system driven by heterotypic repulsion (Fig. 5*B*), the two phases are enriched in either components. Here, we focus on the region of phase space where $\theta > \pi/4$ so that the dense phase is higher in ϕ^1. In this region, $D^1 > 1$ throughout and this implies that $D^2 = 1 - D^1 < 0$. This is explained by the fact that by phase-separating, the change in chemical potential of ϕ^1 due to partitioning of ϕ^2 releases free energy to balance out the unfavorable contribution from ϕ^2 itself, and thus ϕ^2 becomes a depletent. This behavior is expected of real crowders (32, 33), and it has been reported recently by Chauhan et al. (34) for the phase

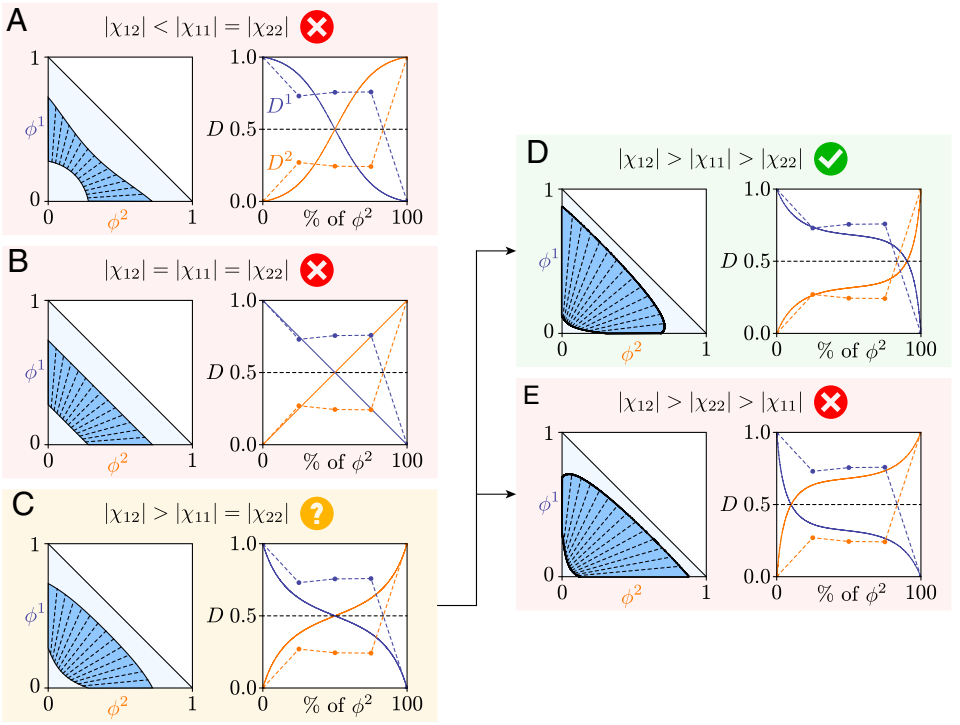

Fig. 4. Characterizing interaction hierarchy by combining dominance with the Flory–Huggins model. We compute two-dimensional phase diagrams from the Flory–Huggins model (*Left* panels) and dominance values of the two solutes (*Right* panels). Dominance values calculated from simulations in ref. 28 are scatter points connected by dashed lines. In (*A*), (*B*), and (*C*), the homotypic interactions are the same, and the heterotypic interaction is weaker than, the same as, or stronger than the homotypic interactions, respectively. Comparing the dominance trend to the A1/FUS system suggests the heterotypic interaction is likely stronger than homotypic attractions, and by further varying one of the homotypic interactions in (*D*) and (*E*), we deduce the homotypic interaction among ϕ^1 is stronger than that among ϕ^2.

behavior of transcription factors in the presence of polyethylene glycol (PEG). It is also worth noting that in both cases, as the system moves toward a critical point, D^1 approaches a singularity because $\Delta f = 0$ at the critical point, and the dominance becomes ill defined.

The computational models highlight the insights that can be gained if the dilute phase concentrations regarding multiple solutes can be measured. If D^α and ratios of n_α and k^α can be measured, then entries in the Hessian $\partial_\alpha\partial_\beta f(\psi_-)$ can be obtained as well, and these correspond to interaction energies of solutes in the dilute phase. This is based on the observation that $n_\alpha \propto \sum_{\beta=1}^{N} k^\beta \partial_\alpha\partial_\beta f(\psi_-)$. To illustrate this point, we note that $\partial_\alpha\partial_\beta f(\psi_-)$ consists of two terms: a constant, symmetric interaction energy matrix $\chi_{\alpha\beta}$ and a diagonal entropic term $\mathrm{diag}\left(\frac{1}{L^1\psi_-^1}, \frac{1}{L^2\psi_-^2}, \ldots, \frac{1}{L^N\psi_-^N}\right)$, where L^α are sizes of the solute molecules. Here, the contributions from the entropy of the solute α to the free energy of the mixture is assumed to be $\frac{\psi_-^\alpha}{L^\alpha}\ln\psi_-^\alpha$. Treating the interaction matrix and sizes L^α as unknowns, we have $N(N+1)/2+N = \frac{1}{2}N^2+\frac{3}{2}N$ fitting parameters. On the other hand, at each point ψ_-^α, we have $N-1$ linearly independent equations using ratios of n_α's. For an N-component system, by sampling $\sim \frac{1}{2}N$ points, the interaction constants $\chi_{\alpha\beta}$ and molecular sizes L^α can be deduced. Measurements of both k^α and n_α are needed, since the Hessian is contained in n_α through k^α. Practical implementation of this procedure can however be difficult, given the large number of relevant solutes. Nonetheless, the prospect of obtaining $\chi_{\alpha\beta}$ in the dilute solution not only helps with identifying the role of different solutes to phase separation but also quantifies protein interactions in general (35).

Dilute Phase Response Functions and Tie Lines

The two-dimensional systems investigated so far serve as illustrative examples of the dominance formulation. We now turn to the realistic scenario where the number of solute species can be more than two. In addition, concentrations of many of these species can be difficult to quantify, so full phase diagrams and tie lines are rarely accessible. We observe, however, that in a typical experiment, there will be at least one macromolecular solute, say α, whose dilute phase concentration is quantifiable (14, 31). Furthermore, one can also change the total starting concentrations of any solute β, so an accessible quantity of interest is the response $R_\beta^\alpha \equiv \frac{\partial\psi_-^\alpha}{\partial\phi^\beta}$. Via mass conservation, one can show that changes in ψ_-^α, in response to changes in ϕ^β, evaluated at the phase boundary, can be expressed as (*SI Appendix*, section 5)

$$\lim_{v\to 0} R_\beta^\alpha = \delta_\beta^\alpha - \frac{n_\beta}{n_\alpha}D^\alpha. \qquad [5]$$

δ_β^α is the Kronecker delta. The appearance of D^α indicates that only the dominance of α is accessible when the dilute phase concentration of α is measured, regardless of which solute β is varied to measure the response. This expression clearly shows that in order to measure the extent to which a solute α contributes to the phase behavior of the mixture, one must measure the dilute phase concentration of the solute in question. These measurements must be performed as the concentration of other solutes is perturbed. Interestingly, by combining two response functions, we can further deduce information regarding the tie line components k^α and k^β.

Previously (31), we established a tie line measurement approach between a solute with index 1 (without loss of

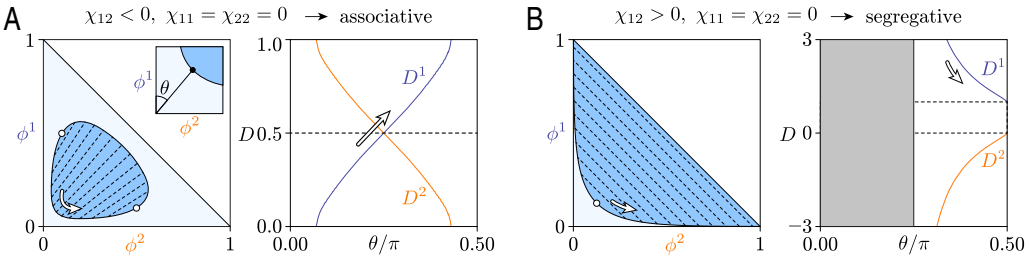

Fig. 5. Dominance trends in heterotypic interaction-driven Flory–Huggins model. When phase separation is driven purely by heterotypic interactions, both attraction (*A*) and repulsion (*B*) can lead to demixing. *Left* panels show the two-dimensional phase diagrams with hollow markers as critical points, and computed D^α values are plotted in the *Right* panels. *Left Inset*: θ is defined as the angle formed between the dilute phase point where the dominance is evaluated, and the ϕ^1 axis. As the point of evaluation approaches the critical points, the D^α values encounter a singularity as Δf goes to 0. An interesting observation in (*B*) is that here, D^1 takes on values not in the range of $(0, 1)$, reflecting the abnormal behavior of a repulsion-driven phase-separating system.

generality), and another solute with index 2 through measurement of ψ_-^1 alone. Using the response function formulation we can formalize this concept into the following question: How should ϕ^1 and ϕ^2 be changed so that ψ_-^1 stays constant? Mathematically, this requires $\delta\psi_-^1 = 0$. Recall that $\frac{\partial\psi_-^\alpha}{\partial\phi^\beta} \equiv R_\beta^\alpha$, and $\delta\psi_-^1 = R_1^1\delta\phi^1 + R_2^1\delta\phi^2$ since all other components are fixed $\delta\phi^3 = \delta\phi^4 = \cdots = \delta\phi^N = 0$. The space defined by $\delta\psi_-^1 = 0$ on the 1−2 plane is a line, and its slope is $K_2^1 \equiv \left(\frac{\partial\phi^1}{\partial\phi^2}\right)_{\psi_-^1} = -\frac{R_2^1}{R_1^1}$. Close to the dilute phase boundary we can use Eqs. **4** and **5** to rewrite K_2^1 as

$$\lim_{v\to 0} K_2^1 = \frac{n_2}{n_1}\frac{D^1}{1-D^1} = \frac{k^1}{k^2}\left[\frac{1}{1+\sum_{\beta\neq 1,2}^N D^\beta/D^2}\right], \quad [6]$$

such that, to leading order, K_2^1 provides an approximation of the ratio of tie line components k^1/k^2 close to the phase boundary. Terms in the square bracket of Eq. **6** represent the relative deviation of K_2^1 from k^1/k^2, and in a true two-component system we have $\lim_{v\to 0} K_2^1 = k^1/k^2$. On the 1−2 plane, this dilute phase contour defined by $\delta\psi_-^1 = 0$ can be thought of as an N-dimensional tie line reduced to the two-dimensional plane.

Measuring K_2^1 allows an estimation of the relative solute partitioning to be made. Notably, if the dominance of all solutes falls within the range $(0, 1)$, the sign of K_2^1 is entirely determined by the sign of n_2/n_1, and thus the sign of k^1/k^2 can be deduced by observing the effect of the solute 2 on condensates. To put this in practical terms, if solute 2 dissolves condensates ($n_2/n_1 < 0$) it partitions out of condensates ($k^1/k^2 < 0$) and vice versa. This makes intuitive sense since it suggests that condensate-favoring modulators partition into condensates while condensate-disfavoring modulators are excluded from condensates. It is important however to point out here that some D^1 can fall outside the 0 and 1 range, and this can happen when a solute is able to change the chemical potentials of other solutes to make phase separation especially favorable, even though its intrinsic energetic contribution opposes phase separation. This is the case of true crowders, as explained in the previous section.

Response Measurement Using Line-Scans

We now investigate practical applications of dominance analysis and interpretations in a range of systems. The response function expression Eq. **5** suggests that measurements of concentrations in the dilute phase are necessary to obtain the dominance of a solute. In the following, we use the index 1 to denote the solute that is measured. Experimentally, a series of samples can be prepared with varying ϕ^1 (homotypic line-scan) or some other ϕ^2 (heterotypic line-scan) while keeping all other ϕ's constant and measuring ψ_-^1 (Fig. 6). This corresponds to traversing the phase space along a line parallel to ϕ^1 or ϕ^2 axis and measuring the response functions R_1^1 or R_2^1. By performing the measurement close to the dilute phase boundary we can rewrite response expressions as

$$\begin{aligned} R_1^1 &= 1 - D^1, \\ R_2^1 &= -\frac{n_2}{n_1}D^1. \end{aligned} \quad [7]$$

In the following, we analyzed multicomponent experimental data using the response formulae to obtain dominance values and partitioning information where applicable.

Phase separation of the optoG3BP1 protein in cells was studied in ref. 22, where the dilute phase [optoG3BP1] was measured as a function of total [optoG3BP1]. We use the index 1 to denote optoG3BP1. We have $R_1^1 = 1$ outside the two-phase regime, and the response is modulated beyond some threshold value ϕ_c^1 (Fig. 7*A*). This is in essence the apparent threshold or saturation concentration $c_{\text{sat,app}}$. In a pseudo one-component system featuring a solution in a mean-field solvent, for $\phi^1 > \phi_c^1$ we expect $R_1^1 = 0$. Instead, we find $R_1^1 > 0$ for optoG3BP1 because of the multicomponent nature of the system (22). Using a two-segment linear fit for optoG3BP1 data we obtain $D^1 = 0.55 \pm 0.14$. In contrast, in another system involving the protein optoFUS (36), the dominance of optoFUS is $D^1 = 0.98 \pm 0.07$ (Fig. 7*B*). This is consistent with the observation that FUS phase separates on its own (37, 38) and is less reliant on other components. This contrasts with G3BP1, where other macromolecules such as RNA are typically added to G3BP1 solution to trigger phase separation (39, 40). A simple dominance measurement thus allows a quick assessment of the multicomponent character, similar to what Guillen-Boixet et al., demonstrated (39).

Next, we explored how the dominance of a component responds to the addition of modulators. For this, we investigated the phase behavior of poly(rA)-RNA upon addition of salt, which phase separates in the presence of PEG and at high [KCl] (41). We perform what we refer to as homotypic line-scans [dilute phase concentrations of poly(rA) are measured as total poly(rA) is increased] with 4% PEG (molecular weight 10 kDa), and 800 or 850 mM of KCl (*Materials and Methods*). We use index 1 to denote poly(rA) and index 2 for KCl. The difference in total KCl

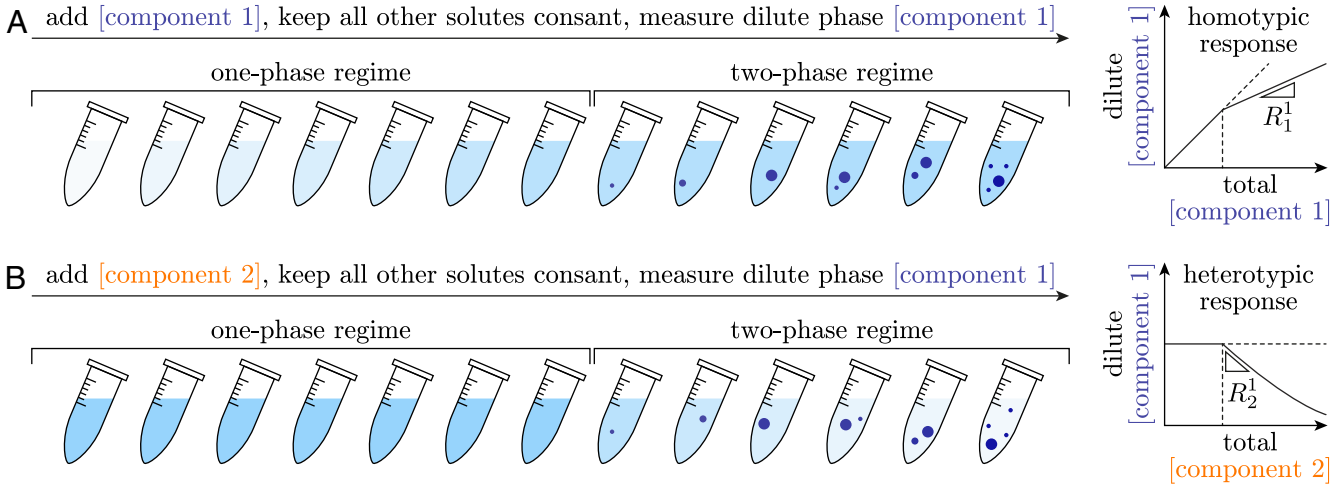

Fig. 6. Experimental measurement of response functions. (*A*) For measuring the homotypic response of component 1, one first prepares a series of samples by varying total [component 1] while keeping all other component concentrations fixed (*Left*). Dilute phase concentration of component 1 needs to be determined next, and this can be done by spinning down the samples to remove the dense phase, followed by fluorescent intensity measurement. Finally, plotting the dilute [component 1] against total [component 1] produces the homotypic response curve (*Right*), and the slope of the response at the onset of phase separation is the R_1^1 of interest. (*B*) The heterotypic response of component 1 against another component 2 can be determined in a similar fashion.

concentrations in the two scans is $\Delta\phi^2 = 50$ mM. Line-scan data show a smaller R_1^1 after phase separation at higher KCl, so D^1 increases from 0.49 ± 0.05 to 0.66 ± 0.04 (Fig. 7*C*). The increase in poly(rA) dominance at higher KCl can be explained by the increased screening of negative charges on poly(rA), facilitating inter-poly(rA) interactions and rendering other interactions, for

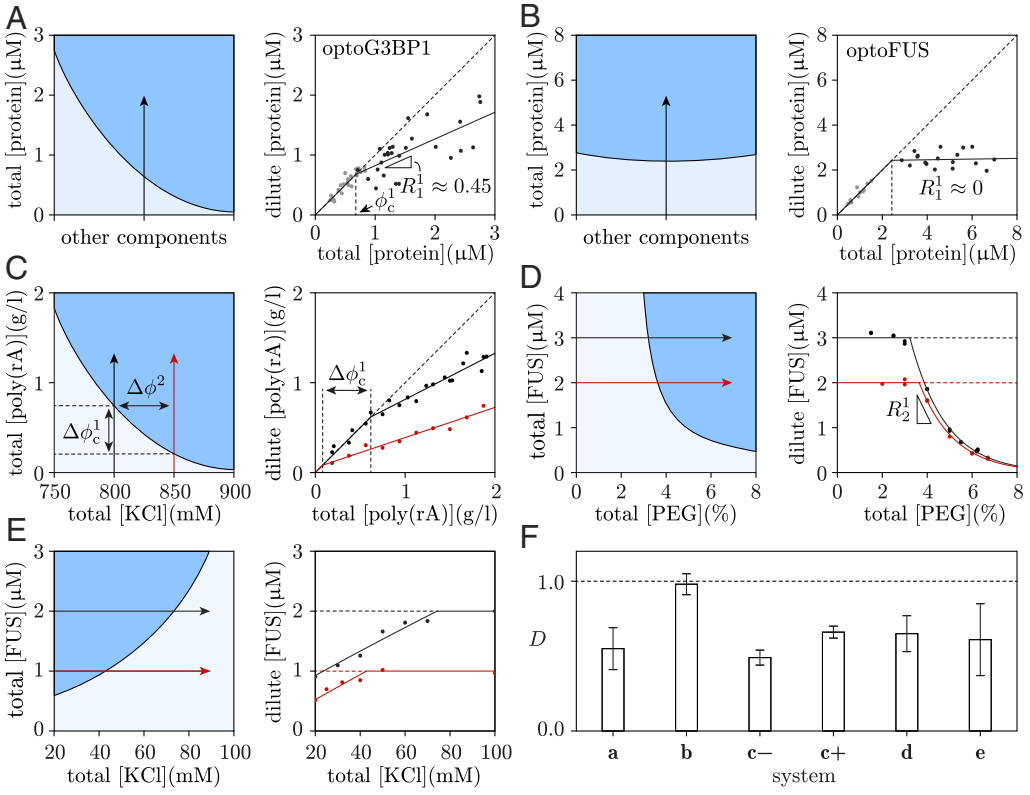

Fig. 7. Extracting dominance and partitioning information from line-scan data. In (*A–E*), *Right* panels illustrate cross-sections of phase spaces, *Left* panels present the corresponding data. Arrows in illustrations represent how line-scans are traversing the phase spaces. In data panels, scatter points are data and solid lines are phenomenological fits. Dotted lines indicate trivial responses in the absence of phase separation. (*A*) The dilute phase response of optoG3BP1 (22) exhibits multicomponent character, so phase separation is likely to depend on other components, emphasized with a phase boundary that depends on other components. (*B*) In the optoFUS system (36) however the protein has much stronger single-component character. (*C*) Poly(rA)-RNA phase-separates in the presence of PEG and KCl. At higher [KCl] the increased charge screening allows inter-RNA interaction to be more favorable, leading to an increase in D^1 and decrease in ϕ_c^1 when [KCl] is increased. (*D*) Heterotypic line-scan of FUS against PEG (31). At low [PEG] no phase separation is observed, so the dilute phase [FUS] is simply a constant. When the line-scan trajectory enters the phase-separated region the dilute phase [FUS] starts decreasing, and this initial slope is taken as R_2^1. (*E*) Heterotypic line-scan of FUS against KCl (12). Here, the plateau appears at high [KCl] because KCl dissolves condensates. (*F*) Overview of dominance values collected. − and + labels for (*C*) correspond to low and high [KCl], respectively.

instance poly(rA)/PEG, less significant. The onset of phase separation shifts to a lower poly(rA) concentration when salt concentration is increased, with $\Delta\phi_c^1 = -0.54 \pm 0.12$g/L. This can be interpreted as an increase in the overall drive for phase separation, and it is consistent with the picture that inter-poly(rA) repulsion weakens with increased screening at higher salt concentrations. By combining $\Delta\phi^2$ with $\Delta\phi_c^1$ we can infer the ratio of surface normals $\frac{n_2}{n_1}$ in the poly(rA)-KCl plane. By definition, $n_1\Delta\phi_c^1 + n_2\Delta\phi^2 + n_3\Delta\phi^3 + \cdots + n_N\Delta\phi^N = 0$, and since the line-scans are performed while fixing $\Delta\phi^3 = \cdots = \Delta\phi^N = 0$ we have

$$\frac{n_2}{n_1} = -\frac{\Delta\phi_c^1}{\Delta\phi^2}, \quad [8]$$

and this gives $\frac{n_2}{n_1} = -\frac{\Delta\phi_c^1}{\Delta\phi^2} = 10.8 \pm 2.4 \frac{\text{g/l}}{\text{M}}$. By combining this with the average D^1, we can further calculate $K_2^1 = 9.4 \pm 1.8 \frac{\text{g/l}}{\text{M}}$. The simple inference is that KCl partitions preferentially into poly(rA) condensates. In reality, we expect the interphase partitioning of K^+ and Cl^- to be different and dependent on the total concentration of KCl, and this information can only be gained by measurements of individual ionic species. Mathematically, treating KCl as a single solute is equivalent to performing an axis rotation, where the K^+ and Cl^- axes are mixed into a new axis. This does not affect $D^1 = \frac{n_1 k^1}{\sum_{\gamma=1}^{N} n_\gamma k^\gamma}$ since the denominator is coordinate-independent and the numerator is not affected. Furthermore, since k^α and ϕ^α transform in the same way, the measured K_2^1 represents the summed partitioning of K^+ and Cl^-, but it is blind to the partitioning of individual ionic species.

Although a homotypic line-scan is a natural path toward quantifying D^1, in experiments the heterotypic line-scan can also be performed (12, 31). To illustrate the approach, we take two sets of data: FUS/PEG from ref. 31 and FUS/KCl from ref. 12. In both cases dilute phase FUS protein (index 1) concentration is measured, while the total concentrations of PEG or KCl (index 2) are varied. In the FUS/PEG line-scan data from ref. 31, two line-scans were performed at 2 and 3 μM [FUS] while adding PEG up to 8% weight fraction. The dilute phase [FUS] measurements showed a plateau at low [PEG] and an exponential decrease as some threshold [PEG] is reached (Fig. 7*D*). The two points at the onset of the decrease are where the lines cross the phase boundary, and from these, we estimate $\frac{n_2}{n_1} = 2.4 \pm 0.4 \frac{\mu\text{M}}{\%}$. The response function R_2^1 can be evaluated at the drop-off points as well, and averaging over the two scans gives $R_2^1 = -1.58 \pm 0.07 \frac{\mu\text{M}}{\%}$. Combining $\frac{n_2}{n_1}$ and R_2^1 gives $D^1 = 0.65 \pm 0.12$. This dominance value indicates there are other solutes actively participating in phase separation and a likely candidate is PEG. This is also reflected in the fact that the FUS/PEG phase boundary measured using PhaseScan has a strong dependence on [PEG] (42). Furthermore, substituting $\frac{n_1}{n_2}$ and the average D^1 into Eq. **6**, we get $K_2^1 = 4.5 \pm 1.8 \frac{\mu\text{M}}{\%}$, providing an estimate of the relative partitioning between FUS and PEG.

In the FUS/KCl system (12), KCl has the opposite effect compared to PEG as condensates are dissolved when more KCl is added to the system (12). Line-scan data show a steady increase of dilute phase [FUS] with increasing [KCl] (Fig. 7*E*). We fit the dilute phase [FUS] after phase separation using a linear function. Repeating the same calculations we obtain $\frac{n_2}{n_1} = -0.033 \pm 0.009 \frac{\mu\text{M}}{\text{mM}}$ and $R_2^1 = 0.020 \pm 0.003 \frac{\mu\text{M}}{\text{mM}}$ to give $D^1 = 0.61 \pm 0.24$. This again suggests a FUS-dominated, multicomponent system. The K_2^1 value obtained is $-0.05 \pm 0.04 \frac{\mu\text{M}}{\text{mM}}$. The implication is that overall, KCl partitions outside condensates, exerting a destabilizing effect. All dominance values obtained are summarized in a bar plot (Fig. 7*F*). Heterotypic line-scans are thus an alternative to homotypic line-scans.

Discussion

Assessing Energetic Contributions of Proteins in More Complex Condensates. Cellular condensates such as stress granules are mesoscale, multicomponent entities. Yang et al., used systematic knockdown experiments to identify G3BP1/2 as the central node in the network of proteins that make up stress granules (40). Sanders et al., showed that different nodes in the network contribute differently to the G3BP1/2-mediated assembly of stress granules (43), which is potentiated by stress-induced polysomal runoff that generates protein-free, unfolded RNA molecules (39). As proteome-scale investigations become more widely used (44–46), it becomes increasingly feasible to identify nodes in protein–protein and protein-nucleic acid interaction networks that are key contributors to condensate assembly. Assessing the contributions of different condensate components will likely involve genetic manipulations of expression levels that enable homotypic and/or heterotypic scanning (22, 43, 44, 47–49). However, the number of components within the condensate will always vastly exceed the number of independent axes along which solute concentrations can be titrated within cells (50). Deployment of dominance analysis introduced in this work, and generalization of this analysis will help quantify the contributions of titratable versus nontitratable and hence hidden components to the driving forces for condensation.

Relationship of Energy Dominance with Other Thermodynamic Measures? The dominance value formulated in this work provides a practical way to rank the importance of the contributions of different components in a system. However, there are many other measures that can also be used to compare the relative contributions of different components. For instance, the difference $\psi_+^\alpha - \psi_-^\alpha$, in units of volume fraction, is a possible candidate, as is the ratio $\psi_+^\alpha/\psi_-^\alpha$ or a combination of the two. Such measures quantify different aspects of phase separation: The difference $\psi_+^\alpha - \psi_-^\alpha$ quantifies interphase concentration gradients that derive from equalizing species-specific chemical potentials, while the ratio $\psi_+^\alpha/\psi_-^\alpha$ encodes the difference in the entropy of α. Ultimately, however, they all contribute to the free energy decrease Δf, so basing the dominance measure on Δf ensures that we can view it as a combination of all possible measures, weighted based on thermodynamic considerations. The dimensionless nature of dominance also means that physical characteristics such as molecular size, charge, and hydrophobicity, are all encompassed in this measure, allowing molecules of vastly different properties to be compared on an equal footing.

Combining the Assumption-Free Dominance Formulation with Other Theoretical Models. In the first half of this work, we used the Flory–Huggins model to generate phase diagrams and interpret dominance values in terms of interaction energies, and this approach can be extended to other models too. Recent studies have adapted the polyphasic linkage formalism of Wyman and Gill (51) to understand how ligands that bind site-specifically can exert an influence on the driving forces for phase separation

via linkage between binding and phase equilibria (52–54). The formalism of binding polynomials, which underlies linkage relationships, is also free of any assumptions regarding the models used to describe either phase separation or binding. Combining the formalism of polyphasic linkage, which rests on quantifying how ligands modulate the driving forces for phase separation, and dominance analysis, which dissects the energetic contributions of different solutes, is likely to be a useful route for uncovering differences in binding modes across the phase boundary. Moreover, in multicomponent systems, the interplay between homotypic and heterotypic interactions determines the shapes of phase boundaries and the slopes of tie lines. Lin et al., recently showed that the shapes of phase boundaries, mapped in terms of concentrations of interaction motifs rather than molecules, can be used to dissect the relative contributions of homotypic versus heterotypic interactions (55). A comparative assessment of the shape analysis introduced by Lin et al. and the dominance analysis introduced here would be valuable, especially for analyzing low-dimensional representations of phase boundaries of multicomponent systems, which are inherently multidimensional in nature. Other contributions including the concept of solubility products, which was recently introduced by Chattaraj et al. (56), and the work of Deviri and Safran (57), both of which focused on buffering in the presence of heterotypic interactions, can be augmented by the dominance analysis introduced here.

Generalizing the Formulation to Multiphase Coexistence. The dominance formulation presented here is built on the assumption of two coexisting phases. Extending the algebra to multiphase scenarios is straightforward. An interesting result that follows for an M-phase, N-component system is that the trace of the response is $\sum_{\alpha=1}^{N} R_{\alpha}^{\alpha} = N - M + 1$. However, a discussion of this slightly more complex scenario requires additional considerations. First, a multiphase system comprising M phases with one dilute phase such as the nucleoplasm or cytosol can be viewed as the concatenation of $(M-1)$ two-phase systems, with each dense phase in equilibrium with a common dilute phase. Understanding dominance within each of the two-phase systems is thus equivalent to understanding dominance in the multiphase system. Second, the relationship between the response R_{β}^{α} and the energetics of condensate formation is exact in the $v \to 0$ limit, practically taken by varying the concentration of a component in line-scan toward condensate dissolution. A natural consequence of this protocol is that in an experiment, even if the system has a multiphase regime, close to compositions where almost all condensates have dissolved, all condensates will be of the same type, rendering a multiphase model irrelevant for practical purposes. This is because in the phase space, an M-phase region is completely surrounded by $(M-1)$-phase regions, which are in turn surrounded by $(M-2)$-phase regions, and so on. This is experimentally observed (41) and computationally illustrated (26). As such, if a line-scan is performed where the system can potentially form more than 1 dense phase, the response R_{β}^{α} immediately outside the homogeneous region almost surely (with probability 1) still corresponds to a case of two-phase coexistence, before multiphase effects set in. Mathematically, out of all compositions on the $(N-1)$-dimensional phase boundary separating the homogeneous region and phase-separated region in concentration space, compositions that correspond to 3-and-more phase equilibrium reside on manifolds of lower dimensions than the two-phase equilibrium manifold. As a result, if we consider the $(N-1)$-dimensional Lebesgue measure defined on the phase boundary, manifolds comprising 3-and-more phase equilibrium compositions have Lebesgue measure 0, and thus the experimentally measured responses near the phase boundary are almost surely (with probability 1) those that originate from two-phase coexistence. As a result, as one moves toward the homogeneous region, the number of phases steadily decreases to 2, before the last remaining type of condensate dissolves.

Conclusions. We have derived and applied an energy dominance analysis to elucidate the relative importance of the interactions of specific components in a multicomponent mixture that undergoes phase separation. We demonstrated the accuracy and flexibility of the framework by deploying it to analyze simulation results for mixtures of low-complexity domains as well as experimentally measured dilute phase response functions. Formally, the dimensionless nature of the dominance parameter allows direct comparison between solutes that are distinct from one another. This makes it possible to compare the dominance of a macromolecular solute with that of a small molecule species, if the dilute phase concentration of the latter can be measured. Dominance analysis takes on special relevance given the likely contributions of metabolites, osmolytes, other naturally occurring and drug-like molecules (20, 58). Combining the dominance framework with detailed theoretical models can further facilitate understanding of phase-separating systems. As such, the dominance formalism presents a conceptual as well as practical advancement in the condensate field.

Materials and Methods

Poly(rA)-RNA with a molecular weight of 700 to 3,500 kilodaltons is purchased from Sigma, and a stock solution of 4 g/L prepared by dissolving lyophilized powder in Milli-Q water. PEG with molecular weight of 10 kilodaltons is purchased from Sigma and dissolved in 1 M KCl, 50 mM HEPES at pH 7.3 at 20% by weight. The working solutions are at 8% PEG, 1.6 M or 1.7 M KCl, 50 mM HEPES, and pH 7.3. These are prepared by mixing the 20% PEG solution with solutions of 4 M KCl, 50 mM HEPES, and 0 M KCl, 50 mM HEPES.

Experimental samples are prepared by mixing the 8% PEG working solution, Poly(rA) stock solution, and Milli-Q water in PCR tubes, with a total volume of 20 μL. After mixing and vortexing, the samples are spun down at 13,400 rpm for 2 min and 2 μL of the supernatant is taken from each sample to measure the dilute phase [Poly(rA)] in a NanoDrop machine (ThermoFisher).

Data, Materials, and Software Availability. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. This work was funded by grants from Global Research Technologies, Novo Nordisk A/S (H.A., T.P.J.K.), the European Research Council under the European Union's Horizon 2020 research and innovation programme through the Marie Sklodowska-Curie grant MicroREvolution (Agreement ID 101023060; T.S.), the Newman Foundation (T.S. and T.P.J.K.), the European Research Council under the European Union's Horizon 2020 research and innovation program through the ERC grant DiProPhys (Agreement ID 101001615; T.P.J.K.), the US NIH (R01NS121114 to R.V.P), and the US NSF (MCB-2227268 to R.V.P).

Author affiliations: [a]Centre for Misfolding Diseases, Yusuf Hamied Department of Chemistry, University of Cambridge, CB2 1EW Cambridge, United Kingdom; [b]Department of Biomedical Engineering and Center for Biomolecular Condensates, Washington University in St. Louis, St. Louis, MO 63130; and [c]Cavendish Laboratory, Department of Physics, University of Cambridge, CB3 0HE Cambridge, United Kingdom

1. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
2. Y. Shin, C. P. Brangwynne, Liquid phase condensation in cell physiology and disease. *Science* **357**, aaf4382 (2017).
3. R. V. Pappu, S. R. Cohen, F. Dar, M. Farag, M. Kar, Phase transitions of associative biomacromolecules. *Chem. Rev.* **123**, 8945–8987 (2023).
4. J. W. Gibbs, On the equilibrium of heterogeneous substances. *Am. J. Sci. Arts* **16**, 441 (1874).
5. P. J. Flory, Thermodynamics of high polymer solutions. *J. Chem. Phys.* **10**, 51–61 (1942).
6. M. L. Huggins, Solutions of long chain compounds. *J. Chem. Phys.* **9**, 440 (1941).
7. M. Muthukumar, Thermodynamics of polymer solutions. *J. Chem. Phys.* **85**, 4722–4728 (1986).
8. G. Raos, G. Allegra, Chain collapse and phase separation in poor-solvent polymer solutions: A unified molecular description. *J. Chem. Phys.* **104**, 1626–1645 (1996).
9. D. Qian, T. C. Michaels, T. P. Knowles, Analytical solution to the Flory–Huggins model. *J. Phys. Chem. Lett.* **13**, 7853–7860 (2022).
10. M. Farag, A. S. Holehouse, X. Zeng, R. V. Pappu, FIREBALL,, A tool to fit protein phase diagrams based on mean-field theories for polymer solutions. *Biophys. J.* **122**, 2396–2403 (2023).
11. M. R. King *et al*., Macromolecular condensation organizes nucleolar sub-phases to set up a pH gradient. *Cell* **187**, 1889–1906.e24 (2024).
12. H. Ausserwoger *et al*., Condensate partitioning governs the mechanism of action of FUS phase separation modulators. bioRxiv [Preprint] (2023). https://doi.org/10.1101/2023.05.31.543137 (Accessed 5 June 2023).
13. H. Ausserwöger *et al.*, Biomolecular condensates sustain pH gradients at equilibrium driven by charge neutralisation. bioRxiv [Preprint] (2024). https://www.biorxiv.org/content/10.1101/2024.05.23.595321v1 (Accessed 25 June 2024).
14. A. Bremer *et al*., Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. *Nat. Chem.* **14**, 196–207 (2022).
15. A. Bremer *et al*., Quantifying coexistence concentrations in multi-component phase-separating systems using analytical HPLC. *Biomolecules* **12**, 1–12 (2022).
16. J. M. Choi, A. A. Hyman, R. V. Pappu, Generalized models for bond percolation transitions of associative polymers. *Phys. Rev. E* **102**, 1–6 (2020).
17. W. Xing, D. Muhlrad, R. Parker, M. K. Rosen, A quantitative inventory of yeast P body proteins reveals principles of composition and specificity. *eLife* **9**, 1–63 (2020).
18. A. Hubstenberger *et al*., P-body purification reveals the condensation of repressed mRNA regulons. *Mol. Cell* **68**, 144–157.e5 (2017).
19. R. Milo, R. Phillips, *Cell Biology by the Numbers* (Garland Science, 2015).
20. M. Kar *et al*., Solutes unmask differences in clustering versus phase separation of FET proteins. *Nat. Commun.* **15**, 4408 (2024).
21. A. R. Subramanya, C. R. Boyd-Shiwarski, Molecular crowding: Physiologic sensing and control. *Annu. Rev. Physiol.* **86**, 1–24 (2024).
22. J. A. Riback *et al*., Composition-dependent thermodynamics of intracellular phase separation. *Nature* **581**, 209–214 (2020).
23. A. N. Semenov, M. Rubinstein, Thermoreversible gelation in solutions of associative polymers. 1. Statics. *Macromolecules* **31**, 1373–1385 (1998).
24. J. M. Choi, A. S. Holehouse, R. V. Pappu, Physical principles underlying the complex biology of intracellular phase transitions. *Annu. Rev. Biophys.* **49**, 107–133 (2020).
25. Y. H. Lin, J. Wessén, T. Pal, S. Das, H. S. Chan, "Numerical techniques for applications of analytical theories to sequence-dependent phase separations of intrinsically disordered proteins" in *Phase-Separated Biomolecular Condensates* (2023), pp. 51–94.
26. S. Mao, D. Kuldinow, M. P. Haataja, A. Košmrlj, Phase behavior and morphology of multicomponent liquid mixtures. *Soft Matter* **15**, 1297–1311 (2019).
27. M. Heidenreich *et al*., Designer protein assemblies with tunable phase diagrams in living cells. *Nat. Chem. Biol.* **16**, 939–945 (2020).
28. M. Farag, W. M. Borcherds, A. Bremer, T. Mittag, R. V. Pappu, Phase separation of protein mixtures is driven by the interplay of homotypic and heterotypic interactions. *Nat. Commun.* **14**, 5527 (2023).
29. D. D. Lee, J. H. Choy, J. K. Lee, Computer generation of binary and ternary phase diagrams via a convex hull method. *J. Phase Equil.* **13**, 365–372 (1992).
30. J. Wolff, C. M. Marques, F. Thalmann, Thermodynamic approach to phase coexistence in ternary phospholipid–cholesterol mixtures. *Phys. Rev. Lett.* **106**, 1–4 (2011).
31. D. Qian *et al*., Tie-line analysis reveals interactions driving heteromolecular condensate formation. *Phys. Rev. X* **12**, 041038 (2022).
32. A. Testa *et al*., Sustained enzymatic activity and flow in crowded protein droplets. *Nat. Commun.* **12**, 1–8 (2021).
33. J. M. Choi, F. Dar, R. V. Pappu, LASSI: A lattice model for simulating phase transitions of multivalent proteins. *PLoS Comput. Biol.* **15**, e1007028 (2019).
34. G. Chauhan, A. Bremer, F. Dar, T. Mittag, R. V. Pappu, Crowder titrations enable the quantification of driving forces for macromolecular phase separation. *Biophys. J.* **123**, 1376–1392 (2024).
35. R. V. Pappu, S. R. Cohen, F. Dar, M. Farag, M. Kar, Phase transitions of associative biomacromolecules. *Chem. Rev.* **123**, 8945–8987 (2023).
36. M. T. Wei *et al*., Nucleated transcriptional condensates amplify gene expression. *Nat. Cell Biol.* **22**, 1187–1196 (2020).
37. M. Kar *et al*., Phase-separating RNA-binding proteins form heterogeneous distributions of clusters in subsaturated solutions. *Proc. Natl. Acad. Sci. U.S.A.* **119**, 1–12 (2022).
38. G. Krainer *et al*., Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun.* **12**, 1085 (2021).
39. J. Guillén-Boixet *et al*., RNA-induced conformational switching and clustering of G3BP drive stress granule assembly by condensation. *Cell* **181**, 346–361 (2020).
40. P. Yang *et al*., G3BP1 is a tunable switch that triggers phase separation to assemble stress granules. *Cell* **181**, 325–345.e28 (2020).
41. N. A. Erkamp *et al*., Spatially non-uniform condensates emerge from dynamically arrested phase separation. *Nat. Commun.* **14**, 684 (2023).
42. W. E. Arter *et al*., Biomolecular condensate phase diagrams with a combinatorial microdroplet platform. *Nat. Commun.* **13**, 7845 (2022).
43. D. W. Sanders *et al*., Competing protein–RNA interaction networks control multiphase intracellular organization. *Cell* **181**, 306–324.e28 (2020).
44. A. Patil *et al*., A disordered region controls cBAF activity via condensation and partner recruitment. *Cell* **186**, 4936–4955.e26 (2023).
45. S. Sridharan *et al*., Systematic discovery of biomolecular condensate-specific protein phosphorylation. *Nat. Chem. Biol.* **18**, 1104–1114 (2022).
46. K. M. Ruff *et al*., Sequence grammar underlying the unfolding and phase separation of globular proteins. *Mol. Cell* **82**, 3193–3208.e8 (2022).
47. D. Bracha *et al*., Mapping local and global liquid phase behavior in living cells using photo-oligomerizable seeds. *Cell* **175**, 1467–1480.e13 (2018).
48. A. R. Strom, C. P. Brangwynne, The liquid nucleome–Phase transitions in the nucleus at a glance. *J. Cell Sci.* **132**, 1–7 (2019).
49. D. S. Lee, N. S. Wingreen, C. P. Brangwynne, Chromatin mechanics dictates subdiffusion and coarsening dynamics of embedded condensates. *Nat. Phys.* **17**, 531–538 (2021).
50. T. Mittag, R. V. Pappu, A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell* **82**, 2201–2214 (2022).
51. J. Wyman, S. J. Gill, Ligand-linked phase changes in a biological system: Applications to sickle cell hemoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5239–5242 (1980).
52. K. M. Ruff, F. Dar, R. V. Pappu, Ligand effects on phase separation of multivalent macromolecules. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2017184118 (2021).
53. K. M. Ruff, F. Dar, R. V. Pappu, Polyphasic linkage and the impact of ligand binding on the regulation of biomolecular condensates. *Biophys. Rev.* **2**, 021302 (2021).
54. S. K. K. Galagedera *et al*., Polyubiquitin ligand-induced phase transitions are optimized by spacing between ubiquitin units. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2306638120 (2023).
55. A. Z. Lin *et al*., Dynamical control enables the formation of demixed biomolecular condensates. *Nat. Commun.* **14**, 7678 (2023).
56. A. Chattaraj, M. L. Blinov, L. M. Loew, The solubility product extends the buffering concept to heterotypic biomolecular condensates. *eLife* **10**, 1–18 (2021).
57. D. Deviri, S. A. Safran, Physical theory of biological noise buffering by multicomponent phase separation. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2100099118 (2021).
58. H. R. Kilgore, R. A. Young, Learning the chemical grammar of biomolecular condensates. *Nat. Chem. Biol.* **18**, 1298–1306 (2022).