Title: Dimerization thermodynamics govern biological phase separation

Authors: J. Matthew Dubach

Affiliations:
Institute for Innovation in Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA

Lead contact:
dubach.matt@mgh.harvard.edu

Summary:
Many intrinsically disordered proteins/regions (IDR) can phase separate into a liquid state (biomolecular condensates) and/or a solid state (aggregates and amyloids). De-mixing in liquid-like phase separation carries an entropic cost that must be balanced by an energetic advantage of condensing. Yet, the mechanism of this energetic advantage remains unclear. Here, it is shown that an increase in dimerization binding interactions through liquid-like condensation can create overall energetic minimums when a system phase separates. Yet, the dimer binding interactions must be rapid explaining why IDRs are so prominent in phase separations. Numerous experimental examples that can only be described through dimer interactions support the dimer model presented here. The transition from condensates to aggregates and amyloids under certain conditions also appears to be driven by dimer abundance. Overall, the dimer model provides a novel theory that solves the thermodynamic mechanism of protein phase separation into biomolecular condensates, aggregates and amyloids.

Keywords:
Phase separation; biomolecular condensates; protein aggregation

Introduction:
Intrinsic disorder is abundant in the proteome[1]. While IDR function remains less characterized than structured proteins[2], the ability of IDRs to phase separate within the cell[3-5] has stoked interest into why phase separation arises and how it impacts cellular function[6,7]. Biomolecular condensates are liquid-like phase separations that flow, merge and dissolve[8], rapidly exchange components with the surrounding environment[9], are dependent on environmental conditions[10], and can incorporate oligonucleotides[11] or multiple other proteins[12]. Aggregates and amyloids are solid structures that often cannot be solubilized and are thought to arise from protein misfolding[13]. Intriguingly, numerous IDR proteins that aggregate in neurodegenerative disease have also been found to form biomolecule condensates under certain conditions[14-16], suggesting that liquid and solid phase separation may have a common energetic origin.

Thermodynamically, phase separation produces a lower entropy of mixing, therefore, since separation is spontaneous, energetic advantages that overcome this entropic cost are required. The reversibility of biomolecular condensates presents a seemingly graspable thermodynamic phenomenon whereas the expectation that aggregation is driven by protein misfolding[13] suggests more complex protein folding thermodynamics are involved. Initial considerations of biomolecular condensate thermodynamics modeled their formation as pure liquid-liquid phase separation (LLPS)[8]. But aspects of condensates do not fit canonical LLPS models[17,18], therefore properties of polymer physics, such as Flory-Huggins interaction energies[19,20], have been used to describe the energetic landscape of condensates. Discovery that the abundance of polar amino acids found in IDRs are drivers of condensate formation[21,22] suggested that Flory-Huggins was a good model for condensation. Yet, observations that altered order of amino acids[23] and replacement of charged amino acids[24] prevents condensate formation led to the "stickers and spacers" model where LLPS, multivalent interactions[25], and percolation govern condensation[21,26]. Sequence dependency on dense phase concentration and material properties of condensates[27,28] as well as loss of just a single sticker impacting the concentration at which condensates arise[29] suggest that binding interactions are essential. However, in the sticker and spacer model, why condensates spontaneously form, the impact of the complex biological environment[18] and why single amino acid mutations[24] disrupt condensate formation remain unanswered. Furthermore, how IDRs can form both condensates and amyloids remains largely unexplained[30]. Tau amyloids can be highly structured[31], suggesting that binding interactions in solid phase separation can be very specific and not complex entanglements.

Recent results show that IDR proteins can bind to form dimer pairs with high affinity and rapid dynamics[32-34]. However, the rapid dynamics of IDR protein interactions prevent *de novo* discovery in typical protein interaction pulldown screens. Thus, IDR binding interactions need to be explicitly measured, which likely limits recognition of the prevalence of these interactions in biology. Therefore, here, the energetic impact of dimerization is considered as a driver of phase separation (the dimer model).

Results:

Dimerization thermodynamics predict observed phase separation

Similar to disordered proteins and peptides, low molecular weight compounds can undergo liquid-like phase separation[35]. One molecule, diphenylalanine capped with a methoxy group (FFome), undergoes phase separation with a pH dependency[36]. At low pH the amine is protonated and the dipeptide is soluble up to 50 mM, remaining diffuse (Fig. S1A). At 25 mM FFome is diffuse at pH 3.5, forms liquid-like phase separation at pH 9 and forms aggregate crystals at pH 13 (**Fig. 1A**). The formation of condensates at pH 9 is also dependent on concentration with a decreasing abundance at 10 mM, and, surprisingly, less abundance at 45 mM (**Fig. 1B**). FFome phase separation is very dynamic as well. A rapid and large increase in pH transitions the system from diffuse to formation of condensates to crystal structures within seconds (Fig. S1B). These dynamics are highlighted by local addition of base into a diffuse solution (**Fig. 1C**). Locally raising pH by adding 2 µl of 10 M NaOH to 100 µl 50 mM FFome at pH 2 generates condensates at the site of addition, yet over time condensates dissolve as the NaOH diffuses away and mixes with the rest of the solution, producing a solution pH < 3. However, some crystal aggregates are formed and remain in solution demonstrating that transient deprotonated FFome formed a small fraction of irreversible structures similar to those observed at stable high pH.

Why FFome phase separates in a pH dependent manner is unknown. The small size of the dipeptide suggests that simple binding interactions must be drivers of phase separation. Since all molecules have an intermolecular force of interacting[37], FFome in solution has a binding affinity to form a homo-dimer, which can be expressed as a dissociation constant (K_d). Thermodynamically, the K_d is a measure of the decrease in Gibbs free energy of the system that arises through binding. The steady-state fraction bound for any interacting molecules can be plotted as a function of the total concentration to K_d ratio ([protein]/K_d), where higher ratios generate more binding (**Fig. 2A**). Critically, since the bound fraction is a steady-state value it applies to interactions with any binding dynamics. It is expected that protonated, positively charged FFome has a lower dimer binding affinity than deprotonated FFome. Likewise, the abundance of FFome impacts how much exists in a transient dimer state. Therefore, FFome concentration and protonation state, two aspects that govern condensation experimentally, will impact the dimerization fraction as well.

Liquid-like phase separation generates a high concentration dense phase and a low concentration dilute phase, which is described by the partition coefficient ([dense]/[dilute]). Through a higher concentration the dense phase will have a higher abundance of dimerization than the dilute phase (**Fig. 2A).** Mass balance calculations (supplemental information) of total steady-state fraction bound as dimers in a system at a [protein]/K_d ratio of 0.01 reveals that phase separation produces an overall higher fraction bound (**Fig. 2B**) - the amount of binding is dependent on the partition coefficient and volume fraction of the dense phase. Increased dimer abundance in the presence of condensates generates a lower effective dissociation constant, which can be used to calculate the impact on the Gibbs free energy of binding (**Fig. 2C** & supplemental information). Here, the system has a boundary condition that limits the total concentration possible in the dense phase (supplementary information).

However, the entropic cost of phase separation generates an increase in system free energy that is also dependent on the volume fraction and partition coefficient (**Fig. 2D** & supplemental information). Yet, this entropic cost is shifted from the Gibbs free energy of binding in partition coefficient vs. volume fraction space. Summing the two energies reveals that liquid-like phase separation in system with a [protein]/K_d ratio of 0.01 creates an energetic minimum state (**Fig. 2E**). Here, the energetic advantage of phase separation achieved by increased dimer abundance overcomes the entropic cost of de-mixing at specific partition coefficients and volume fractions. Calculating the minimum achievable energy through condensation as a function of [protein]/K_d ratio demonstrates that phase separation is only energetically favorable within a specific range of [protein]/K_d ratios (**Fig. 2F**). At low and high [protein]/K_d ratios the reduction in energy through increased total dimer abundance upon phase separation does not overcome the entropic cost at any volume fraction and partition coefficient. This limited range suggests that 25 mM FFome at pH 3.5 is below the phase separating [protein]/K_d ratio, while 25 mM FFome at pH 13 is above the liquid-like phase separating ratio. Overall, these calculations suggest that condensation can be driven by simple dimer binding interactions.

Hetero-interacting peptides behave in a manner predicted by the dimer model

Hetero-interacting molecules can also undergo phase separation. Using previously demonstrated[38] simple mixtures of polylysine (polyK) and polyglutamic acid (polyE), phase separation was confirmed to be dependent on the peptide length (**Fig. 3A**). Equimolar mixtures of 50mer peptides formed aggregates at the concentrations where 30mer and 20mer peptides form condensates, while 4mer peptides showed no phase separation. The 20mer and 30mer peptide condensates exhibited both merging and spontaneous disappearance (Fig. S4A). To determine if aggregation of the 50mer peptides arose from excess mass, lower concentrations of equimolar mixtures of 50mer polyK and polyE were generated, but these also produced aggregates (**Fig. 3B**). However, mixtures of equimolar E_{30} and E_{20} with K_{50} formed condensates at 2 mM (**Fig. 3C**). Thus, despite the same overall abundance of lysine and glutamic acid as the 50mer mixture at 2 mM, simply cutting one peptide into two sections transitioned the mixture from solid aggregates to liquid-like condensates. Charge neutral mixtures of 20mer or 30mer polyE with K_{50} also formed condensates at 2 mM K_{50} (**Fig. 3C**). Yet, charge neutral mixtures of E_4 with K_{50} failed to form any phase separation at 20 mM K_{50} (Fig. S4B). Aggregates and not condensates formed at higher concentrations of equimolar E_{30}, E_{20} and K_{50} mixtures (**Fig. 3C**). Aggregates also formed at higher concentrations of charge neutral mixtures of E_{30} and K_{50} and E_{20} and K_{50} (**Fig. 3C**).

At charge neutrality the only difference between E_{30} and/or E_{20} mixed with K_{50} compared to E_{50} mixed with K_{50} is the length of the polyE, suggesting that length and not capacity to form multivalent interactions is the critical factor dictating phase separation. PolyK and polyE in solution can bind to form dimers with binding affinities dependent on the length of the peptide[39], thus 50mer peptides will dimerize with a higher affinity than 30mer or 20mer peptides. Therefore, mixtures of E_{30} and E_{20} with K_{50} will have a lower dimer abundance than mixtures of E_{50} with K_{50}. The condensation dependency on both length and concentration suggests that the dimer abundance of peptide mixtures governs the ability to form liquid-like condensates in the same manner as FFome interactions.

Similar to homo-binding interactions, energy calculations of hetero-interactions predict a [protein]/K_d ratio range where condensates are energetically favorable (**Fig. 3D**, Fig. S5 & supplemental information). Furthermore, the predicted volume fractions that achieve the energetic minimum are calculated to remain below 1%. These findings are in stark contrast to existing theories that predict condensates will increase in size with increasing concentration until the dense phase becomes system spanning. Prior research on the concentration dependence of condensate formation commonly describes the concentration above which condensates appear (C_{sat})[24] but does not establish a system spanning dense phase. Fortunately, the concentration dependence of condensate dissolution, as predicted here in

the dimer model, is readily testable through increasing the protein concentration of a system with stable condensates. Thus, equimolar mixtures of K_{20} and E_{20} peptides in 350 mM NaCl were created at increasing peptide concentrations. In this system, condensation demonstrated a C_{sat} at 1 mM of each peptide, but also dissolution at 15 mM of each peptide (**Fig. 3E**), before system spanning dense phase predicted by the existing model.

Polyion peptide phase separation is driven by dimerization

Salt can induce condensate reentrance for phase separating peptides with charged amino acids[40], suggesting that the abundance of ions impacts the energy of condensate formation. Intriguingly, dimerization of two oppositely charged polyions primarily arises from the gain in entropy of counterion release and can be disrupted by high counterion concentrations that lower the energetic advantage of binding[39]. Therefore, if binding is the driver of condensate formation, the presence of counterions will disrupt the energetic advantage of condensation. To test if counterion concentration impacts condensate formation, equimolar mixtures of K_{20} and E_{20} peptides at various peptide and salt concentrations were created and the presence of condensation was determined through DIC imaging. An automated image analysis algorithm quantifying the non-background pixels was used to identify the presence of condensates (Fig. S4C). A plot of condensation as a function of salt and peptide concentration demonstrates that condensates reenter solution with a similar energy pattern to the molar free energy calculated from the dimer model for hetero-dimer interactions (**Fig. 4A & 3D**). At higher concentrations, condensates formed upon vigorous mixing but lost homogenous condensation within 10 minutes (Fig. S6A-D). The volume fraction was also dependent on both the peptide concentration and salt concentration (**Fig. 4B**), as predicted by the dimer model. The formation of condensates as a function of salt and peptide concentration was also determined for K_{30} and E_{30} at equimolar ratios. The 30mer reentrance boundary curve (**Fig. 4C**) displayed a shift toward lower peptide concentrations and higher salt concentrations compared to the 20mer curve, suggesting a higher dimer affinity as expected. Similar results were found with binding between 1:1 mixtures of K_{10} and D_{10}, where condensates were disrupted by lower salt concentrations (**Fig. 4C**). However, for the 10mer peptides, a sharp decrease in the reentrance boundary curve at high concentrations could not be established due to limiting counterion concentrations in the peptide stock solutions. Yet, 10mer peptide condensates at the upper concentration limit generated smaller condensates (Fig. S7E), similar to observed salt concentration reentrance of 20mer peptides (**Fig. 4B**), confirming the impact of counterions at these concentrations. Critically, the correlation of peptide length with salt concentration needed to prevent condensate formation confirms that phase separation proceeds through dimer interactions. Otherwise, if phase separation arose through multivalent interactions, for example, the concentration of salt that prevents condensation at 1 mM 30mer peptides would be the same as the concentration that disrupts 1.5 mM 20mer peptides (equal total monomer concentration), but that is not the case.

The mechanism of binding-induced counterion release driving condensate formation was further confirmed in unequal molar ratios of E_{30} and K_{30} (Fig. S7A). Here, counterions from the stock solution were unbalanced and condensates were not formed. However, addition of NaCl, which provided ions for release upon binding, enabled condensate formation at these unbalanced ratios. The dimer model also predicts that the presence of one condensate component above the dissolution concentration will prevent condensate formation (Fig. S7B). This was confirmed experimentally with K_{20} and E_{20} where E_{20} was held above the dissolution concentration and lower concentrations of K_{20} were unable to generate condensates (Fig. S7C). However, when both E_{20} and K_{20} were at concentrations below the dissolution concentration, the imbalanced ratio from lowering the concentration of just K_{20} did not prevent condensate formation (Fig. S7C).

Binding affinity and dynamics govern condensation and aggregation

Seemingly, both the dimer model presented here and the existing theory of biomolecular condensates do not provide compelling mechanisms for aggregate and not condensate formation observed with 50mer peptides. Decreasing the peptide concentration reduced the size of the aggregates but did not generate condensates (**Fig. 3B**). Similarly, increasing the salt concentration also reduced the size of the aggregates, suggesting that aggregates form through polyion binding, but did not generate any condensates (Fig. S8). The NaCl concentration (850 mM) needed to prevent aggregates was higher than the concentration (500 mM) that prevented of E_{30}-K_{30} condensates at the same peptide concentration, confirming higher affinity binding between the 50mer peptides.

Here, the kinetics of binding interactions that drive condensate formation in the dimer model are considered. Proteins readily exchange between the dense and dilute phase[41-47], yet condensation is stable in steady state demonstrating that total flux is balanced. In the dimer model, condensates consist of both bound and unbound protein. In the binding model, unbound protein flux is balanced by the entropic force to leave the condensate and the energetic advantage of existing at a high condensate concentration to increase the likelihood of binding (**Fig. 4D**). However, bound protein does not have any energetic advantage to enter nor remain within the condensate and is only subject to the gain in entropy upon leaving, as well as a decrease in chemical potential at high dense phase concentrations. Since flux between the two phases is diffusion driven and the rates of diffusion for bound and unbound protein will be similar, this imbalanced force will seemingly prevent condensation. Yet, diffusion out of condensates is a dynamic process, and, over time, a bound protein both diffuses and dissociates (**Fig. 4E**). Recent results demonstrate that IDR protein binding in a condensate occurs through extremely fast dynamics, with partner exchange rates suggesting k_{off} rates greater than 10^6 1/s[48]. Measured diffusivity rates inside condensates have been reported between 10^{-3} to 10^{-1} $\mu m^2/s$[41-46]. Therefore, it is expected that molecules within condensates can exist with dynamics where bound protein dissociates before significant diffusion relative to the size of the typical condensate (**Fig. 4F**). This phenomenon enables steady state condensate formation and suggests that the dissociation rate of K_{50}-E_{50} is too slow to form condensates. Since both 50mer and shorter peptides are expected to have the same diffusion-limited association rates, the higher affinity of the longer peptides arises solely through a slower dissociation rate. Therefore, the slower dissociation rate prevents liquid-like condensate formation. Instead, dimerized 50mer polyK and polyE will be hydrophobic as the charged components are interacting with each other, thus aggregation likely forms when multiple dimers interact.

The dimer model captures protein-oligonucleotide phase separation

Condensates can also be generated from mixtures of proteins and oligonucleotides[49,50]. Mixtures of K_{20} and a 20mer single stranded DNA (DNA_{20}) produced condensates (**Fig. 5A**). The solubility limitations of DNA prevent 1:1 charge ratio at higher concentrations, thus determination of concentration induced condensate dissolution was performed at uneven molar mixtures. Dissolution was not observed at lower concentrations with the same charge ratio (**Fig. 5B**), confirming that loss of condensates at high concentration is driven by concentration and not charge ratio. Both the observed C_{sat} and dissolution concentration where higher for K_{20}-DNA_{20} than K_{20}-E_{20}, suggesting that K_{20} binds DNA_{20} with a lower affinity than E_{20}. However, mixtures of K_{20} and a 20mer RNA (RNA_{20}) were unable to generate condensates, instead producing aggregates (**Fig. 5C**). Absence of condensates was also found for 1:1 charge ratios of RNA_{20} with K_{50}, K_{30}, K_{10} and K_4 (**Fig. 5C**), with all except K_4 producing aggregates. The absence of condensates remained at lower concentrations of polyK and RNA as well (Fig. S9), suggesting that polyK binding to RNA has a much higher affinity than binding to DNA or polyE.

Condensates were formed, however, between polyK and poly(A) RNA, a long RNA polymer, (**Fig. 5D**), demonstrating that polyK can form condensates with RNA. As a large polymer, a single molecule of poly(A) can bind multiple molecules of polyK. Therefore, each poly(A) molecule is unlikely to be

saturated with polyK, leaving sections unbound, creating a more balanced force to remain in the condensate (**Fig. 4D**). The dimer model also predicts that for binding interactions with one molecule having multiple binding sites the dissolution concentration is higher than for one-to-one binding (Fig. S10A). This increased dissolution concentration arises from the lower entropic cost of RNA condensation. All polyK peptides showed loss of condensation at both high and low polyK:poly(A) concentration ratios (Fig. S10B). Existing models of condensate formation between oppositely charged multivalent molecules have predicted that charge imbalance is the driver of dissolution[51-55]. However, intriguingly, here the span of charge ratios where condensates were stable was dependent on the length of the polyK peptide (**Fig. 5E**), demonstrating that reentrance with multimeric interactions is likely due to dimer forming thresholds and not a charge imbalance. At extreme concentration imbalances the likelihood of the lower concentration component being bound in solution is much higher than at equal ratios. Therefore, the energetic advantage of phase separation versus diffuse distribution is lost at large concentration imbalances and the system remains diffuse.

Polyions can form amyloid fibrils at high dimer abundance

Experimental results with polyion peptides demonstrate that high concentrations where condensates are not present generated aggregates. For the 50mer peptides, which do not form condensates at any concentration tested, aggregation produced higher ordered structure with increasing concentration (**Fig. 6A**). And, while difficult to quantify, higher order structures appeared to be more abundant at higher concentrations for 10mer, 20mer and 30mer polyion mixtures (**Fig. 6A**). Generation of these highly ordered, amyloid-like structures at high concentrations was immediate. However, concentrations close to the condensate dissolution concentration initially formed condensates upon peptide mixing that became smaller in size over time and eventually disappeared leaving small aggregates (**Fig. 6B**). This observed reduction in condensate size over time was similar to concentration dependent condensate size (**Fig. 3E**), suggesting that the soluble fraction of peptide was reduced over time as peptide formed aggregates. Furthermore, dissolution of condensates during transition into aggregates did not directly produce aggregates but instead reentered solution (**Fig. 6C**), also suggesting that these condensates are not precursors to aggregates.

The dimer model and supporting experimental results suggest that the fraction of steady-state dimerized IDR governed by the [protein]/K_d ratio dictates phase separation (**Fig. 6D**). At low bound fractions the decrease in the free energy of binding that can be achieved through condensation overcomes the entropic cost of de-mixing and biomolecular condensates produce the lowest energetic state, provided the binding interactions are fast. However, at higher [protein]/K_d ratios the diffuse dimer abundance is also higher and the relative increase in binding within higher concentration condensates does not outweigh the entropic cost of de-mixing so condensates are not energetically favorable. Yet, the abundance of steady-state dimerized protein in solution enables a new energetic minimum when the dimerized IDR forms aggregates. Aggregation likely arises from the low entropy of solvent interacting with dimerized peptide outweighing the entropic cost of de-mixing. At even higher [protein]/K_d ratios, the increased abundance of dimerized IDR appears to enable a lower energetic minimum to be reached through more structured amyloid formation. Critically, with the peptides used here, aggregation is not a product of condensation, rather it occurs when condensates are not energetically favorable or dynamically possible, demonstrated by aggregation of the 50mer peptides (**Fig. 3A**).

Discussion:

The energetic advantage of dimerization, either homotypic or heterotypic has not been previously explored as a contributor to phase separation. However, binding interactions between highly flexible molecules are essential to condensate formation[56]. And, experimentally, condensation has been shown to be dependent on system temperature[57], IDR phosphorylation[58], and buffering of electrostatic

interactions[3,40,59], diverse properties that alter binding. The recent demonstrations that IDRs dimerize with extremely fast dynamics while maintaining a disordered state[32,33,48,60] suggests that dimerization binding interactions are more abundant than currently appreciated. Here, consideration of the energetic impact of dimerization captures the majority of literature results on liquid-like condensates. The dimer model accurately captures salt induced reentrance[40] and indicates that binding energy limitations at imbalanced molar ratios, and not charge imbalance[52], induces dissolution, a finding that may clarify transcriptional condensate function[61]. The model also describes complex coacervation[62] and the energetic forces governing condensation between oppositely charged molecules[51]. The reliance of condensate formation on very fast dissociation kinetics also explains why biomolecular condensates are nearly unique to disordered proteins.

The dimer model suggests that rather than physical crosslinks among multivalent proteins being the cause of condensate formation[63], simple dimerization induces condensates. Therefore, sequence mutations in disordered proteins that prevent condensate formation[24] do not remove stickers, but rather alter the binding affinity. However, the dimer model does not exclude the presence of multivalent interactions, when possible, within a condensate. The dimer model simply demonstrates that the energy gained from being bound to one or many molecules and the higher likelihood of being bound within a higher concentration environment drives phase separation. Indeed, multiple binding interactions that molecular dynamic simulations suggest can happen within a condensate[48] could prevent aggregation. Dimerization as a driver of condensation also demonstrates how IDRs can form aggregates and amyloids. Stacked β-sheets[64] are the core component of fibrillar structures[31,65], demonstrating that IDRs interact with precise alignment. Thus, the steady-state abundance of dimers at high [protein]/K_d ratios provides aligned precursor for amyloid stacking. And, the presence of aggregate seeds[66] provides a substrate for dimers to grow fibrils through dimer-seed interactions at a rate that will be dependent on the [protein]/K_d ratio in the system. The dimer model, however, does not require that the entire length of an IDR be bound, congruent with the observed fuzziness of tau fibrils[67].

Experimental measurements based on dimer model predictions also confirm that the dimer model accurately captures phase separation. Increasing protein or peptide concentrations well beyond C_{sat} has not been previously explored since existing models suggest condensates would persist, however the dimer model suggests otherwise and performing experiments at high peptide concentrations confirms this prediction. Several outstanding questions remain and the dimer model is certainly incomplete. The size of the condensing components will impact the condensate volume fraction, while the surface tension will impact the number and size of condensates. However, overall, the dimer model presented and validated here predicts observed phase separation that cannot be described by other theories. Viewing IDR phase separation phenomena through the lens of the dimer model suggests that protein sequence[68] and post translational modifications[58,69] impact phase separation by altering binding affinities – a vantage point that could clarify the mechanism of protein aggregation in neurodegenerative disease[70,71] as well as other biological phase separation.

Acknowledgements: Many thanks to Mariane Le Fur for technical guidance and Eric Gale and Alykhan Shamji for thoughts on the manuscript.

Funding: This work was supported by a grant from the National Institutes of Health (R01CA241179) and utilized equipment funded by the Massachusetts Life Science Center.

Competing interests: No conflicts of interest are declared.

Methods

Dimer model calculations.

All calculations were performed in MATLAB. MATLAB code is available at https://github.com/dubachLab/phaseSeparation. Plots were generated in Graphpad Prism.

Condensate formation.

Unless otherwise noted all chemicals were obtained from Sigma Aldrich. FFome was obtained from Ambeed. Peptides were obtained from Alamanda Polymers or Genscript (4mer peptides). DNA (ACGTATCCGAATGCCAGTTG) and RNA (ACGUAUCCGAAUGCCAGUUG) were obtained from IDT. FFome was dissolved in 5 mM HEPES buffer at pH<4 and 500 mM KCl. pH was controlled by addition of HCl and NaOH or KOH. All peptides and oligonucleotides were reconstituted in 10 mM TRIS buffer, pH 8.0. pH was adjusted using NH_3OH or HCl. Phase separation was achieved through mixtures of stock solutions to generate the desired concentrations and imaged in 96 well glass bottom well plates (CellVis) on a Nikon Ti2 widefield microscope in differential interference contrast (DIC) with a 60x NA 1.49 oil immersion or 20x NA 0.75 objective.

Data and materials availability: All data are available in the main text or the supplementary figures. All materials are commercially available.

References

1 Oldfield, C. J. *et al.* Comparing and combining predictors of mostly disordered proteins. *Biochemistry* **44**, 1989-2000 (2005). https://doi.org:10.1021/bi047993o

2 Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **6**, 197-208 (2005). https://doi.org:10.1038/nrm1589

3 Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123-133 (2015). https://doi.org:10.1016/j.cell.2015.09.015

4 Tompa, P. Structural disorder in amyloid fibrils: its implication in dynamic interactions of proteins. *Febs j* **276**, 5406-5415 (2009). https://doi.org:10.1111/j.1742-4658.2009.07250.x

5 Fuxreiter, M. & Vendruscolo, M. Generic nature of the condensed states of proteins. *Nature Cell Biology* **23**, 587-594 (2021). https://doi.org:10.1038/s41556-021-00697-8

6 Banani, S. F. *et al.* Genetic variation associated with condensate dysregulation in disease. *Dev Cell* **57**, 1776-1788 e1778 (2022). https://doi.org:10.1016/j.devcel.2022.06.010

7 Selkoe, D. J. Cell biology of protein misfolding: the examples of Alzheimer's and Parkinson's diseases. *Nat Cell Biol* **6**, 1054-1061 (2004). https://doi.org:10.1038/ncb1104-1054

8 Brangwynne, C. P. *et al.* Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-1732 (2009). https://doi.org:10.1126/science.1172046

9 Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018). https://doi.org:doi:10.1126/science.aar2555

10 Nott, Timothy J. *et al.* Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell* **57**, 936-947 (2015). https://doi.org:https://doi.org/10.1016/j.molcel.2015.01.013

11 Maharana, S. *et al.* RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360**, 918-921 (2018). https://doi.org:10.1126/science.aar7366

12 Xing, W., Muhlrad, D., Parker, R. & Rosen, M. K. A quantitative inventory of yeast P body proteins reveals principles of composition and specificity. *eLife* **9**, e56525 (2020). https://doi.org:10.7554/eLife.56525

13 Soto, C. & Pritzkow, S. Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nature Neuroscience* **21**, 1332-1340 (2018). https://doi.org:10.1038/s41593-018-0235-9

14 Wegmann, S. *et al.* Tau protein liquid–liquid phase separation can initiate tau aggregation. *The EMBO Journal* **37**, e98049 (2018). https://doi.org:https://doi.org/10.15252/embj.201798049

15 Qamar, S. *et al.* FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. *Cell* **173**, 720-734.e715 (2018). https://doi.org:10.1016/j.cell.2018.03.056
16 Zambusi, A. *et al.* TDP-43 condensates and lipid droplets regulate the reactivity of microglia and regeneration after traumatic brain injury. *Nature Neuroscience* **25**, 1608-1625 (2022). https://doi.org:10.1038/s41593-022-01199-y
17 McSwiggen, D. T., Mir, M., Darzacq, X. & Tjian, R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev* **33**, 1619-1634 (2019). https://doi.org:10.1101/gad.331520.119
18 Mittag, T. & Pappu, R. V. A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol Cell* **82**, 2201-2214 (2022). https://doi.org:10.1016/j.molcel.2022.05.018
19 Flory, P. J. Thermodynamics of high polymer solutions. *The Journal of chemical physics* **10**, 51-61 (1942).
20 Huggins, M. L. Some properties of solutions of long-chain compounds. *The Journal of Physical Chemistry* **46**, 151-158 (1942).
21 Brangwynne, Clifford P., Tompa, P. & Pappu, Rohit V. Polymer physics of intracellular phase transitions. *Nature Physics* **11**, 899-904 (2015). https://doi.org:10.1038/nphys3532
22 Das, R. K., Ruff, K. M. & Pappu, R. V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr Opin Struct Biol* **32**, 102-112 (2015). https://doi.org:10.1016/j.sbi.2015.03.008
23 Martin, E. W. *et al.* Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694-699 (2020). https://doi.org:10.1126/science.aaw8653
24 Wang, J. *et al.* A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **174**, 688-699 e616 (2018). https://doi.org:10.1016/j.cell.2018.06.006
25 Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340 (2012). https://doi.org:10.1038/nature10879
26 Harmon, T. S., Holehouse, A. S., Rosen, M. K. & Pappu, R. V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife* **6** (2017). https://doi.org:10.7554/eLife.30294
27 Wei, M. T. *et al.* Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat Chem* **9**, 1118-1125 (2017). https://doi.org:10.1038/nchem.2803
28 Alberti, S., Gladfelter, A. & Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **176**, 419-434 (2019). https://doi.org:10.1016/j.cell.2018.12.035
29 Choi, J.-M., Holehouse, A. S. & Pappu, R. V. Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annual Review of Biophysics* **49**, 107-133 (2020). https://doi.org:10.1146/annurev-biophys-121219-081629
30 Michaels, T. C. T. *et al.* Amyloid formation as a protein phase transition. *Nature Reviews Physics* **5**, 379-397 (2023). https://doi.org:10.1038/s42254-023-00598-9
31 Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185-190 (2017). https://doi.org:10.1038/nature23002
32 Borgia, A. *et al.* Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61-66 (2018). https://doi.org:10.1038/nature25762
33 Sottini, A. *et al.* Polyelectrolyte interactions enable rapid association and dissociation in high-affinity disordered protein complexes. *Nat Commun* **11**, 5736 (2020). https://doi.org:10.1038/s41467-020-18859-x
34 Stadmiller, S. S., Aguilar, J. S., Waudby, C. A. & Pielak, G. J. Rapid Quantification of Protein-Ligand Binding via 19F NMR Lineshape Analysis. *Biophysical Journal* **118**, 2537-2548 (2020). https://doi.org:https://doi.org/10.1016/j.bpj.2020.03.031
35 Abbas, M., Lipiński, W. P., Nakashima, K. K., Huck, W. T. S. & Spruijt, E. A short peptide synthon for liquid-liquid phase separation. *Nat Chem* **13**, 1046-1054 (2021). https://doi.org:10.1038/s41557-021-00788-x

36 Cao, S. *et al.* Dipeptide coacervates as artificial membraneless organelles for bioorthogonal catalysis. *Nature Communications* **15**, 39 (2024). https://doi.org:10.1038/s41467-023-44278-9
37 Hunter, C. A. Quantifying Intermolecular Interactions: Guidelines for the Molecular Recognition Toolbox. *Angewandte Chemie International Edition* **43**, 5310-5324 (2004). https://doi.org:https://doi.org/10.1002/anie.200301739
38 Cakmak, F. P., Choi, S., Meyer, M. O., Bevilacqua, P. C. & Keating, C. D. Prebiotically-relevant low polyion multivalency can improve functionality of membraneless compartments. *Nature Communications* **11**, 5949 (2020). https://doi.org:10.1038/s41467-020-19775-w
39 Ou, Z. & Muthukumar, M. Entropy and enthalpy of polyelectrolyte complexation: Langevin dynamics simulations. *The Journal of Chemical Physics* **124** (2006). https://doi.org:10.1063/1.2178803
40 Krainer, G. *et al.* Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nature Communications* **12**, 1085 (2021). https://doi.org:10.1038/s41467-021-21181-9
41 Burke, Kathleen A., Janke, Abigail M., Rhine, Christy L. & Fawzi, Nicolas L. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Molecular Cell* **60**, 231-241 (2015). https://doi.org:https://doi.org/10.1016/j.molcel.2015.09.006
42 Elbaum-Garfinkle, S. *et al.* The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proceedings of the National Academy of Sciences* **112**, 7189-7194 (2015). https://doi.org:doi:10.1073/pnas.1504822112
43 Zhang, H. *et al.* RNA Controls PolyQ Protein Phase Transitions. *Molecular Cell* **60**, 220-230 (2015). https://doi.org:https://doi.org/10.1016/j.molcel.2015.09.017
44 Gopal, P. P., Nirschl, J. J., Klinman, E. & Holzbaur, E. L. F. Amyotrophic lateral sclerosis-linked mutations increase the viscosity of liquid-like TDP-43 RNP granules in neurons. *Proceedings of the National Academy of Sciences* **114**, E2466-E2475 (2017). https://doi.org:doi:10.1073/pnas.1614462114
45 Taylor, N. O., Wei, M.-T., Stone, H. A. & Brangwynne, C. P. Quantifying Dynamics in Phase-Separated Condensates Using Fluorescence Recovery after Photobleaching. *Biophysical Journal* **117**, 1285-1300 (2019). https://doi.org:https://doi.org/10.1016/j.bpj.2019.08.030
46 Folkmann, A. W., Putnam, A., Lee, C. F. & Seydoux, G. Regulation of biomolecular condensates by interfacial protein clusters. *Science* **373**, 1218-1224 (2021). https://doi.org:doi:10.1126/science.abg7071
47 Miné-Hattab, J. *et al.* Single molecule microscopy reveals key physical features of repair foci in living cells. *eLife* **10**, e60577 (2021). https://doi.org:10.7554/eLife.60577
48 Galvanetto, N. *et al.* Extreme dynamics in a biomolecular condensate. *Nature* (2023). https://doi.org:10.1038/s41586-023-06329-5
49 Roden, C. & Gladfelter, A. S. RNA contributions to the form and function of biomolecular condensates. *Nature Reviews Molecular Cell Biology* **22**, 183-195 (2021). https://doi.org:10.1038/s41580-020-0264-6
50 Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357** (2017). https://doi.org:10.1126/science.aaf4382
51 Besteman, K., Van Eijk, K. & Lemay, S. G. Charge inversion accompanies DNA condensation by multivalent ions. *Nature Physics* **3**, 641-644 (2007). https://doi.org:10.1038/nphys697
52 Milin, A. N. & Deniz, A. A. Reentrant Phase Transitions and Non-Equilibrium Dynamics in Membraneless Organelles. *Biochemistry* **57**, 2470-2477 (2018). https://doi.org:10.1021/acs.biochem.8b00001
53 Zhang, F. *et al.* Reentrant Condensation of Proteins in Solution Induced by Multivalent Counterions. *Physical Review Letters* **101**, 148101 (2008). https://doi.org:10.1103/PhysRevLett.101.148101
54 Nguyen, T. T., Rouzina, I. & Shklovskii, B. I. Reentrant condensation of DNA induced by multivalent counterions. *The Journal of Chemical Physics* **112**, 2562-2568 (2000). https://doi.org:10.1063/1.480819
55 Banerjee, P. R., Milin, A. N., Moosa, M. M., Onuchic, P. L. & Deniz, A. A. Reentrant Phase Transition Drives Dynamic Substructure Formation in Ribonucleoprotein Droplets. *Angew Chem Int Ed Engl* **56**, 11354-11359 (2017). https://doi.org:10.1002/anie.201703191
56 Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285-298 (2017). https://doi.org:10.1038/nrm.2017.7
57 Fritsch, A. W. *et al.* Local thermodynamics govern formation and dissolution of Caenorhabditis elegans P granule condensates. *Proceedings of the National Academy of Sciences* **118**, e2102772118 (2021). https://doi.org:doi:10.1073/pnas.2102772118

58 Sridharan, S. *et al.* Systematic discovery of biomolecular condensate-specific protein phosphorylation. *Nature Chemical Biology* **18**, 1104-1114 (2022). https://doi.org:10.1038/s41589-022-01062-y
59 Bremer, A. *et al.* Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. *Nat Chem* **14**, 196-207 (2022). https://doi.org:10.1038/s41557-021-00840-w
60 Heidarsson, P. O. *et al.* Release of linker histone from the nucleosome driven by polyelectrolyte competition with a disordered protein. *Nat Chem* **14**, 224-231 (2022). https://doi.org:10.1038/s41557-021-00839-3
61 Henninger, J. E. *et al.* RNA-Mediated Feedback Control of Transcriptional Condensates. *Cell* **184**, 207-225.e224 (2021). https://doi.org:10.1016/j.cell.2020.11.030
62 Muthukumar, M. 50th Anniversary Perspective: A Perspective on Polyelectrolyte Solutions. *Macromolecules* **50**, 9528-9560 (2017). https://doi.org:10.1021/acs.macromol.7b01929
63 Choi, J.-M., Dar, F. & Pappu, R. V. LASSI: A lattice model for simulating phase transitions of multivalent proteins. *PLOS Computational Biology* **15**, e1007028 (2019). https://doi.org:10.1371/journal.pcbi.1007028
64 Daebel, V. *et al.* β-Sheet Core of Tau Paired Helical Filaments Revealed by Solid-State NMR. *Journal of the American Chemical Society* **134**, 13982-13989 (2012). https://doi.org:10.1021/ja305470p
65 Gremer, L. *et al.* Fibril structure of amyloid-β(1–42) by cryo–electron microscopy. *Science* **358**, 116-119 (2017). https://doi.org:doi:10.1126/science.aao2825
66 Goedert, M. Alzheimer's and Parkinson's diseases: The prion concept in relation to assembled Aβ, tau, and α-synuclein. *Science* **349**, 1255555 (2015). https://doi.org:doi:10.1126/science.1255555
67 Wegmann, S., Medalsy, I. D., Mandelkow, E. & Müller, D. J. The fuzzy coat of pathological human Tau fibrils is a two-layered polyelectrolyte brush. *Proceedings of the National Academy of Sciences* **110**, E313-E321 (2013). https://doi.org:doi:10.1073/pnas.1212100110
68 Morin, J. A. *et al.* Sequence-dependent surface condensation of a pioneer transcription factor on DNA. *Nature Physics* **18**, 271-276 (2022). https://doi.org:10.1038/s41567-021-01462-2
69 Hofweber, M. & Dormann, D. Friend or foe-Post-translational modifications as regulators of phase separation and RNP granule dynamics. *J Biol Chem* **294**, 7137-7150 (2019). https://doi.org:10.1074/jbc.TM118.001189
70 Buratti, E. TDP-43 post-translational modifications in health and disease. *Expert Opin Ther Targets* **22**, 279-293 (2018). https://doi.org:10.1080/14728222.2018.1439923
71 Despres, C. *et al.* Identification of the Tau phosphorylation pattern that drives its aggregation. *Proceedings of the National Academy of Sciences* **114**, 9080-9085 (2017). https://doi.org:doi:10.1073/pnas.1708448114
72 Jarmoskaite, I., AlSadhan, I., Vaidyanathan, P. P. & Herschlag, D. How to measure and evaluate binding affinities. *eLife* **9**, e57264 (2020). https://doi.org:10.7554/eLife.57264
73 Prausnitz, J. & Foose, L. Three frontiers in the thermodynamics of protein solutions. *Pure Appl. Chem.* **79**, 1435-1444 (2007). https://doi.org:doi:10.1351/pac200779081435
74 Bomont, J. M. A consistent calculation of the chemical potential for dense simple fluids. *J Chem Phys* **124**, 206101 (2006). https://doi.org:10.1063/1.2198807

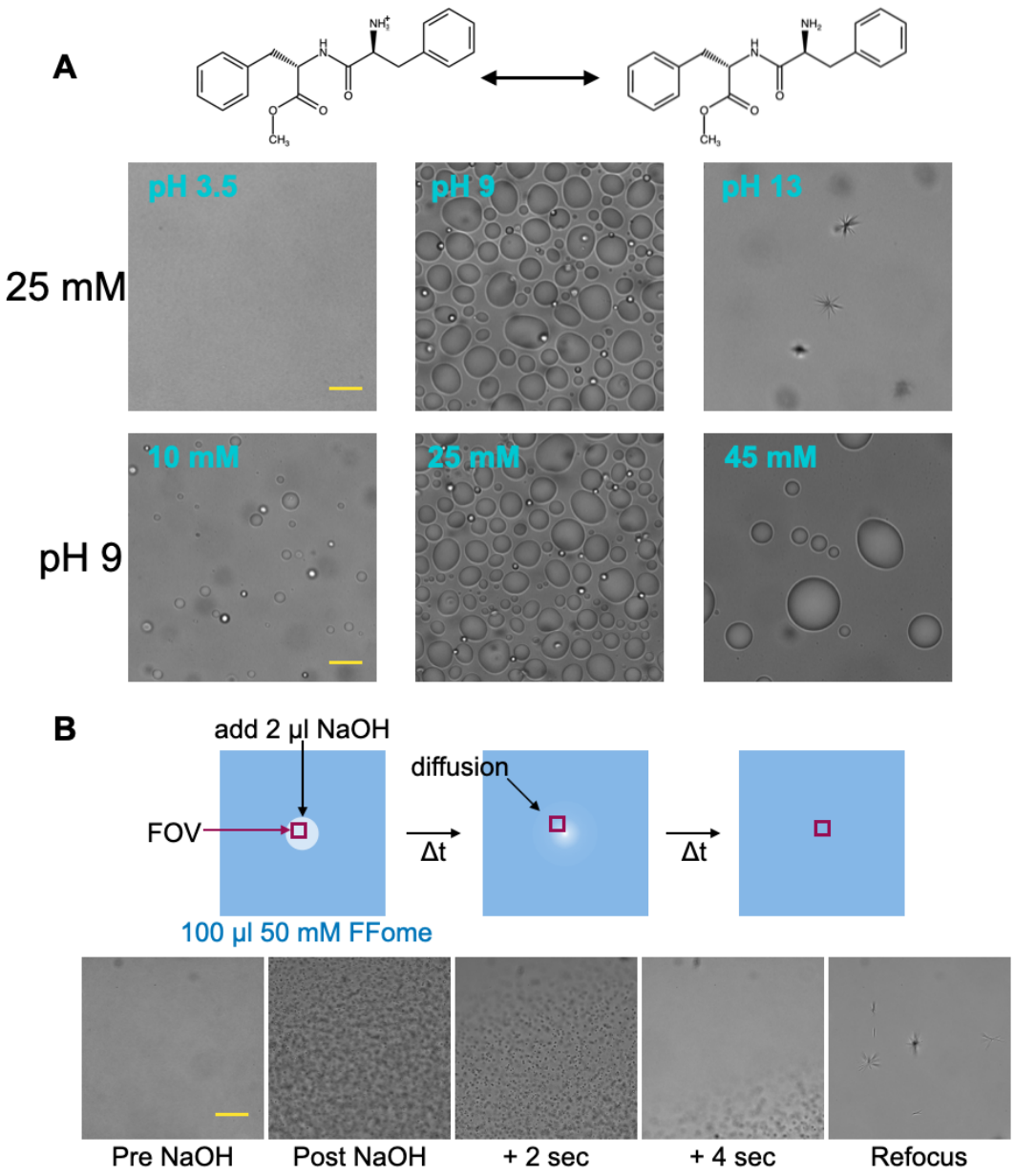

Figure 1. Phase separation of FFome (**A**) FFome will be protonated in a pH-dependent manner. Representative DIC images of FFome in solution at specific pH and concentrations, scale bars = 30 µm. (**B**) Schematic for local addition of high concentration NaOH in a solution of FFome. Representative DIC images of FFome before NaOH addition and post addition and multiple time points as well as refocusing in the sample after condensation dissolved, scale bar = 40 µm.

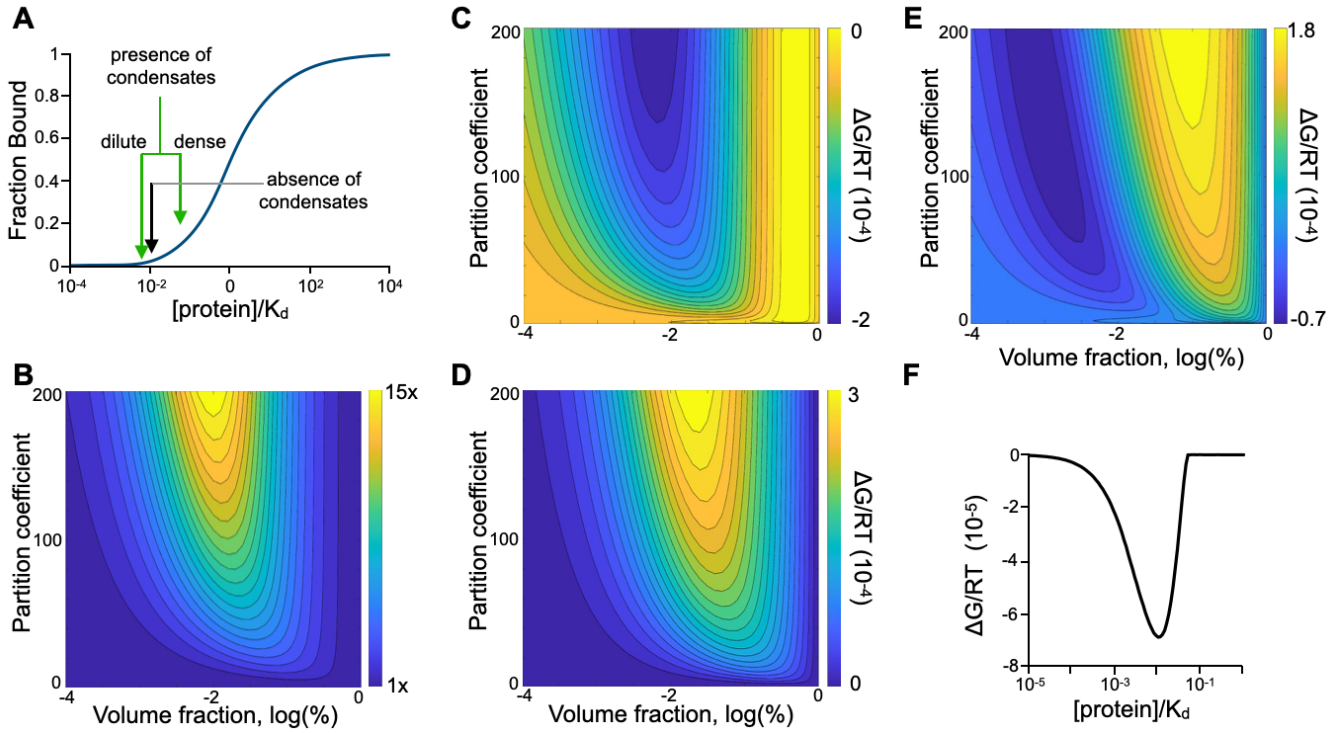

Figure 2. The dimer model of phase separation predicts system energy minimums upon liquid-like condensation within a limited [protein]/K_d ratio range. (**A**) The steady state fraction bound as a function of the [protein]/K_d ratio. Condensation generates two phases that reside on the fraction bound curve. (**B**) Calculations for total homo-dimerization fraction bound upon liquid-like phase separation as a function of dense phase volume fraction and partition coefficient. (**C**) Calculations for the change in the Gibbs free energy of binding upon liquid-like phase separation as a function of dense phase volume fraction and partition coefficient. (**D**) Calculations for the change in the entropy of de-mixing upon liquid-like phase separation as a function of dense phase volume fraction and partition coefficient. (**E**) Calculations for the change in total system energy upon liquid-like phase separation as a function of dense phase volume fraction and partition coefficient. (**F**) Calculations for the minimum change in total energy that can be achieved through liquid-like phase separation as a function of the [protein]/K_d ratio of homo-dimerization. B-F, K_d = 10 mM.

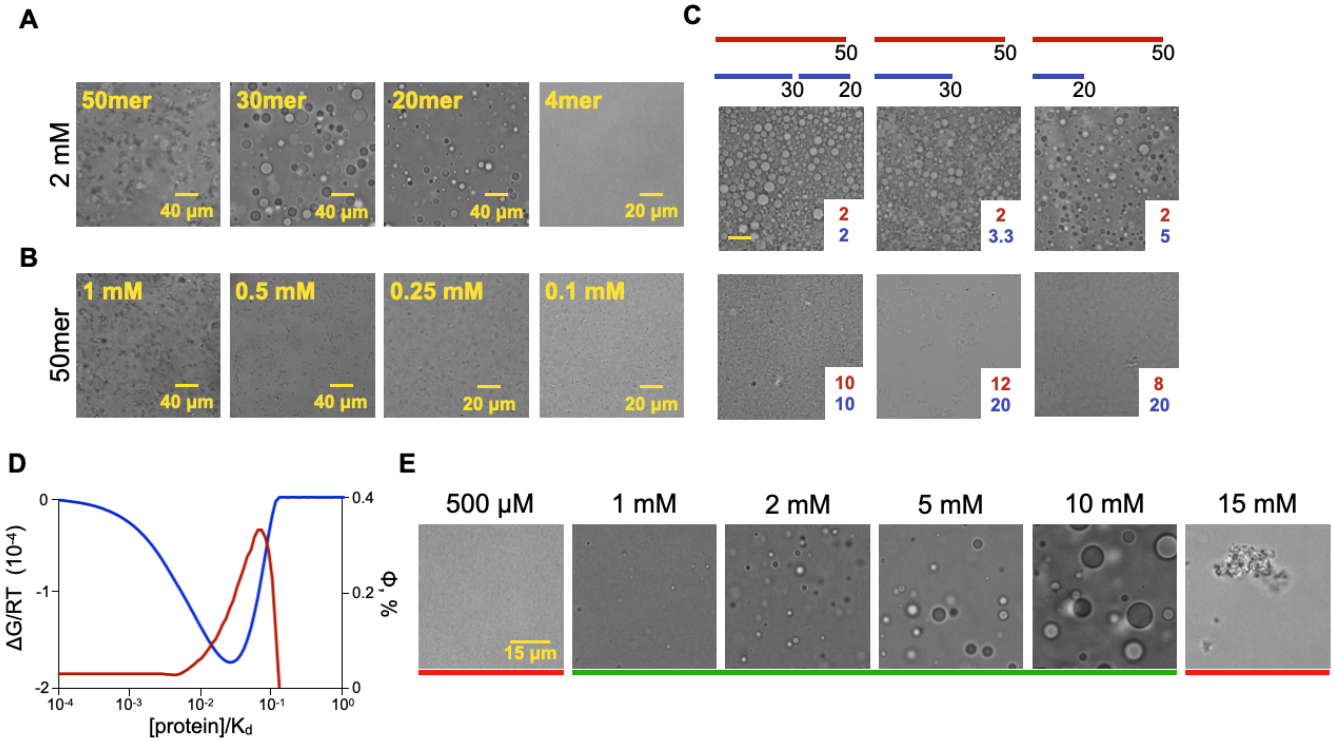

Figure 3. Polyion condensate formation follows the dimer model predictions. (**A**) Representative DIC images of equimolar polyK and polyE mixtures at 2 mM. (**B**) Representative DIC images of equimolar K_{50} and E_{50} mixtures at multiple concentrations. (**C**) Representative DIC images K_{50} (red) and polyE (blue) mixtures at equal charge ratios, shown overlaying each image is the concentration of K_{50} (red) and polyE (blue) in mM. (**D**) Calculated energy minimum for condensate formation (blue) and volume fraction (red) at the minimum energy as a function of protein concentration (single protein) for an equimolar mixture of two binding proteins, K_d = 10 mM. **(E)** Representative DIC images of K_{20}-E_{20} peptide condensates at equimolar concentrations in the presence of 350 mM NaCl, with red and green indicating the absence or presence of condensates, respectively.

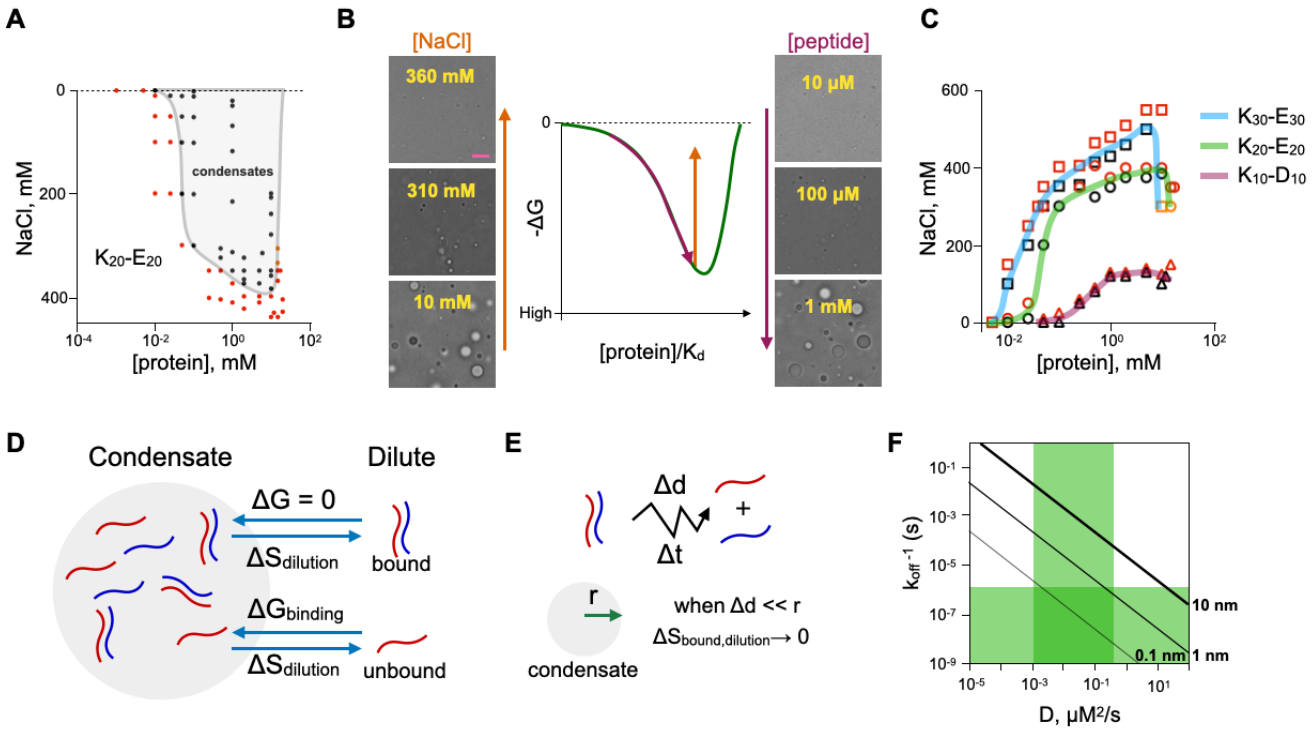

Figure 4. Condensate salt disruption and binding dynamics in the dimer model. (**A**) Formation (black) or absence (red) of K_{20}-E_{20} condensates at varying NaCl concentrations and equimolar peptide concentrations. A trace between the presence and absence of condensates highlights condensate forming conditions. Orange data points are condensates that stabilize into aggregates. (**B**) Representative DIC images of K_{20}-E_{20} peptide condensates as a function of NaCl (orange, 1 mM peptide) and peptide (purple, [NaCl] = 20x[peptide]) concentration, scale bar = 10 μm. (**C**) Presence (black) and absence (red) of condensates for K_{20}-E_{20} (circles – green trace), K_{30}-E_{30} (squares – blue trace) and K_{10}-D_{10} (triangles – purple trace) with traces of the boundary for condensate reentrance as a function of [NaCl] and equimolar peptide concentration. (**D**) The energetic forces impacting bound and unbound molecules that form condensates. **(E)** The dissociation rate and diffusivity govern the diffusion distance of a bound molecule. (**F**) Reported (green) dissociation rates of intrinsically disordered proteins in condensates and diffusivity measured in condensates. Isolines (black), determined by Fick's law of diffusion, are plotted for average diffusion distance of bound protein before dissociation.

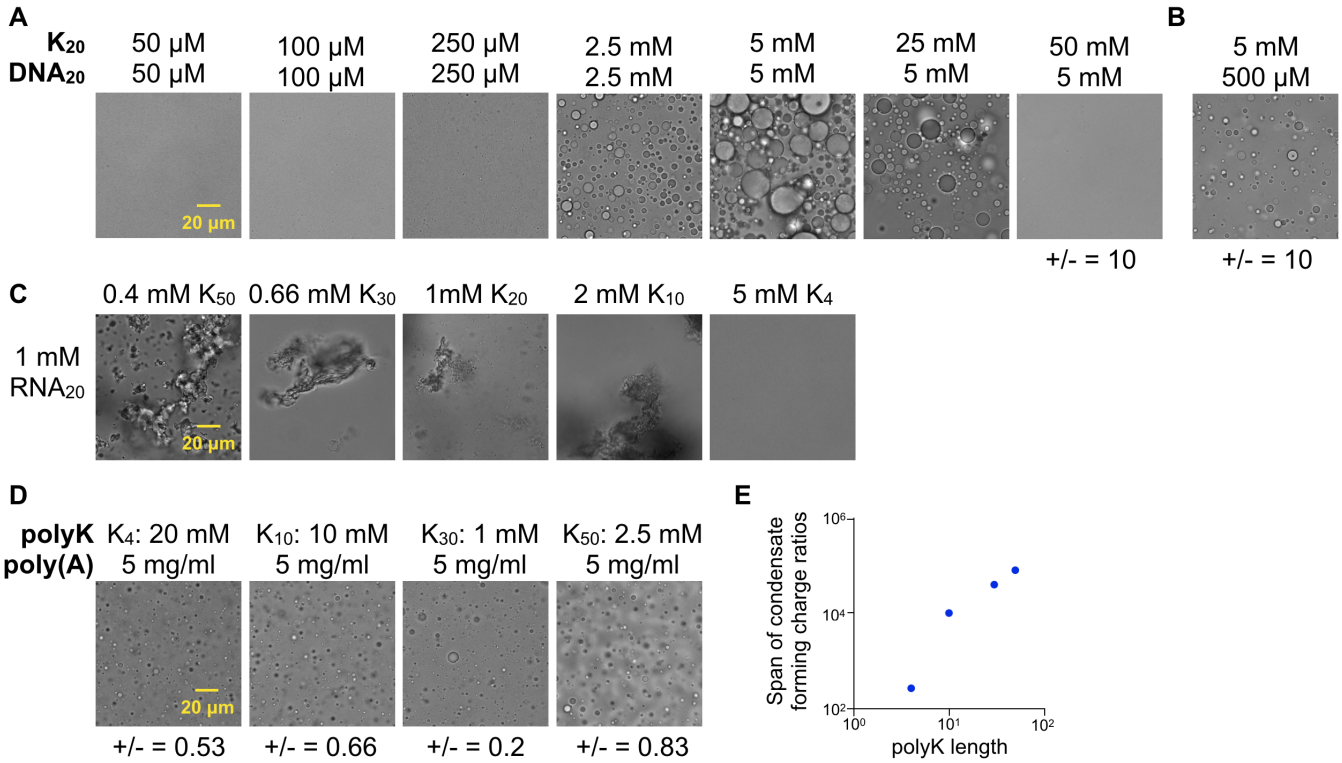

Figure 5. Condensates formed by peptides and oligonucleotides follow the dimer model. (**A**) Representative DIC images of K_{20}-DNA_{20} mixtures. (**B**) Representative DIC image of a K_{20}-DNA_{20} mixture at equal charge ratio as the highest concentration in (A). (**C**). Representative DIC images of polyK-RNA_{20} mixtures. (**D**) Representative DIC images of polyK and poly(A) mixtures. (**E**) The charge ratio span (fold difference between the positive and negative extremes) for polyK and poly(A) mixtures over which condensates form as a function of polyK peptide length.

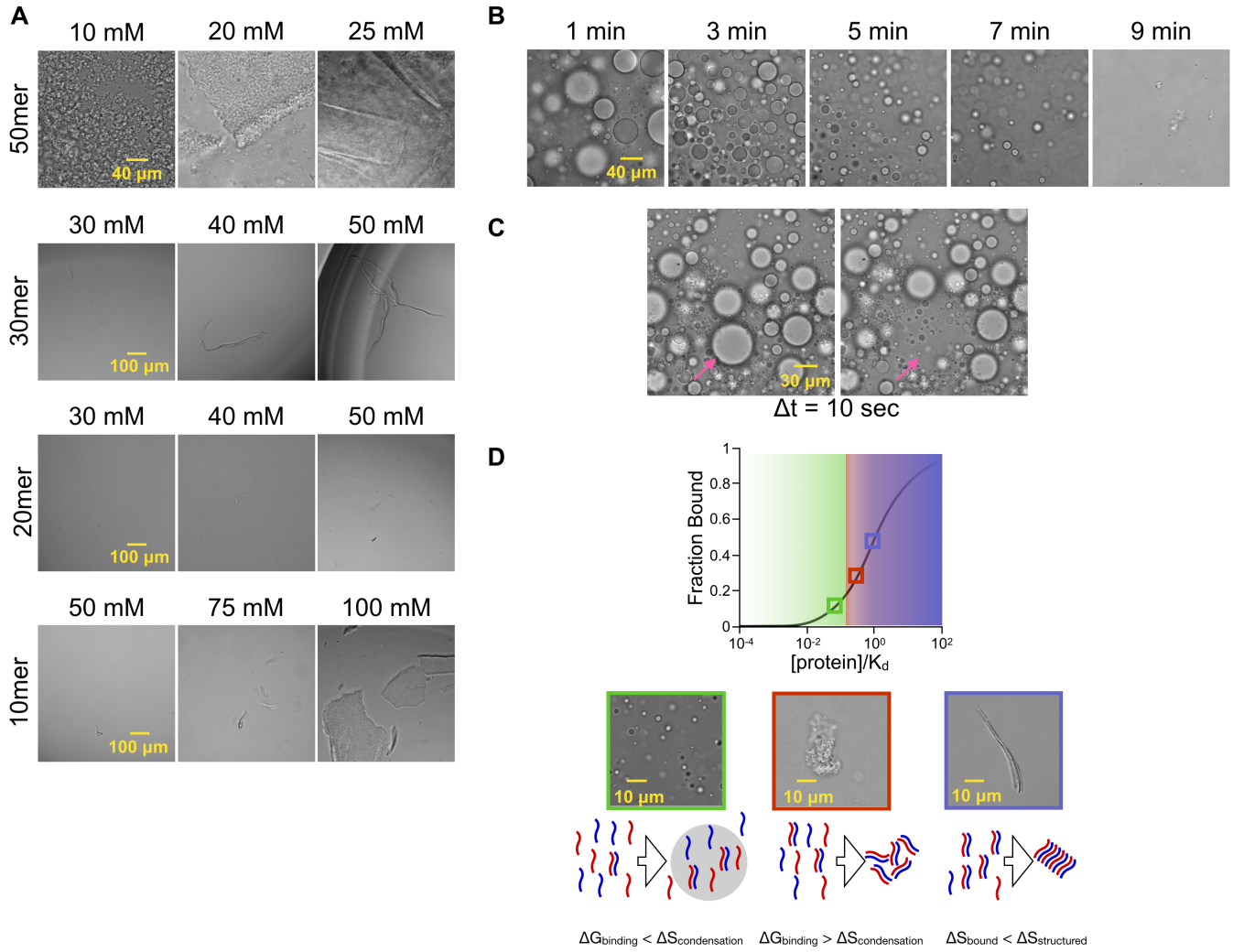

Figure 6. Polyions form structured aggregates at higher concentrations. (**A**) Representative DIC images of K_{50}-E_{50}, K_{30}-E_{30}, K_{20}-E_{20}, and K_{10}-D_{10} equimolar mixtures. (**B**) Representative DIC images of K_{20}-E_{20} equimolar mixture at 15 mM over time. (**C**) DIC images of K_{30}-E_{30} equimolar mixture at 10 mM at 5 minutes and 10 seconds later, arrow indicating a condensate that dissolves. (**D**) Model of phase separation. At [protein]/K_d < 10^{-1}, condensation produces an energy minimum through increased binding. Just above [protein]/K_d = 10^{-1}, condensation is not energetically favorable and the steady state abundance of dimerized protein induces aggregation. At higher [protein]/K_d values more dimerized protein generates higher order structures such as amyloid fibrils.

Supplementary Information

Thermodynamic Calculations:
Spontaneous dimer binding of any molecules arises through a reduction in free energy with an affinity that is defined by a dissociation constant. The steady-state fraction bound of interacting molecules is a function of the total concentration and the dissociation constant[72].

For a homo-binding (dimer) interaction, the dissociation constant is defined as:

$$K_d = \frac{[A]^2}{[AA]} \tag{1}$$

where, $[A]$ is the concentration of the free protein A and $[AA]$ is the concentration of the bound (dimerized) protein. From the total concentration $[A]_{total}$ (free and bound) in the system:

$$[A] = [A]_{total} - 2[AA] \tag{2}$$

and, through substitution:

$$K_d[AA] = ([A]_{total} - [AA])([A]_{total} - [AA]) \tag{3}$$

Through the quadratic equation the total bound concentration as a function of total protein and dissociation constant of dimer binding generates the following equation:

$$[AA] = \frac{1}{8}\left(4[A]_{total} + Kd - \sqrt{(-4[A]_{total} - K_d)^2 - 16[A]_{total}{}^2} \right) \tag{4}$$

The total free concentration can then be calculated through mass balance.

Liquid phase separation in a system generates both dilute and dense regions with concentrations that are defined by the partition coefficient and the dense phase volume fraction. The partition coefficient is defined as:

$$\alpha = \frac{[A]_{condensate,total}}{[A]_{dilute,total}} \tag{5}$$

where, α is the partition coefficient, $[A]_{condensate,total}$ is the total concentration of A (bound and unbound) in the condensate (dense phase) and $[A]_{dilute,total}$ is the total concentration of A (bound and unbound) in the dilute phase. The concentration in each phase can be calculated from the total concentration in the system. First:

$$[A]_{total} = \phi[A]_{condensate,total} + (1 - \phi)[A]_{dilute,total} \tag{6}$$

where, ϕ is the volume fraction of the dense phase and $[A]_{total}$ is the total concentration (bound and unbound) in the system. Here, the number and size of condensates are not considered, rather the total volume of each phase. Rearranging, the concentration in the condensate can be calculated from:

$$[A]_{condensate,total} = \frac{[A]_{total}}{\left(\phi + \frac{1}{\alpha}(1-\phi)\right)} \quad (7)$$

Conversely the dilute phase concentration is:

$$[A]_{dilute,total} = \frac{[A]_{total}}{(\phi(\alpha - 1) + 1\,)} \quad (8)$$

The total bound and unbound concentrations in each phase of the separation (condensate or dilute) can be directly calculated from the total protein concentration within that phase ($[A]_{condensate,total}$ and $[A]_{dilute,total}$) and the binding dissociation constant using Equation 4 and mass balance.

By definition, condensates have a higher concentration than diffuse distribution in the absence of phase separation, and thus a higher fraction bound within the condensate, while dilute regions have a lower concentration than the diffuse state, which reduces the fraction bound in the dilute region (**Fig. 2A**). Summing the volume-weighted, K_d-governed binding in both regions produces the total fraction bound in the system upon phase separation, which generates an effective dissociation constant:

$$K_{d,effective} = \frac{[A]_{sum}{}^2}{[AA]_{sum}} \quad (9)$$

where, $[A]_{sum}$ and $[AA]_{sum}$ are the concentrations of unbound and bound protein in the system.

The effective K_d is lower in the presence of condensates due to an increase in overall bound protein in the system. In the absence of phase separation, the effective K_d is the same as the binding K_d. The effective K_d can then be used to determine the impact of condensation on the Gibbs free energy of the system through the energy of binding:

$$\Delta G = RTln\left(\frac{K_{d,effective}}{K_d}\right) \quad (10)$$

where, ΔG is the change in free energy of binding that arises due to the presence of condensates.

However, there is an entropic cost to de-mixing that occurs when condensates are present. The cost is dependent on the concentrations in the dense and dilute regions and can be calculated from:

$$\Delta G = RT\big(\phi[A]_{condensate}ln([A]_{condensate}) + (1-\phi)[A]_{dilute}ln([A]_{dilute}) - [A]_{total}ln([A]_{total}) \quad (11)$$

where, $[A]_{condensate}$ is the concentration of unbound A in the condensate, $[A]_{dilute}$ is the concentration of unbound A in the dilute phase, and $[A]_{total}$ is the concentration of unbound A in the absence of condensates. Calculating entropic cost as the change in entropy from the system without condensates also captures the decrease in entropy that arises from a reduction of total species moles that occurs through increased dimer formation. The entropic cost is calculated for each protein species (A and AA) and summed to obtain the total entropic cost of condensation.

Summing the change in Gibbs free energy of binding and the entropic cost of de-mixing generates the total energetic impact of phase separation. Yet, to be physically realistic a boundary condition that limits

the concentration in the dense phase is required. In the absence of a boundary condition, at low affinity and high concentration parameters, energetic minimums occur with dense phase concentrations that are not possible. Therefore, a third thermodynamic force needs to be considered - the chemical potential of the protein. The total chemical potential in the presence and absence of phase separation does not obviously change as the total amount of material is the same. However, chemical potential is a function of density[73] and is non-ideal, increasing exponentially as the density approaches concentration saturation[74]. Furthermore, the physical size of the protein will limit the achievable concentration in the dense phase. To capture the energetic impact of dense phase concentration dependent chemical potential the following expression is used:

$$\Delta\mu = RT\theta\phi\left(\exp\left(\left(\frac{C_{max}}{C_{max} - C_{condensate}}\right)^2\right) - \exp\left(\left(\frac{C_{max}}{C_{max} - C_{total}}\right)^2\right)\right) \quad (12)$$

where, $\Delta\mu$ is the change in chemical potential upon condensate formation, θ is a constant, C_{max} is the saturation (maximum) total protein concentration in solution that can be achieved, $C_{condensate}$ is the total protein concentration in the dense phase and C_{total} is the total concentration in the system. This representation was used to generate a chemical potential of protein with non-linear dependence on concentration that approaches infinity at the saturation concentration.

Varying the values of θ and C_{sat} demonstrates that inclusion of $\Delta\mu$ restricts the condensate concentration to physically achievable values, particularly at higher protein concentrations, but has little impact at lower protein concentrations (Fig. S2 and S3). Thus, C_{sat} and θ values were chosen to be 100 mM and 10^{-2}, respectively.

For a hetero-binding interaction with a 1:1 stoichiometry, the thermodynamic calculations will be similar except for the mass balance. The dissociation constant is defined as:

$$K_d = \frac{[A][B]}{[AB]} \quad (13)$$

where, $[A]$ and $[B]$ are the concentrations of free (unbound) proteins A and B, respectively, and $[AB]$ is the bound concentration. The total concentrations are known and thus the concentration of free protein is:

$$[A] = [A]_{total} - [AB] \quad (14)$$

$$[B] = [B]_{total} - [AB] \quad (15)$$

substituting into the previous equation:

$$K_d[AB] = ([A]_{total} - [AB])([B]_{total} - [AB]) \quad (16)$$

and, through the quadratic formula:

$$[AB] = \frac{1}{2}\left([A]_{total} + [B]_{total} + Kd - \sqrt{(-[A]_{total} - [B]_{total} - K_d)^2 - 4[A]_{total}[B]_{total}} \right) \quad (17)$$

The concentrations of A and B can subsequently be determined through mass balances. Under phase separation, the concentrations in the dense and dilute phases can be calculated as described from homo-dimerization for both A and B.

For the calculations here the concentration of A and B are equal. Thus, the Gibbs free energy of binding in the presence of phase separation can be calculated as described for homo-dimer interactions. The entropic cost of de-mixing is the sum of the costs for A, B and AB. And the same chemical potential boundary condition is applied.