



Thermodynamically Driven Assemblies and Liquid-liquid Phase Separations in Biology

Journal:	<i>Soft Matter</i>
Manuscript ID	SM-EMA-11-2018-002285.R1
Article Type:	Review Article
Date Submitted by the Author:	21-Dec-2018
Complete List of Authors:	Falahati, Hanieh; Yale University School of Medicine Haji-Akbari, Amir; Yale School of Engineering and Applied Science, Chemical and Environmental Engineering

Cite this: DOI: 10.1039/xxxxxxxxxxx

Thermodynamically Driven Assemblies and Liquid-liquid Phase Separations in Biology[†]

Hanieh Falahati,^a and Amir Haji-Akbari^{*b}Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxxx

www.rsc.org/journalname

The sustenance of life depends on the high degree of organization that prevails through different levels of living organisms, from subcellular structures such as biomolecular complexes and organelles to tissues and organs. The physical origin of such organization is not fully understood, and even though it is clear that cells and organisms cannot maintain their integrity without consuming energy, there is growing evidence that individual assembly processes can be thermodynamically driven and occur spontaneously due to changes in thermodynamic variables such as intermolecular interactions and concentration. Understanding the phase separation *in vivo* requires a multidisciplinary approach, integrating the theory and physics of phase separation with experimental and computational techniques. This paper aims at providing a brief overview of the physics of phase separation and its biological implications, with a particular focus on the assembly of membraneless organelles. We discuss the underlying physical principles of phase separation from its thermodynamics to its kinetics. We also overview the wide range of methods utilized for experimental verification and characterization of phase separation of membraneless organelles, as well as the utility of molecular simulations rooted in thermodynamics and statistical physics in understanding the governing principles of thermodynamically driven biological self-assembly processes.

1 Introduction

According to the second law of thermodynamics, universe is inevitably moving towards increasing its entropy, which is usually interpreted as lack of order and organization¹. Living organisms, however, maintain a high level of intercellular and subcellular organization and compartmentalization by consuming energy. Subcellular organization provides cells with specialized micro-environments for different cellular functions, while intercellular organization within multicellular organisms makes the formation and functioning of specialized tissues and organs possible. Due to this preponderance of order, it was generally thought that the emergence of order in biological cells can only occur through active processes. A new paradigm that is becoming increasingly popular recently, however, questions this widely accepted viewpoint, and argues that even though energy needs to be consumed for maintaining the integrity of biological cells and organisms, individual assembly processes within them can still occur via thermodynamically-driven phase transitions².

For instance, consider the structural organization of biological

cells, which are comprised of two types of organelles. Many intracellular organelles are membrane-bound, and their composition is maintained through active transport of molecules and ions across their surrounding membranes. Examples include endoplasmic reticulum³, mitochondria⁴ and lysosomes⁵. A second class of intracellular organelles, such as nucleoli^{6,7}, Cajal bodies⁸, and stress granules⁹, lack bounding membranes, and are instead comprised of highly concentrated assemblies of different proteins and RNAs. The question of how membraneless organelles form has fascinated biologists since the initial discovery of the nucleolus, the quintessential membraneless organelle, in the 18th century. For instance, as early as 1898, Montgomery conducted a comprehensive investigation of nucleoli in different cell types, and presented his findings in the form of 346 hand-drawn figures (a sample shown in Figure 1A)⁶. He characterized nucleoli as “*masses of varying dimensions, which may be either globular or irregular in shape, according as they are fluid or viscid in consistency*“. He further described the nucleoli to form via “*coalescence of numerous small portions of nucleolar substance*“, consistent with its fluidity⁶. However, this model faded away due to advancements in cell biology and genetics, which demonstrated that nucleoli form around ribosomal DNA (rDNA) repeats, which are sites of active transcription and processing of ribosomal RNA (rRNA), and ribosomal biogenesis. By the end of the 20th century, membraneless

^a Department of Neuroscience, Yale School of Medicine, New Haven, CT, 06510; E-mail: hanieh.falahati@yale.edu

^b Department of Chemical & Environmental Engineering, Yale University, New Haven, CT, 06520, USA; E-mail: amirhajiakbaribalou@yale.edu

organelles were commonly thought to form via active processes, commensurate with their active biological function. A number of influential works by Sear², Brangwynne *et al.*¹⁰, Li *et al.*¹¹, and Kato *et al.*¹², however, redirected the focus to the liquid/gel nature of membraneless organelles, and the possibility that thermodynamically driven liquid-liquid phase separation (LLPS) might be responsible for their *in vivo* assembly^{13–25}.

Another possible example of thermodynamically-driven phase separation in living systems is the reorganization of lipids and proteins within biological membranes (Figure 1C–D). In principle, the biological membranes that encompass cells and bounded organelles are spatially inhomogeneous two-dimensional mixtures of a wide variety of proteins and lipids. Lipid drafts, or microdomains enriched in cholesterol and saturated lipids such as sphingomyelin, are the most widely known manifestations of such heterogeneity. They are proposed to play an important role in the function and localization of certain membrane proteins such as ion channels, and are suspected to form as a result of thermodynamically driven phase separation occurring within two-dimensional liquid-like membranes^{27,30–32}.

Phase separations in biology are not limited to intracellular processes, and have also been invoked²⁸ to explain the organization of tissues within multicellular organisms (Figure 1E–H). For instance, the dissociated embryonic cells originating from the same tissue are known to form spherical bodies commensurate with a tissue constituting a distinct mesoscopic thermodynamic phase. Furthermore, mixtures of cells originating from different tissues tend to separate into clusters rich in cells belonging to distinct tissues, and liquid-like cell aggregates with lower effective surface tensions always envelop clusters with higher effective surface tensions. This mesoscopic demixing of cells is also concentration- and composition-dependent²⁸. All these features are consistent with a liquid-liquid phase separation, which is thought to be mediated by the differential adhesiveness of cells originating from different tissues. This difference in adhesion propensity arises from different adhesion molecules such as Cadherins^{29,33} at the surface of cells.

It is necessary to emphasize that biological phase separations are not limited to LLPS, and can sometimes culminate in the formation of crystalline solids. Unlike liquids, molecules in crystalline solids exhibit long range order. A notable example is a process known as *biomineralization*, which, for instance, results in the formation of bones and teeth³⁴. The formation of actin filaments^{35,36} and microtubules³⁷ is also a phase separation resulting in the formation of solids.

In addition to its suggested role in cellular organization under normal circumstances, phase separation can also be pathological^{38–40}, and diseases such as Alzheimer's⁴¹, Amyotrophic Lateral Sclerosis (ALS)^{42,43}, Parkinson's⁴⁴ and cataract⁴⁵ occur as a result of the emergence of pathological assemblies within cells and tissues. Understanding the role of thermodynamics in the formation of such assemblies is key to identifying effective treatments for these medical conditions.

In recent years, several excellent reviews^{13–25,46} have been published on the subject of biomolecular liquid-liquid phase separation. This current review has been written from a molecu-

lar thermodynamics perspective and aims at providing a brief overview of thermodynamically driven phase separations in biological systems, with a particular focus on the role of LLPS in membraneless organelle assembly. Understanding LLPS in biological systems requires an interdisciplinary approach that is built upon the theory of phase separation rooted in classical thermodynamics and polymer physics, and applying the cutting-edge experimental and computational approaches to address the complexity of biological systems. This paper is aimed at providing minimal conceptual ingredients of such an exploration and is organized as follows. We dedicate Sections 2.1 and 2.2 to discuss the thermodynamics and kinetics of phase separations in multi-component systems, respectively. Section 3 discusses experimental evidence for biological LLPS, with *in vitro* and *in vivo* studies highlighted in Sections 3.1 and 3.2, respectively. Section 4 is dedicated to molecular simulations, and their usefulness in understanding biomolecular phase separations. Finally, we put forward some major unaddressed questions about biological self-assembly and phase separation, and discuss some potential areas of future exploration in Section 5.

2 Thermodynamics and Kinetics of Phase Separation

2.1 Thermodynamics of Phase Separation

Phase separation refers to a process that occurs in multi-component mixtures, and culminates in the formation of new phases with densities and/or compositions different from the original phase. The maximum number of distinct coexisting phases that can emerge within a mixture is given by the *phase rule*, which can be derived from classical thermodynamics. According to phase rule, a mixture of k nonreactive components can coexist in a maximum of $k + 2$ distinct phases. In mixtures with reactive components, this maximum is decreased by r , the number of linearly independent chemical reactions in the system, which is equal to the rank of the stoichiometry matrix⁴⁷. We will primarily focus on the coexistence of two distinct phases, as three- or more- phases can only coexist over a narrower range of thermodynamic variables. In a single-component system, for instance, it will only be at the triple point where three phases can coexist with one another.

In general, phase separation usually starts within liquid (or gaseous) mixtures. One notable exception constitutes multi-component crystals, such as metallic alloys⁴⁸ or colloidal mixtures⁴⁹, which can in principle separate into two or more crystalline solids with distinct compositions and symmetries. The new phase that would emerge as a result of phase separation within a liquid mixture can, however, be a solid or a liquid. The corresponding phase transitions, which are referred to as *precipitation* and *liquid-liquid phase separation*, respectively, can both occur in biological systems. Precipitation plays a pivotal role in the formation of ordered structures such as actin filaments^{35,36} and microtubules³⁷, as well as bio-mineralization (e.g., bone formation)³⁴. It can also be pathological e.g., in the formation of amyloid plaques in Alzheimer's disease^{41,50}. Precipitation is also a key separation process in structural biology, utilized for protein

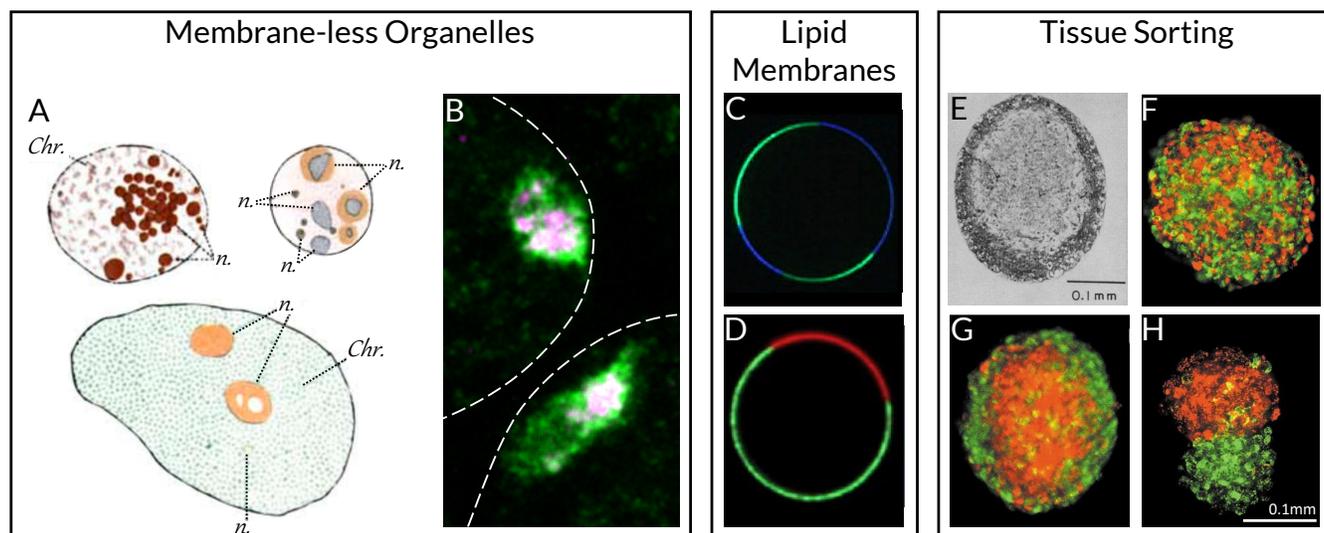


Fig. 1 Phase separation in biological systems. The formation of membraneless organelles is proposed to be driven by LLPS. **(A)** Hand drawing of the nucleolus (*n.*) in the nuclei of different cell types by Montgomery in 1898⁶. Based on their shape and behavior, he noticed that the nucleoli are liquid or viciid. *Chr.*: Chromatin. **(B)** Nucleolus is not homogeneous. Two fluorescently-tagged nucleolar proteins, Fibrillarlin (magenta) and Modulo (green), exhibit different localization patterns within nascent nucleoli of *Drosophila* embryos. Images were obtained using expansion microscopy²⁶, and the dashed lines show the boundaries of nuclei. **(C-D)** Lipid membranes are capable of separating into inhomogeneous subdomains through LLPS. Macroscopic phases of ordered and disordered liquids separated in giant unilamellar vesicle (GUV) **(C)** or vesicles derived from plasma membranes **(D)** (From 27). **(E-H)** LLPS is also proposed to be able to drive cellular organization. **(E)** A mixture of dissociated 5-day heart ventricle cells and 5-day liver cells of chick embryo demix into a sphere, with the liver tissue enveloping a core of heart tissue (from Steinberg 1963²⁸). **(F-H)** Equilibrium distribution of cells expressing different levels and kinds of Cadherins. **(F)** Mixture of N5A cells expressing the same type and level of Cadherin remain intermixed and form a sphere. **(G)** Demixing of N5A cells (red) that express N-cad at levels 50% higher than N2 cells (green). **(H)** A mixture of B-cad (green) or R-cad (red) expressing cells separate into clusters with red cells partially capping a B-cad-expressing mass. Adopted from Duguay *et al.* 2003²⁹.

crystallization^{51,52}.

The focus of this paper is, however, on LLPS, which is thought to play an important role in the assembly of membraneless organelles, the formation of lipid rafts in membranes, and cell sorting in biological tissues. In general, the term 'liquid' usually refers to a phase that is amorphous, i.e., that its molecular structure lacks long-range translational order. The particular term utilized for describing such amorphous states of matter sometimes depends on its mechanical and transport properties, i.e., its relaxation dynamics. In a biological context, in particular, the term liquid usually refers only to amorphous phases that relax quickly and do not therefore withstand shear deformation. This mechanical definition is of particular interest to biology, due to the functional importance of fast dynamics within the new phase (e.g., for transport of ions and macromolecules). Amorphous phases that relax more slowly, however, constitute a wide range of materials, from gels to glasses, and exhibit interesting dynamical properties such as aging^{53–55} (i.e., time dependent autocorrelation). The processes such as aggregation and physical gelation that result in the formation of such amorphous states fall into the general category of liquid-liquid phase separation.

In order for LLPS to be thermodynamically favored, it needs to result in a decrease in the mixture's Gibbs free energy. Let $\mathbf{x}^i \equiv (x_1^i, x_2^i, \dots, x_k^i)$ be the composition of a homogeneous liquid mixture \mathcal{C} , with x_j^i 's the mole fractions of individual components. Upon phase separation, \mathcal{C} will separate into two coexisting

phases, the original phase with composition \mathbf{x}^f , and a new phase y with composition \mathbf{y}^f (Figure 2A). Note that \mathbf{x}^i will be located on a tie line connecting \mathbf{x}^f and \mathbf{y}^f , i.e., there will be a partition constant $0 \leq \lambda \leq 1$ so that $\mathbf{x}^i = \lambda \mathbf{x}^f + (1 - \lambda) \mathbf{y}^f$. Since the two phases x and y will be at equilibrium, $\boldsymbol{\mu}_x(\mathbf{x}^f) = \boldsymbol{\mu}_y(\mathbf{y}^f)$ with $\boldsymbol{\mu}_\alpha \equiv (\mu_{\alpha,1}, \mu_{\alpha,2}, \dots, \mu_{\alpha,k})$ the chemical potential vector for phase $\alpha = x, y$. The change in the free energy of the system as a result of phase separation will therefore be given by:

$$\begin{aligned} \Delta g &= \lambda g_x(\mathbf{x}^f) + (1 - \lambda) g_y(\mathbf{y}^f) - g_x(\mathbf{x}^i) \\ &= \lambda \mathbf{x}^f \cdot \boldsymbol{\mu}_x(\mathbf{x}^f) + (1 - \lambda) \mathbf{y}^f \cdot \boldsymbol{\mu}_y(\mathbf{y}^f) - \mathbf{x}^i \cdot \boldsymbol{\mu}_x(\mathbf{x}^i) \\ &\stackrel{\boldsymbol{\mu}_x(\mathbf{x}^f) = \boldsymbol{\mu}_y(\mathbf{y}^f)}{=} \mathbf{x}^i \cdot [\boldsymbol{\mu}_x(\mathbf{x}^f) - \boldsymbol{\mu}_x(\mathbf{x}^i)] \end{aligned} \quad (1)$$

Here $g_\alpha(\cdot)$ is the molar Gibbs free energy of phase $\alpha = x, y$. The described phase separation will be thermodynamically favored if $\mathbf{x}^i \cdot [\boldsymbol{\mu}_x(\mathbf{x}^f) - \boldsymbol{\mu}_x(\mathbf{x}^i)] \leq 0$. Also, the constraints imposed on chemical potentials and the overall composition of the two phases imply that \mathbf{x}^f , \mathbf{y}^f and λ can be uniquely determined from the temperature, pressure and composition of the original mixture. In general, it is possible to obtain empirical or semi-empirical expressions for $\boldsymbol{\mu}_\alpha$, which can then be utilized to predict the possibility– or lack thereof– phase separation for any given mixture. For the special case of a binary system with components A and B , the compositions of coexisting phases can be uniquely determined for a given temperature and pressure, and the phase

separation will be thermodynamically favored if:

$$\int_{x_A^i}^{x_A^f} \left[\frac{x_A^i}{1-x_A^i} - \frac{x_A}{1-x_A} \right] \left(\frac{\partial \mu_A}{\partial x_A} \right)_{P,T} dx_A < 0 \quad (2)$$

Eq. (2) follows from (1) and the Gibbs-Duhem equation, i.e., $d\mu_B = -x_A d\mu_A / (1-x_A)$. Eqs. (1) and (2) describe the necessary thermodynamic conditions for phase separation, without telling us anything about its underlying physics. The latter can be attained by decomposing Δg into its enthalpic and entropic contributions, which correspond to the strength of intermolecular interactions, and the number of accessible microstates, respectively. Letting $\Delta g = \Delta h - T\Delta s$ and assuming that neither Δh nor Δs are strong functions of temperature, one can imagine the following scenarios. In the case of $\Delta h > 0$ and $\Delta s < 0$, phase separation cannot occur under any circumstances, and the system will remain mixed at all conditions. If $\Delta h < 0$ and $\Delta s > 0$, the system will always be demixed and a homogeneous mixture will be unstable. We will consider the other two scenarios in which the system can remain homogeneous, or phase separate depending on thermodynamic conditions.

The most common scenario is when $\Delta h < 0$ and $\Delta s < 0$, and corresponds to situations in which molecules within the demixed phases experience stronger intermolecular interactions. The system, however, experiences a decline in the number of accessible microstates due to demixing. This usually occurs when there are unfavorable intermolecular interactions between some components within the original mixture, and results in the separation of the components that do not "like" each other into distinct phases. Such separation, however, eliminates all the microstates in which molecules of different types are randomly distributed within a homogeneous liquid, and results in a decrease in entropy. The free energetic penalty associated with such entropic loss, $-T\Delta s$, will be proportional to temperature. Therefore, beyond a critical temperature known as *upper critical solution temperature (UCST)*, demixing will become thermodynamically unfavorable, and the system can only exist in a mixed state.

Figure 2A depicts the prototypical phase diagram for a binary mixture with a UCST. According to the phase rule, temperature and the composition of the original homogeneous mixture are sufficient for determining whether phase separation occurs, and if so for predicting the compositions of coexisting liquids. As can be seen in Figure 2A, at concentrations outside the shaded dome, the system will remain fully mixed. What is notable about binary- or pseudo-binary- mixtures is that changing the concentration within the dome will not affect the compositions of the coexisting liquids, and will only alter the partition constant, λ . In biological systems, examples of UCST have been observed *in vitro* for the low-complexity domain of the RNA binding protein FUS⁵⁶, a disordered Nauge protein⁵⁷, and lipid bilayers⁵⁸, and *in vivo* for the nucleolar proteins Fibrillarin, Nopp140, RNA polymerase I, and Pitchoune⁵⁹.

The other less common situation occurs when the entropy of demixing is positive, i.e., when the mixture is more ordered than the separated phases. This can occur for one of the following reasons. One possibility is the existence of very strong

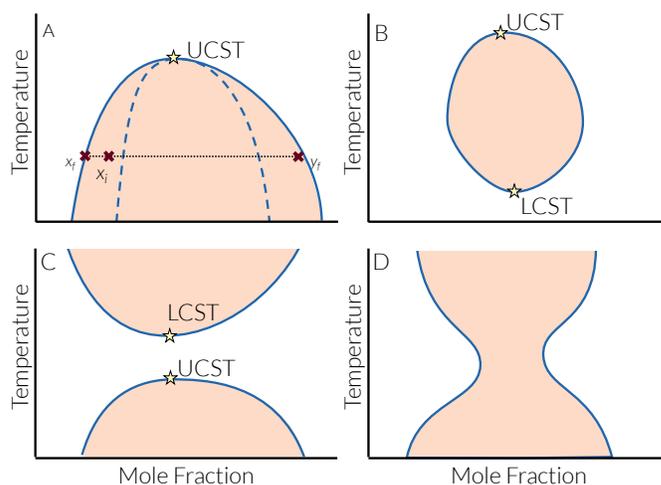


Fig. 2 Schematic representation of different types of phase diagrams. A solution that exhibits a phase separation: (A) with a UCST and a dome-shaped phase diagram; (B-C) with both an LCST and a UCST. (B) When the LCST is smaller than the UCST, the phase diagram will be loop-shaped, while (C) if LCST is larger, it will have a regular and an inverted dome; (D) with no critical temperature and an hourglass-shaped phase diagram. In either case, at permissive temperatures, phase separation will occur only if the mole fraction, x^i , is within the light orange region of the phase diagram, and will culminate in the formation of two phases with mole fractions x^f and y^f , respectively. Note that x^f and y^f only depend on temperature, and not the initial composition of the mixture. Dotted line shows the tie-line and dashed line shows the spinodal line.

and/or highly directional intermolecular interactions between non-identical components in the original mixture. In a binary mixture, for instance, this can occur if hydrogen bonds can only form between A and B molecules. At low temperatures, demixing will be thermodynamically unfavorable since the enthalpic cost of breaking such strong interactions between unlike molecules cannot be compensated by the entropic gain due to demixing. Beyond a critical temperature known as the *lower critical solution temperature (LCST)*, however, the entropic term will become dominant and the system will phase separate. The existence of LCST can therefore be attributed to effective interactions that depend on temperature⁴⁶. Such mixing-induced ordering will eventually fade at sufficiently high temperatures due to the weakening of the original directional interactions, so a UCST will also exist, as depicted in Figure 2B. In other words, when the LCST temperature is smaller than UCST, the dome of Figure 2A will be replaced with a loop. This behavior has been observed *in vitro* for a spindle-associated protein, BuGZ⁶⁰.

Another type of mixing-induced ordering can occur for polymeric solutions close to the vapor-liquid critical point of the pure solvent. Under such circumstances, critical fluctuations in the solvent will decrease the number of configurations accessible to polymeric chains, henceforth culminating in a negative entropy of mixing. This scenario, which is uncommon in biological systems, results in phase diagrams shown in Figure 2C and 2D⁶¹ depending on the separation between the low-temperature demixing region and the critical temperature of the solvent.

Despite this lack of universality in how thermodynamically driven phase separations are affected by changes in temperature, they are distinct from active assembly processes in that they are all *reversible*. In other words, active assembly processes are driven by irreversible enzymatic reactions and proceed at higher rates when temperatures are higher. This means that changing temperature will only impact their kinetics, and not their overall direction. Thermodynamically driven assemblies, however, can be reversed by changing temperature, e.g., increasing it in the case of a UCST or decreasing it in the case of an LCST. Such reversibility has been shown to be pivotal in determining whether a particular *in vivo* assembly process is thermodynamically driven or active, e.g., by monitoring its response to oscillations in temperature⁵⁹.

At the end of this section, it is necessary to emphasize that while we have only discussed phase separation in mixtures, both liquid-liquid^{62–65} and solid-solid^{66–72} transitions can also occur in single-component molecular^{62–66,68,69} and colloidal^{67,70–72} systems. Unlike phase separation in mixtures, the coexisting phases in single-component systems can only differ in density or symmetry (i.e., the arrangement of their constituent molecules). The ability of a pure substance to exist in distinct crystalline and amorphous forms is referred to as *polymorphism* and *polyamorphism*, respectively, and has been extensively studied in experiments^{63,66,68,69,72} and simulations^{62,64,65,67,70,71}.

2.2 Kinetics of Phase Separation

The thermodynamic framework discussed above is only sufficient for describing the equilibrium behavior of a mixture, and does not provide any information about the kinetics of the underlying phase separation process. In order to fully understand the role of LLPS in biological self-assembly, it is necessary to characterize its kinetics, as the kinetics of a particular LLPS process will determine whether it can occur over "biologically relevant" timescales. For instance, stress granules need to form swiftly when an organism is under stress, and their assembly timescale needs to be commensurate to this function. As another example, consider the relatively short duration of cell cycles at early stages of embryogenesis, which imposes a functional upper limit on the assembly timescales of membraneless organelles such as nucleoli and histone-locus bodies. Moreover, the pathogenicity of aggregates assembled through phase separation in aging-related diseases will depend on the timescales of their formation. Therefore, the biological implications of thermodynamically driven phase separations cannot be thoroughly understood without accounting for their kinetics, as such assembly processes can only be relevant if they are faster than the corresponding biological clock of interest.

As a first-order phase transition, LLPS can occur via two distinct mechanisms, depending on the magnitude of the thermodynamic driving force Δg . For small Δg 's, e.g., close to the phase boundaries of Figure 2, demixing will occur through *nucleation and growth* (Figure 3). During nucleation, which is an activated stochastic process, instantaneous thermal and compositional fluctuations result in the formation of small nuclei of the new phase within the old mixture. Such nuclei will be thermodynamically unstable and will therefore be more likely to melt due to their

large specific surface areas, unless they are larger than a critical size. The free energetic cost of forming such a critical nucleus, ΔG_{nuc} , is referred to as the *nucleation barrier*. The likelihood that a critical nucleus will form within a mixture is proportional to $e^{-\Delta G_{\text{nuc}}/kT}$, with k the Boltzmann constant. Therefore a larger ΔG_{nuc} will imply a lower likelihood for the formation of a critical nucleus. Upon its formation, however, a critical nucleus can grow until the system reaches the predicted partition constant λ , and the growth timescale will scale with transport properties of the original mixture. The separation between the nucleation and growth timescales will be larger when $\Delta G_{\text{nuc}} \gg kT$, and phase separation will become a nucleation-limited rare event. The closer the nucleation barrier is to kT , the smaller will such a separation of timescales be. In the limit of $\Delta G_{\text{nuc}} \rightarrow kT$, LLPS will be growth-limited. Whenever the nucleation rate is large, e.g., in the growth-limited regime, multiple nucleation events will occur within a single mixture, and the completion of LLPS will proceed through several simultaneous non-equilibrium processes⁷³. One such process that occurs as a result of diffusion within the continuous phase is known as *coarsening* or *Ostwald ripening*⁷⁴ in which larger nuclei will grow in the expense of smaller nuclei that melt due to their lower thermodynamic stability. Another important process is called *coalescence* in which smaller nuclei collide and join to form a larger nucleus.

This simplified picture is the essence of *classical nucleation theory (CNT)*^{75,76}, which is the most widely used quantitative framework for understanding nucleation. CNT assumes that nucleation is a single-step process, and a steady-state distribution of precritical nuclei is established at timescales that are orders of magnitude shorter than the nucleation time. According to CNT, nucleation barrier will be given by:

$$\Delta G_{\text{nuc}} = \frac{16\pi\sigma^3}{3|\Delta g|^2} \quad (3)$$

with σ the interfacial tension between the two phases, which only depends on temperature and the compositions of the two coexisting phases. It can be generally stated that the nucleation barrier is expected to decrease upon increasing Δg . It is necessary to emphasize that the key assumptions of CNT might be violated in some systems, and alternatives of CNT, such as multi-step nucleation^{77,78}, have been formulated in the literature.

Like other first-order phase transitions, nucleation in LLPS can be homogeneous or heterogeneous. In homogeneous nucleation, the critical nucleus forms endogenously within the mixture, while in heterogeneous nucleation, an external entity such as an insoluble impurity provides a template for nucleation, and results in a decrease in the size of the critical nucleus and the magnitude of the nucleation barrier. Heterogeneous nucleation generally occurs at higher rates, and is a means of exerting spatiotemporal control on phase separation in biological systems⁷⁹. An extension of CNT for heterogeneous nucleation was proposed by Turnbull⁸⁰, and predicts that $\Delta G_{\text{het}} = f(\theta_c)\Delta G_{\text{homo}}$. Here $f(\theta_c)$ is called the *potency factor* and depends on θ_c the contact angle between the two phases and the external surface. From a physical perspective, potency factor is a measure of differential attractive-

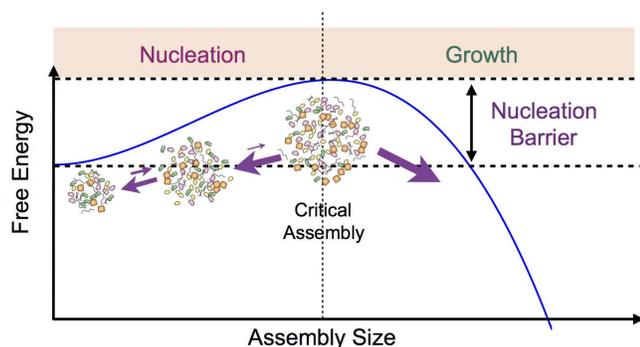


Fig. 3 Nucleation and growth in phase separation processes. A first-order phase separation starts with nucleation during which tiny assemblies of the new phase emerge within the old phase. Smaller assemblies are thermodynamically unstable due to their large specific surface areas, and are therefore more likely to shrink unless they reach a critical size. If the nucleation barrier is large, then the formation of new assembly becomes nucleation-limited. Otherwise, this process becomes growth limited.

ness of the interaction between the nucleating surface and the two liquids⁸¹. Lower potency factors correspond to situations in which the external surface has a higher propensity for the new phase, which results in smaller contact angles, and considerably larger nucleation rates at identical Δg 's.

Nucleation and growth is the mechanism of demixing when the original mixture is metastable, i.e., when an arbitrary concentration fluctuation will always be dampened unless its amplitude is sufficiently large (e.g., in a critical nucleus). Another regime, however, is possible in which the mixture becomes mechanically unstable. From a thermodynamic perspective, this will occur if the Hessian of $g(\mathbf{x})$ — or its generalized second-order derivative with respect to its independent arguments— has at least one negative eigenvalue. For a binary system, this will imply a situation in which $g''(x_A) < 0$. Under such circumstances, concentration fluctuations with sufficiently large wavelengths will be amplified, which will in turn drive the system into a demixed state. Several mean-field theories have been developed to describe this phenomena, which is typically known as *spinodal decomposition*^{82–84}. Within these theories, it is predicted that a transition from the nucleation-and-growth regime to the spinodal regime occurs at a spinodal line (Figure 2A), which is the loci of inflection points of $g(\mathbf{x})$. In the case of binary mixtures, this will correspond to loci of $g''(x) = 0$. During spinodal decomposition, fluctuations with different wavelengths get exponentially amplified, but with different time constants. As a result, the morphology of the system is dominated by the fastest growing wavelength, resulting in a characteristic length scale for phase separated domains. There is extensive experimental⁸⁵ and computational^{86–88} evidence for spinodal decomposition. Spinodal decomposition has been experimentally observed in the phase separation of Ddx4 protein⁸⁹. Due to the mean-field approximation inherent in the spinodal model, however, it is not trivial to pinpoint the boundary at which the phase separation mechanism transforms from nucleation and growth to spinodal decomposition.

3 Phase Separations in Biological Systems: Experimental Investigations

As mentioned in Section 1, phase separation can occur in many different biological settings. The resurgent interest in the role of LLPS in biological assembly can, however, be primarily attributed to their potential role in the formation of membraneless organelles. Understanding the mechanisms by which membraneless organelles form, is particularly important as these cellular bodies carry out several essential cellular functions⁹⁰. In the nucleus, for instance, Cajal bodies form at snRNA loci and constitute the sites at which the splicing machinery is assembled^{91,92}. Histone-locus bodies contain the macromolecular machinery necessary for the transcription and processing of histone mRNAs⁹², and nucleoli are the sites of ribosomal biogenesis and several other important cellular functions. In the cytoplasm, stress granules and processing bodies (P-bodies) form in response to cellular stresses, and limit the translation of mRNAs⁹³. Each of these organelles has a distinct composition; they are rich in RNAs and proteins necessary for their unique function, and lack the components that might inhibit such roles^{11,94,95}. Furthermore, these organelles are not structurally homogeneous, and the macromolecules within them tend to organize into smaller domains. For instance, depending on the organism, nucleolus can be comprised of two to three micro-domains that specialize in different steps of ribosomal biogenesis (Figure 1A-B)^{6,96–98}. Understanding the underlying mechanisms for the formation of such intricate, specialized structures, and the regulation of the timing and location of their assembly, is therefore critical to our ability to understand the biological functions of such organelles.

In addition to their significant role in normal cellular functions, membraneless organelles can exhibit anomalous morphologies in several pathological disorders. For instance, a hallmark of cancer cells is an increase in the number and size of their nucleoli, a fact that was described as early as 1896 by Giuseppe Pianese^{99,100}. Moreover, mutations in stress granule proteins such as FUS have been associated with neurodegenerative diseases such as ALS or frontotemporal dementia (FTD) in which pathological aggregates form^{21,24,93,101}. This also underlines the necessity for a better mechanistic understanding of the driving forces and the regulatory mechanisms of the membraneless organelle assembly.

In this section we describe the experimental approaches commonly used for studying phase separation in biological systems (summarized in Figure 4). In each case we describe what is measured by each technique, how the results support the LLPS model, and also the limitations of each approach.

3.1 In Vitro Investigation of LLPS

The propensity of protein/salt water mixtures to phase separate *in vitro* is not new in biology. For instance, the phase diagrams of globular proteins such as lysozyme, and γ -crystallin dissolved in aqueous NaCl solutions have been studied since 1970's, and their demixing into liquid-like, amorphous and crystalline phases at different conditions has been reported^{102,103}. Such studies of phase separation were primarily conducted with the aim of

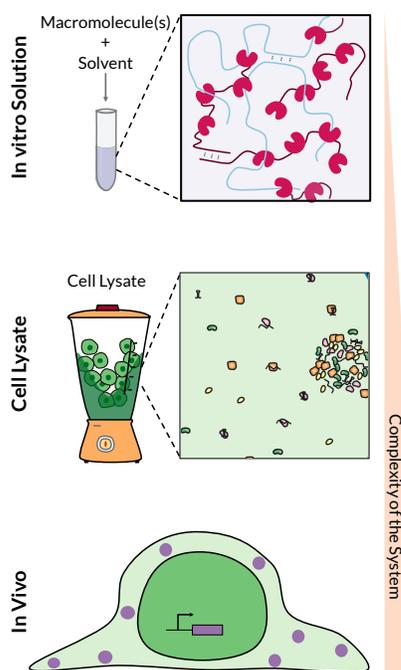


Fig. 4 Experimental investigations of liquid-liquid phase separation in biomolecular systems. *In Vitro* systems are typically comprised of a small number of macromolecules, such as RNAs and/or proteins, dissolved in an aqueous solution, and, due to their well-defined compositions, are ideal for examining the effect of different thermodynamic factors on the phase behavior. Furthermore, *in vitro* studies have made considerable contributions to our current understanding of the molecular properties that drive phase separation, as well as the effect of pathological mutations on the demixing propensity. In order to capture the inherent complexity of biological systems, a number of *in vitro* investigations have used cell lysates to study LLPS. Finally, several studies have directly focused on LLPS and membraneless organelle assembly *in vivo*, in order to account for full complexity of biological systems. Such studies examine three main aspects of membraneless organelles: (i) mechanical and transport properties, (ii) the effect of thermodynamic variables, such as concentration and temperature, on LLPS, and (iii) kinetics and spatiotemporal regulation of membraneless organelle assembly. See text for more details.

understanding and optimizing protein crystallization, which is an important technique in structural biology, and a prerequisite for determining the three-dimensional structures of proteins using X-ray crystallography. In this context, liquid-like and amorphous assemblies of proteins are unwanted products of processes aimed at producing protein crystals.

With resurgent interest in the phase separation model and the possible role of LLPS in membraneless organelle assembly, similar *in vitro* assays are now utilized for studying the condensation of proteins and/or RNAs that localize to such *in vivo* assemblies. These studies primarily focus on mapping out the phase diagrams of mixtures with one- or a few- types of macromolecules, by varying temperature, pH and ionic strength of the solvent, and the concentrations of the constituent macromolecules. Such *in vitro* approaches have proven invaluable in assessing the sensitivity of the thermodynamics and kinetics of phase separation to changes in molecular properties, and comparing the phase separation propensity of different proteins under well-defined condi-

tions. Such molecular properties, for instance, can be tuned by mutating different domains and residues of the phase-separating proteins, or by engineering biology-inspired modules e.g., by introducing sequence repeats^{11,56,104–110}. These approaches have been widely used in understanding the role of mutations in proteins that can form pathological aggregates. Mutations in the coding regions of certain proteins such as FUS^{101,111}, TDP-43¹¹² and hnRNPA1¹¹³, which are all observed in pathological disorders, are shown to change the boundaries of the phase diagram, as well as the kinetics of phase separation (referred to as aging). Other *in vitro* studies examine the effect of introducing reactions that result in post-translational modifications of proteins, such as phosphorylation, methylation or ubiquitination, on the phase boundaries^{25,114–117}. The effect of post-translational modifications has been observed in the assembly of proteins such as Coilin¹¹⁸, nucleolar proteins (reviewed in Ref. 119), MEG proteins¹²⁰, eIF-2 α ¹²¹, FUS¹¹⁵, Ddx4⁵⁷, UBQLN2¹¹⁶, and LAT¹²², and was also used to engineer synthetic RNA-protein assemblies *in vitro*^{109,114}.

In vitro studies have been particularly pivotal in identifying the types of biological macromolecules that can phase separate under physiological conditions²³. Phase separation has been observed in solutions of one or more proteins, mixtures of RNA and protein¹¹³, solutions of RNA¹²³, and mixtures of protein and liposomes¹¹⁷. These phase separating biomolecules can be classified into two distinct categories. The first class is comprised of those with multivalent interactions. Similar to patchy colloids (i.e., colloids with multiple interaction sites at their surface), they typically contain repetitive modules that can attract and form non-covalent bonds with other modules in their binding partners. For proteins, these repetitive modules are often folded domains. Examples include the signaling protein WASP, which has multiple proline-rich motifs that can bind to the three SH3 domains of Nck¹¹, or proteins with multiple RNA binding domains. In the latter case, RNA can function as a cross-linker between the proteins that contain multiple RNA binding domains. The second class is comprised of proteins with intrinsically disordered regions (IDRs), which have also been observed to phase separate *in vitro*¹². Such regions lack well-defined three-dimensional structures and are capable of forming multiple weak interactions with other molecules. Polypeptides that are exclusively comprised of IDRs are typically referred to as *intrinsically disordered proteins (IDPs)*¹²⁴. It is necessary to emphasize that these two categories are not mutually exclusive, as many phase separating proteins contain both repetitive modules and IDRs. In this paper, we will mostly focus on IDPs and IDR-containing proteins and the computational approaches to study their ability to undergo phase separation.

Although such *in vitro* studies have been pivotal in laying out of basic understanding of biological condensates, they are incapable of properly accounting for the complexity of *in vivo* systems that are comprised of millions of different components. This is due to the fact that adding even one extra component to a mixture can potentially result in a change in the phase boundary in favor of mixing or demixing of original components, and can even cause the emergence of new liquid or solid phases that would otherwise not form in the original mixture. For instance, adding

nucleophosmin to a mixture of fibrillarin in buffer will result in the emergence of a new liquid phase⁹⁷. Therefore, it is not clear whether the macromolecules that phase separate *in vitro* would also condense inside cells in the presence of millions of other components.

With the aim of better representing *in vivo* systems, several groups have conducted *in vitro* assays of successively complex multi-component mixtures. For instance, in a seminal work in Rosen's lab, the twelve components of the T-cell receptor signaling pathway were reconstituted *in vitro*, which then segregated into a multi-component assembly containing the kinase and depleted of the phosphatase, and a second aqueous phase of the phosphatase¹²². Similarly, the phase separation of the six components of the post-synaptic density was studied *in vitro*¹²⁵. Interestingly, both these reconstitutions involve lipid bilayers in order to mimic the *in vivo* localization of the constituent proteins, which occurs through establishing attractive interactions between the transmembrane and cytoplasmic proteins. In both cases, only the thermodynamics of phase separation was examined. It is, however, entirely plausible that the surface of the lipid bilayer can act as a heterogeneous nucleation site for the formation of the respective assemblies, and therefore play a role in the spatiotemporal regulation of the formation of such condensates.

The *in vitro* systems that are the most similar to intracellular conditions are cellular extracts, which contain the bulk of components present *in vivo*. For instance, Hancock showed that expanding intact nuclei in a hypotonic medium will result in the disassembly of the nucleoli, and the effect can be reversed by adding crowding agents¹²⁶. This is consistent with the reversibility of a thermodynamically-driven phase separation. Later, Kato *et al.* showed that exposing cell or tissue lysates to a chemical known as biotinylated isoxazole will induce the formation of hydrogels that include many RNA-binding proteins. These proteins mostly contain low complexity domains, or IDRs, which are proposed to drive phase separations *in vivo*¹². Similarly, exposing *Drosophila* egg chambers to a saline solution results in the emergence of novel dynamic nuclear bodies that share several features of other membraneless organelles, such as rapid exchange of components¹²⁷. While these experiments demonstrate the ability of multi-protein mixtures or extracts to undergo phase separation, they also indicate the strong sensitivity of the phase boundary to small changes in composition. Therefore, while *in vitro* studies provide invaluable information about the underlying molecule driving forces for biopolymeric phase separation, *in vivo* studies are necessary to determine whether the assembly of individual proteins and RNAs in living cells is truly an LLPS process.

3.2 *In Vivo* Investigation of LLPS

The complex nature of living cells, which are comprised of a large number of reactive components, makes testing the validity of the LLPS model *in vivo* extremely challenging. In recent years, several studies have attempted to address this grand challenge by examining: (i) the mechanical and transport properties of intracellular assemblies, (ii) their sensitivity to changes in thermodynamic variables, and (iii) the *in vivo* kinetics of the self-assembly

process.

The majority of earlier *in vivo* studies of LLPS focus on examining the mechanical and transport properties of membraneless organelles^{128,129}. For instance, Brangwynne *et al.* reported the liquid-like behavior of P-granules, such as their propensity to wet external surfaces, and their ability to fuse with one another¹⁰. Subsequent studies on other intracellular assemblies, including nuclear bodies^{6,130–132}, and other cytoplasmic RNA granules^{107,123,133,134} demonstrated that this ability to fuse is shared by all these organelles. In addition, the dynamic nature of such organelles has been extensively studied by measuring the mobility of components that localize to those assemblies. This can, for instance, be achieved by tagging the macromolecule of interest with a fluorescent probe, and measuring the recovery rate after its photobleaching. It has been observed that the tagged macromolecules that localize into membraneless assemblies can rapidly move in and out, or within such assemblies^{59,135,136}. However, not all components of the membraneless organelles exhibit this rapid dynamic behavior. For instance, the core of stress granules, and the MEG protein that constitute the shell of P-granules are stable structures with amorphous solid- or crystal-like dynamics^{137,138}.

In addition to these mechanical properties, the sensitivity of such assemblies to changes in thermodynamic variables have also been examined *in vivo*. As discussed in section 2.1, the thermodynamic driving force for phase separation is determined both by the composition of the mixture, and temperature. Most *in vivo* studies examine how phase separation is impacted by changes in concentration, and demonstrate that changing the concentration of the phase-separating component, either genetically^{106,139,140} or optically¹⁴¹, will modulate the phase separation process. Also, globally increasing the concentration of macromolecules in HeLa cells by exposing them to a hypotonic solution, or decreasing the temperature results in the self-assembly of a germ granule protein, Ddx4⁵⁷. Interestingly, locally increasing the concentration of specific nucleolar proteins, by tagging them individually with LacI and targeting them to LacO repeats, can enrich the nucleolar proteins not tagged with LacI¹⁴². Together, these studies confirm the concentration dependence of the assembly, as expected for a thermodynamically-driven phase separation. Yet, active processes also exhibit a similar dependence of rate on concentration, rendering the findings of such studies inconclusive.

Recently, an *in vivo* assay was developed by Falahati and Wieschaus that allows for distinguishing between the LLPS model and an active assembly process for bona fide organelles⁵⁹. This assay is based on the two thermodynamic properties that collectively distinguish LLPS from an active assembly model, namely the temperature dependence and reversibility. This assay was utilized to study the mechanism by which six nucleolar proteins localize to the nucleolus. Using this approach they demonstrated that Fibrillarin, Nopp140, Pitchoune and RNA polymerase I, condense at low temperatures and dissolve at ambient temperatures in *Drosophila* embryos, confirming a reversible phase separation with a UCST that is responsive to rapid changes in temperature. Interestingly, not all nucleolar proteins in this study followed this behavior; During the initial growth stage, the assembly of Nu-

cleostemin1 and Modulo, the fly homologue of Nucléolin, was inhibited at low temperatures, did not show reversibility in response to oscillations in temperature, and became insensitive to changes in temperature when a maximum size (saturation level) was achieved. This is inconsistent with the LLPS mechanism, and can be explained by an active reaction driving the assembly.

In addition to the thermodynamic conditions, the cells also need to modulate the kinetics of condensation, to regulate when and where the assemblies would form. Many membraneless organelles such as nucleoli and histone-locus bodies form strictly at well-defined locations inside the cells, suggesting that the nucleation barrier is smaller at those select locations. Interestingly, the assembly of the phase separating components of nucleolus loses its spatiotemporal precision and exhibits a high degree of spatiotemporal variability in the absence of such nucleolar organizer regions⁷⁹. These nucleolar organizer regions are sites of ribosomal RNA transcription, and when activated, result in the transformation of the assembly process from a nucleation-limited homogeneous assembly to a growth-limited heterogeneous phase separation. Therefore, in addition to the role of ribosomal RNA in modulating the thermodynamics of nucleolus assembly¹⁴⁰, it can also play a role in the spatiotemporal regulation of the nucleolus formation by changing the kinetics of assembly⁷⁹.

4 Phase Separations in Biological Systems: Computational Investigations

Experimental studies provide convincing evidence that thermodynamically driven LLPS processes play a significant role in the assembly of membraneless organelles. The existing experimental techniques can indeed be utilized to characterize the thermodynamics and kinetics of LLPS both *in vitro* and *in vivo*. Their major disadvantage, however, is their inability to probe the molecular level events that culminate in phase separation, due to their limited spatiotemporal resolution. Similarly, experiments are not well equipped to probe the conformational ensembles and dynamics of IDPs and other phase separating proteins. In addition to these limitations, it is fairly expensive to use experiments for exploring large parameter spaces, e.g., assessing the sensitivity of phase separation propensity to changes in molecular properties and environmental conditions. Theoretical and computational approaches, however, do not have any of these limitations. They not only provide accurate molecular-level information about the conformation ensemble of proteins, and the molecular events that culminate in phase separation, but can also be used for large-scale sensitivity analysis studies. In addition, they can be utilized for probing length and timescales not otherwise accessible to experiments.

Theoretical and computational studies of LLPS can be broadly classified into two categories. In *field-based* studies, which are usually continuum in nature, the effect of neighboring molecules on each molecule is represented by an effective field that depends on spatial profiles of thermodynamic variables, such as density and composition. The free energy of the system $\mathcal{F}[\mathbf{c}(\mathbf{r})]$ is then expressed as a functional of the density and/or composition profile $\mathbf{c}(\mathbf{r})$, which is then minimized analytically or nu-

merically using variational methods, subject to proper conservation constraints^{82–84}. Within field-based framework, demixing will occur under conditions at which a non-uniform composition profile minimizes the free energy functional. Such functionals can also be utilized for constructing generalized partial differential equations for transport of individual components, making it possible to predict the spatiotemporal evolution of composition in phase separating macroscopic and mesoscopic systems. Such descriptions have been extensively used to study flow¹⁴³ and phase separation¹⁴⁴ in polymeric mixtures, and microphase separation in block copolymers¹⁴⁵. In general, these, alongside other flavors of field-based methods, such as the self-consistent field theory (SCFT)^{146,147}, have been an integral part of polymer physics in recent decades. Due to their simplicity and versatility, they are excellent tools for understanding the underlying physics of phase separation, and to assess how it is impacted by factors such as mixing enthalpy and surface tension. In recent years, field-based methods have been used for understanding LLPS in biological systems^{140,148,149}. When it comes to phase separation in a mixture of biomolecules with specific sequences, however, standard field-based methods have limited utility, as they lack the specificity and resolution needed to faithfully account for specific interactions and spatial correlations present in such aqueous biomolecular solutions. There have, however, been numerous attempts^{89,98,150} in recent years to alleviate some of these limitations, via applying approaches such as the random phase approximation (RPA) method^{151,152}. Such approaches resolve heterogeneity at a single-chain level, but still depend on the core assumption that single-chain characteristics are predictive of phase separation propensity¹⁵³. The applications of field-based approaches in understanding biomolecular phase separation has been the focus of several recent reviews^{15,16,46}, and will not be discussed further in the current paper.

In *particle-based* methods (Figure 5), which are the main focus of this section, a phase separating mixture is represented as a collection of particles, which interact according to a pre-determined potential energy function known as a *force-field*. A force-field is usually a linear combination of bonded (such as bonds, angles, dihedrals, etc) and non-bonded (e.g., Columbic, dispersion, etc) interactions. The primary advantage of particle-based methods is their ability to account for interatomic and intermolecular correlations without any *a priori* assumptions about their nature, while in field-based methods, the mathematical form of such correlations– or their lack thereof– usually needs to be explicitly incorporated into the model. Particle-based methods– also known as *molecular simulations*– are classified into two categories. While *molecular dynamics (MD)* simulations¹⁵⁴ are based on integrating Newton's equations of motion, *Monte Carlo (MC)* simulations¹⁵⁵ utilize an importance sampling stochastic scheme to generate a statistically representative sequence of configurations commensurate with the corresponding thermodynamic ensemble. Given sufficiently long sampling, both MC and MD are expected to yield identical averages for thermodynamic and structural properties if the underlying system is *ergodic*. They might, however, be different in their efficiency, i.e., the statistical uncertainty of the desired thermodynamic averages vs. the expended computational

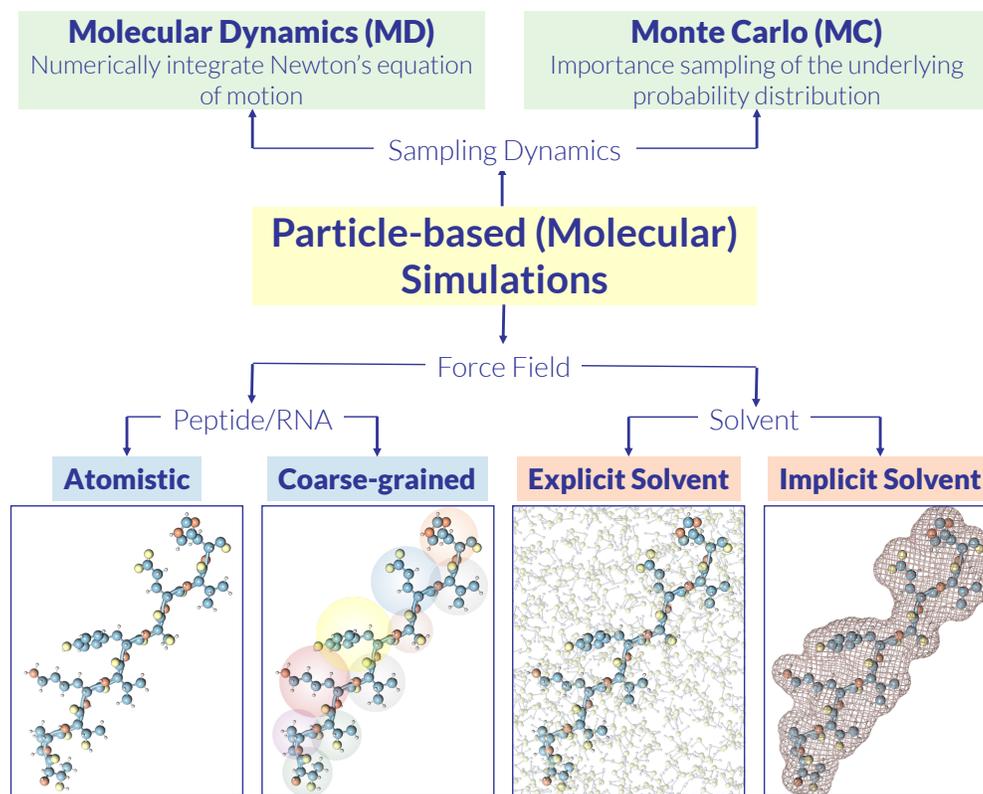


Fig. 5 Classification of particle-based (molecular) simulations of biological systems, based on the sampling dynamics and granularity of the utilized force-fields. Sampling the configuration space can be conducted using stochastic dynamics (in MC), or Newtonian dynamics (in MD). The utilized force-field for the biomolecule can be atomistic or coarse-grained. In both cases, solvent molecules can either be included explicitly in the simulation box, or their impact on the peptide/RNA chain can be represented using an effective field or modified interaction parameters. Coarse-grained models can have different levels of complexity, from each interaction site representing several atoms within a residue, to representing the entire macromolecule.

time. The major advantage of MD is its ability to predict dynamical properties, such as transport coefficients and nucleation kinetics, a task that cannot be achieved with MC due to lack of a rigorous mapping of MC moves to actual time. MC, however, is advantageous when the underlying force field is discontinuous, or when unphysical trial moves are utilized to enhance the sampling efficiency in slowly relaxing systems. Examples include particle swaps in multiphase systems or mixtures¹⁵⁶, and configurational bias approaches in simulations of polymers and peptides¹⁵⁷. Such unphysical moves are widely used in studying the conformational space of peptides, and the thermodynamics of biomolecular LLPS.

Molecular simulations can also be classified based on the resolution of the utilized force-fields. The most detailed force-fields are *atomistic* force-fields^{158–161}, in which every atom is represented by one– or sometimes several– interaction sites. Atomistic simulations of biomolecular systems, however, are computationally expensive, and it is extremely difficult– if not impossible– to simulate systems with more than a million atoms, even with specialized hardware and massive parallelization^{162–164}. There are two not mutually exclusive approaches for tackling these limitations. *Coarse-grained* force-fields expand the range of accessi-

ble length and timescales by averaging out atomistic details, and representing groups of atoms (e.g., residues in a peptide chain) as single interaction sites^{165,166}. In *implicit-solvent* force-fields, which are used in both atomistic and coarse-grained simulations, solvent molecules are excluded, and their impact is accounted for with an effective field, or through modifying interaction parameters between the biomolecules¹⁶⁷. As will be discussed later, coarse-grained force-fields can have different levels of granularity to the extent that a single interaction site can represent anything from several atoms within a residue, to the entire biomolecule.

One of the major limitations of conventional molecular simulations– irrespective of the utilized force-field– is their inability to efficiently probe rare events, such as low-probability conformational rearrangements (e.g., protein folding) and nucleation. For that purpose, a wide range of *advanced sampling techniques* have been developed, which expand the range of accessible timescales. Advanced sampling techniques can be classified into two categories. In *bias-based* methods, such as umbrella sampling¹⁶⁸, flat histogram methods^{169,170} and metadynamics^{171,172} the underlying Hamiltonian is modified in order to preferentially favor certain configurations. In *path sampling*

methods, such as transition path sampling¹⁷³, forward-flux sampling^{174–176}, transition interface sampling^{177,178} and parallel tempering^{179–181}, the underlying Hamiltonian is not modified, and trajectories are instead sampled in a targeted manner.

The remainder of this section will be dedicated to discussing the types of information that molecular simulations can provide about LLPS in biological systems. The studies discussed here utilize a wide variety of simulation techniques, and can be conceptually categorized into two groups. The first group deals with IDPs as main culprits for LLPS in biological systems, and explores their thermodynamics, conformational ensembles and dynamics. The second group, however, directly deals with the question of biological liquid-liquid phase separation, and understanding its thermodynamics and kinetics.

4.1 Computational Investigation of IDPs, the Main Culprits of Biological LLPS

While IDPs and proteins with IDRs have been known for a long time, the origin of their disorder still remains elusive. This question is closely related to a decades-old problem in polymer physics, i.e., the question of what dictates a polymer's conformation in solution. The conditions under which a polymer will collapse onto a compact globule have been extensively studied, and it is generally understood that it is rare for a heteropolymer with a random sequence of solvophobic and solvophilic residues to simultaneously fold into a globular structure and remain solvated^{182,183}. In order for such folding to occur, a tight balance needs to be established between the entropic loss due to adopting a folded structure, and the enthalpic gain due to favorable energetic interactions among the core solvophobic residues, and among the solvophilic residues and the solvent. Such a condition is usually not met for a random heteropolymer¹⁸⁴. Peptides with well-defined folded structures therefore constitute remarkable anomalies. This is not surprising as naturally occurring peptides are not random heteropolymers, and instead have sequences naturally selected through evolution. Understanding the relationship between sequence and folding propensity has been the focus of multiple studies. In a pioneering work, Uversky *et al.*¹⁸⁵ demonstrated that the propensity of a peptide to form well-defined folded structures is dictated by its mean per-residue net charge, and the average hydrophobicity of its constituent residues, and the fact that IDPs do not have well-defined folded structures is due to their high net charge and low hydrophobicity. This minimal model, however, cannot be fully predictive, as folding propensity can be strongly impacted by subtle features such as the linear distribution of charged residues along a sequence¹⁸⁶. Therefore, even though sequence-based methods for predicting folding propensity have become increasingly accurate over years^{187–191}, they are not well-equipped to properly account for all such nuances. Moreover, such approaches are usually incapable of providing molecular-level information about the conformational dynamics of the peptides that they might correctly detect as intrinsically disordered.

In recent decades, molecular simulations have emerged as attractive tools for inspecting the origin of IDP disorder, their con-

formational dynamics and their binding propensity. The predictive power of such studies were, however, limited until recently, primarily due to the tendency of most classical protein force fields to overestimate the formation of secondary structures. This problem is indeed universal across the board, and affects both explicit- and implicit-solvent atomistic and coarse-grained force-fields. This issue has, however, been addressed¹⁹² in more recent all-atom force-fields, e.g., by strengthening water-protein interactions^{193–195} and utilizing more realistic water models¹⁹⁶. Also, newer force-fields, such as the Kirkwood-Buff force field (KBFF)¹⁹⁷, have been developed with a focus on solution properties. Similar improvements have been made to implicit-solvent¹⁶⁷ and coarse-grained¹⁹⁸ models.

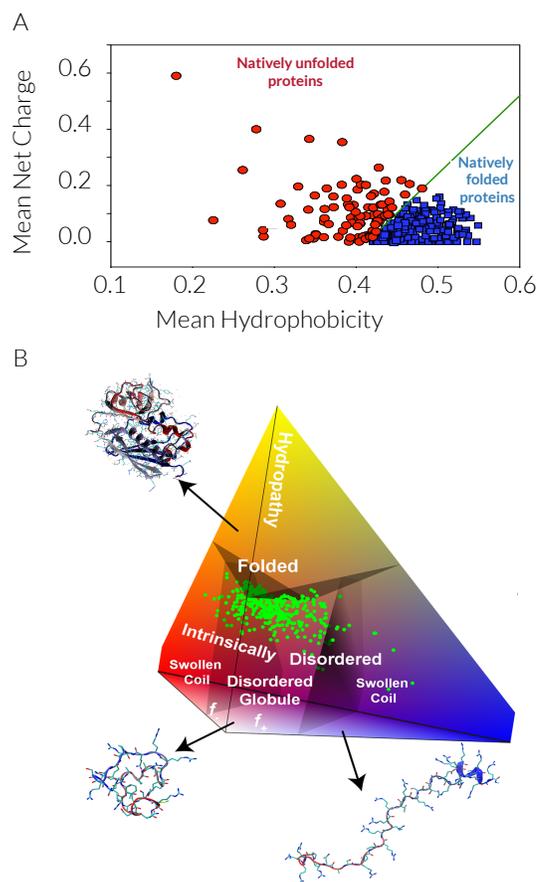


Fig. 6 (A) (Adapted and relabeled from Ref. 199) The original hydrophobicity/mean charge correlation of Uversky *et al.* (B) (Reproduced from Ref. 186) The three-dimensional correlation of Mao *et al.* that describes the behavior of IDPs based on the fraction of positively and negatively charged residues.

One of the most pressing question about IDPs is the extent to which their disorder is impacted by factors such as amino acid sequence, temperature, and the presence of inert crowders. A powerful way to address these questions is to utilize highly coarse-grained force-fields in which distinct amino acids with similar properties (e.g., charge, hydrophobicity) are encoded into a single type of interaction site. Despite not being sequence-specific, such models qualitatively capture the underlying physics

of interactions between a peptide and solvent molecules. One major advantage of using such models is the computational efficiency at which "generic" peptides with a sufficiently wide range of desired solvation and electrostatic properties can be simulated. For instance, the original idea of Uversky *et al.*¹⁸⁵ was explored by Ashbaugh and Hatch²⁰⁰ using a simple bead-spring model of peptides in which each residue can only be of three types: hydrophobic, polar, and positively charged polar. They simulated sequences with a wide range of lengths (N), net charges (q) and hydrophobicities (H), with 15-20 random peptides generated for each N , q and H . Even though they observed that coil-like and collapsed globules occupy distinct regions of the (q, H) space, the boundary between them only becomes N -independent (as predicted by Uversky) when explicit counterions are included for charged residues (vs. utilizing screened charges). A similar approach was utilized by Miller *et al.*²⁰¹ for probing the impact of crowders on the compactness of peptides (e.g., their radii of gyration). In the presence of inert (i.e., repulsive) crowders, they observed that coil-like peptides become more compact, while no significant increase in compactness was observed for globular peptides. In other words, crowding can be used as a means of modulating IDP compactness in biological systems. Note that it is not feasible to use experiments or atomistic molecular simulations to explore the effect of crowders on the conformational ensembles of IDPs, and this is a question that can only be addressed using such highly coarse-grained models.

Despite their usefulness in elucidating the qualitative behavior of IDPs, such coarse-grained models are not suitable for making quantitative predictions about the conformational dynamics and binding propensities of a particular IDP. Such detailed information about IDPs can usually be obtained from atomistic simulations only, which, thanks to recent advancements in computer architecture, have become faster and more popular in recent years²⁰². For instance, numerous atomistic studies have been conducted with the aim of understanding the effect of charge distribution on disorder propensity. As an example, Mao *et al.* utilized the ABSINTH implicit-solvent model¹⁶⁷ and MC simulations to explore the conformational ensemble of peptides with low hydrophobicity²⁰³. As predicted by the Uversky *et al.*'s hydrophobicity/mean net charge correlation¹⁸⁵, such peptides tended to remain unfolded. Depending on the fraction of positively and negatively charged residues, however, they exhibited swollen coil, or disordered globule conformations. In other words, different IDPs can have different levels of chain compactness, which is key to their biological function. IDP conformation can also be impacted by charge decoration, or the linear distribution of positively and negatively charged residues along the sequence, as demonstrated in a recent work by Firman and Ghosh²⁰⁴. They derived an analytical theory for coil-to-globule transition in heteropolymers, and calibrated it against findings of atomistic MC simulations of IDPs in the DisProt²⁰⁵ database. They observed that IDPs with identical f_+ 's and f_- 's can exhibit different conformations, depending on factors such as charge decoration, and post-translational modifications. The role of charge decoration on phase separation propensity has also been demonstrated in on-lattice MC²⁰⁶ and coarse-grained MD²⁰⁷ simulations. Note that it is nontrivial

to account for such effects in sequence-based methods, demonstrating the importance of atomistic simulations in accessing the conformational ensemble of IDPs.

An interesting phenomenon that has been recently studied using molecular simulations is the collapse of some unfolded peptides— including some IDPs— upon increasing temperature^{208–210}. This behavior is in contrast to the tendency of most polymers to swell upon heating²¹¹, and has been attributed to temperature-dependent per-residue hydration energies for such peptides^{212–214}. For instance, Wuttke *et al.*²¹² investigate chain compaction in five proteins that can exist in an unfolded state under physiologically relevant conditions using FRET experiments, theory and molecular simulations. Using a theoretical description proposed by Sanchez²¹¹, they explained the observed heat-induced compaction using a T -dependent effective monomer-monomer interaction parameter. They also conducted molecular simulations using the implicit-solvent ABSINTH model, and confirmed that the heat-induced chain compaction will only occur if the temperature dependence of per-residue hydration energies are explicitly included in the ABSINTH model. Later studies using explicit-solvent force-fields confirm such compaction^{213,214}, and the-dependence of solvent accessible surface area (SASA) on temperature²¹³. From a molecular perspective, this effect can be attributed to the preponderance of nascent structures in unfolded peptides, which makes them distinct from random-coil polymers^{213,214}. Note that it is extremely difficult to attain such granular residue-level information using experiments, and yet such information are pivotal for understanding IDP properties and function^{215,216}. Due to the rugged free energy landscape of the large peptides studied in Refs. 212–214, advanced sampling techniques such as replica exchange molecular dynamics^{217–219} were utilized for generating statistically representative peptide conformations.

Molecular simulations have also proven useful in exploring the effect of probes and chromophores on the conformational ensembles and dynamics of IDPs, an information critical to interpreting data obtained from probe-based techniques. One of these techniques is the Förster resonance energy transfer (FRET)²²⁰ in which two chromophores, one energy donor and one energy acceptor, are covalently placed at select locations along a peptide (usually at the C and N termini). Energy is then transferred from the excited donor to the ground-state acceptor through dipole-dipole coupling. The efficiency of such energy transfer will be proportional to r^{-6} , with r the distance between the two probes. FRET energy transfer efficiency is therefore a sensitive measure of the distance between the two probes, and is widely used for exploring the conformational ensemble and dynamics of IDPs. In order to properly interpret FRET findings, it is necessary to make sensible assumptions about the impact of probes on the conformational ensemble, dynamics, and long-range contact distribution of IDPs. In recent years, this question has been extensively studied using molecular simulations^{221–226}. For instance, Zerze *et al.* investigated²²¹ the effect of fluorophores on the compactness, secondary structure content, and long-range contact distribution of three unfolded proteins (the cold shock protein, CSP) from *Thermotoga maritima*, the DNA-binding domain of l-repressor (LR),

and the N-terminal domain of HIV integrase (IN), using replica exchange MD²²⁷, and they did not find such structural metrics to be strongly impacted by the fluorophores utilized in FRET experiments. Since then, numerous other computational studies have attempted to characterize different aspects of FRET experiments^{222–226}.

4.2 Computational Investigation of Biological LLPS

The computational studies discussed in Section 4.1 involve simulating one (or a handful of) IDP chains, and can therefore be conducted using fully atomistic force-fields. It is, however, far more expensive computationally to study the thermodynamics and kinetics of LLPS for the following reasons. First of all, it is not generally possible to obtain a realistic picture of LLPS without simulating systems comprised of at least several hundred IDPs (or other phase separating biomolecules), a daunting and computationally intractable task if atomistic force-fields are to be utilized. Developing accurate coarse-grained force-fields of proteins is therefore a prerequisite to studying the thermodynamics and kinetics of LLPS in biological systems. Secondly, depending on the magnitude of the thermodynamic driving force, LLPS can be nucleation-limited. Crossing the nucleation barrier will be a rare event under such circumstances, and will occur over timescales much longer than what would take for the conformational rearrangement of individual chains. As discussed earlier, a wide variety of advanced sampling techniques have been developed for studying rare events in recent decades, which need to be used for studying the kinetics and free energy landscape of nucleation-limited LLPS in biological systems.

Due to these complexities, very few computational studies of the actual LLPS process have been conducted. Most such studies employ highly coarse-grained implicit-solvent nonspecific models of proteins, and are thus only designed to provide qualitative information about the thermodynamics and kinetics of phase separation in protein-like systems. Historically, studies of biological phase separation predate recent interest in LLPS, and, like experimental efforts, were originally conducted with the aim of understanding the crystallization of globular proteins. In such representations, individual proteins are treated as colloidal particles interacting via isotropic²²⁸ or patchy^{229,230} potentials. The interest in probing the thermodynamics and kinetics of phase transitions in patchy colloidal systems^{231–236}, however, extends far beyond understanding protein crystallization, as patchy colloids are excellent model systems for inspecting how a competition between isotropic and directional interactions can impact the phase behavior, e.g., the relative stability of different crystals and liquids, as well as the kinetics of self-assembly.

In recent years, several qualitative studies of biological LLPS have been conducted, mostly employing simple on-lattice models. For instance, Jacobs and Frenkel²³⁷ used grand canonical Monte Carlo (GCMC)²³⁸ to simulate an N -component lattice gas, in which the contribution of two adjacent occupied sites of types i and j to the total energy was given by ϵ_{ij} . Inspired by the pioneering theoretical work of Sear and Cuesta⁵⁸, who had used random matrix theory to identify conditions for LLPS in many-component

biological systems, they drew ϵ_{ij} 's from a Gaussian distribution of pre-specified mean and variance, and observed that de-mixing will occur if the standard deviation in interaction strength exceeds a well-defined threshold. This study provides some clarity to the otherwise hopeless question of LLPS *in vivo*, and stipulates that the propensity of a complex biomolecular mixture to undergo LLPS can be tuned by changing the variance of interaction strength between its constituent components.

Despite their utility in providing insight into how demixing propensity is impacted by factors such as interaction strength, such coarse-grained models cannot be quantitatively predictive when a particular protein is concerned, and are particularly unsuitable for studying IDPs due to their poorly defined generic geometries and conformational ensembles. In recent years, however, there have been major advancements^{198,239} in developing accurate and computationally efficient coarse-grained protein force-fields in which each amino acid is represented using a single interaction site, and solvent molecules are only considered implicitly. Also, charged residues that are prevalent in IDPs interact via screened electrostatic potentials^{240,241}, as predicted by the Debye-Hückel theory²⁴². Therefore, due to the short-range nature of such screened electrostatic interactions, there is no need to utilize computationally costly methods such as the Ewald sum²⁴³. This latter fact makes such coarse-grained representations even faster. It is therefore now feasible to simulate hundreds of IDP chains on existing computer architectures. This has resulted in a few studies of direct coexistence of phase separating IDPs. For instance, Dignon *et al.*¹⁹⁸ utilize the slab method^{244,245} to study the phase separation of two proteins, i.e., the low complexity domain of the RNA-binding protein FUS and the DEAD-box helicase protein LAF-1. In the slab method, the coexisting densities and compositions are accurately determined by placing a slab of the peptide-rich phase in contact with vacuum, which represents the solvent-rich phase. By doing so, the thermodynamics of phase separation can be systematically investigated without a need to cross the nucleation barrier. The question of interest in Dignon *et al.*'s work, in particular, was to understand how LLPS is qualitatively affected by FUS mutations, and the presence– or lack thereof– folded domains in LAF-1. In a separate study²⁴⁶, they utilized their coarse-grained force-field alongside the slab method to demonstrate that there is a close correlation between T_C , the critical demixing temperature, T_θ the temperature at which coil-to-globular transition occurs for an isolated chain, and T_B , the Boyle temperature at which the second virial coefficient vanishes. They established such correlation by simulating 20 different IDP sequences with a wide range of properties. This is a very important finding and demonstrates that properties such as T_B and T_θ that can be determined from observing the behavior of isolated chains can be predictive of the propensity of a dense mixture of such chains to undergo phase separation. These findings prove the applicability of the random phase approximation method that assumes the existence of such a correlation¹⁵³.

5 Emerging Questions and Path Forward

Thermodynamically driven phase separations provide an energetically inexpensive means for cells to concentrate certain proteins

and RNAs. Such structures can be liquid-like or solid-like, which determines the types of biological functions that they undertake. For instance, the fact that membraneless organelles are liquid-like allows for fast diffusion of components including substrates and products in and out of such assemblies. Crystalline or amorphous solid assemblies, however, can provide structural support to the cell. They can also attract and retain certain molecules from the pool of readily available cellular components, e.g., in the case of stress. Over all, the phase separation model has been successful in explaining certain aspects of the assembly and function of membraneless organelles. Yet, the complexity of biological systems makes it extremely difficult to validate and characterize the occurrence of LLPS processes *in vivo*.

One such complexity arises from the active nature of living organisms. Despite growing evidence in recent years in favor of the LLPS model, the contribution of active reactions to the formation and disassembly of membraneless organelles *in vivo* cannot be disregarded^{119–121,137,247–250}. For instance, depletion of ATP completely blocks the assembly of stress granules, whose proteome contains a large number of ATP-dependent helicases and protein remodelers¹³⁷. In principle, the large number of active reactions, such as RNA transcription, that occur within membraneless organelles, can be sufficient to locally concentrate macromolecules, and drive the assembly of membraneless organelles^{251–255}. For instance, the local accumulation of the products of such functional reactions can drive the recruitment of their binding partners, such as processing factors, a process that does not need to be a thermodynamically driven phase separation^{79,140}. Another active process that can culminate in the assembly of membraneless organelles is the active transport of their constituent components. For instance, the formation of stress granules and the growth of P-bodies in response to stress, rely on motor proteins^{248,249}. Interestingly, the formation of high concentration assemblies with liquid-like properties have also been observed in active swimmers whose motion is powered by a chemical reaction even in the absence of any attractive forces²⁵⁶. Finally, biochemical processes such as transcription, translation, and post-translational modifications, that result in the formation of the components of membraneless organelles, are all comprised of active chemical reactions that will proceed only in the presence of ATP. While the contribution of such active processes to the formation of membraneless organelles can be through modulation of LLPS, they can also function as the sole driving force for the formation of membraneless organelles. In other words, while each model can explain certain aspects of the membraneless organelle assembly, distinguishing between a solely-active assembly mechanism, and a combination of active and LLPS is particularly challenging and requires a more extensive thermodynamic characterization of the macromolecules that localize to such organelles *in vivo*. In recent years, numerous phenomenological continuum models have been proposed for understanding the role of reactions and concentration gradients in phase separation^{13,257–261}. Developing additional testable theoretical models based on the predictions of the two mechanisms are pivotal for discerning the potential contribution of each mechanism to a particular assembly process.

A second complication arises from the complexity of biomolecules, and the difficulty of probing all timescales relevant to their conformational rearrangements and assembly. Experimentally, this can, in principle, be addressed by using probe-based techniques such as FRET^{115,262,263}, or nuclear magnetic resonance (NMR) spectroscopy^{56,112,115}. Unfortunately, these techniques do not have the sufficient resolution for realtime probing of biomolecular conformations. Developing advanced spectroscopy and microscopy techniques with increased resolution will be particularly helpful in this regard. One possible avenue is to re-engineer techniques such as femtosecond ultrafast scattering^{264,265} and four-dimensional electron microscopy²⁶⁶ to be compatible with biomolecular systems. In addition, this is an area that can benefit immensely from molecular simulations. Despite major breakthroughs in recent decades, we are only witnessing the beginning of using molecular simulations to investigate IDPs and biological LLPS. With more accurate force-fields, more accurate and efficient advanced sampling techniques, and better and faster hardware, it is now possible to investigate conformational rearrangements of a wide range of proteins, as well as their propensity to bind other proteins and ligands. Moreover, recent advancements in systematic coarse-graining have made it feasible to explore biological LLPS in real time. In particular, using such coarse-grained force-fields in conjunction with advanced sampling techniques can provide us with accurate information about the kinetics and mechanism of phase separation, and thus help us address deep fundamental questions about proteins and protein folding and assembly. One of the most pressing questions is the role and importance of hydrophobicity in inducing biological self-assembly. This question has fascinated statistical physicists and biophysicists for decades^{182,267,268}, and has been recently investigated for simpler hydrophobic models using extensive molecular simulations and path sampling techniques^{269–273}. Nevertheless, the precise kinetics and mechanism of hydrophobic assembly in biological systems is far from fully understood, and considering the power of state-of-the-art advanced sampling techniques in elucidating the mechanism of other rare events such as crystallization^{274–280}, they can be very useful in understanding the hydrophobic effect. IDPs are excellent systems in this regard, as they tend to phase separate and assemble despite having low hydrophobicity, and understanding the molecular origins of such behavior is critical to unraveling the longstanding question of hydrophobicity and biological assembly.

Finally, an important challenge in studying LLPS in biological systems is the difficulty of controlling thermodynamic variables (such as temperature, concentration, pH) and operating conditions (such as hydrodynamic shear) at a cellular level. Such a challenge not only makes it difficult to obtain reliable data about the kinetics of biomolecular assembly *in vivo*, but also makes a robust analysis of the sensitivity of LLPS to such factors challenging. In recent years, optogenetic and microfluidic approaches have been successfully utilized to address this issue. For instance, optogenetic approaches can be used for spatial modulation of concentration locally and globally within cells^{141,281}. Another powerful approach that has gained increased popularity in the soft matter community in recent years is microfluidics, which allows for swift

and precise control of operating variables such as temperature, pressure, concentration, pH and shear deformation. This feature makes microfluidics particularly ideal for rate measurements, as they have been successfully used for probing nucleation kinetics in several systems^{282–285}. Microfluidic devices were recently employed for studying the impact of rapid changes in temperature on the *in vivo* assembly of nucleolar proteins in *Drosophilla* embryos⁵⁹. Developing similar experimental techniques that will enable us to explore the molecular driving forces and kinetics of LLPS can be another potential area of future exploration.

Acknowledgements

We are grateful to E. Wieschaus for productive and insightful discussions. We acknowledge E. Wieschaus and R. Samanta for their help in the production of Figure 1B. A.H.-A. gratefully acknowledges the support of the National Science Foundation CAREER Award (Grant No. CBET-1751971). H.F. is a Howard Hughes Medical Institute Fellow of the Life Sciences Research Foundation (LSRF).

Notes and references

- 1 D. Frenkel, *Physica A*, 1999, **263**, 26–38.
- 2 R. P. Sear, *Soft Matter*, 2007, **3**, 680.
- 3 K. R. Porter, A. Claude and E. F. Fullam, *J. Exp. Med.*, 1945, **81**, 233–246.
- 4 C. Benda, *Arch. Anat. Physiol.*, 1898, **73**, 393–398.
- 5 S. Castro-Obrigón, *Nat. Ed.*, 2010, **3**, 49.
- 6 T. S. H. Montgomery, *J. Morphol.*, 1898, **15**, 265–582.
- 7 D. D. Brown and J. Gurdon, *Proc. Natl. Acad. Sci. USA*, 1964, **51**, 139–146.
- 8 S. R. Cajal *et al.*, *Trab. Lab. Invest. Biol. (Madrid)*, 1903, **2**, 129–221.
- 9 E. Gutierrez-Beltran, P. N. Moschou, A. P. Smertenko and P. V. Bozhkov, *Plant Cell*, 2015, **27**, 926–943.
- 10 C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoegge, J. Gharakhani, F. Jülicher and A. A. Hyman, *Science*, 2009, **324**, 1729–1732.
- 11 P. Li, S. Banjade, H.-C. Cheng, S. Kim, B. Chen, L. Guo, M. Llaguno, J. V. Hollingsworth, D. S. King, S. F. Banani, P. S. Russo, Q.-X. Jiang, B. T. Nixon and M. K. Rosen, *Nature*, 2012, **483**, 336–340.
- 12 M. Kato, T. W. Han, S. Xie, K. Shi, X. Du, L. C. Wu, H. Mirzaei, E. J. Goldsmith, J. Longgood, J. Pei, N. V. Grishin, D. E. Frantz, J. W. Schneider, S. Chen, L. Li, M. R. Sawaya, D. Eisenberg, R. Tycko and S. L. McKnight, *Cell*, 2012, **149**, 753 – 767.
- 13 A. A. Hyman, C. A. Weber and F. Jülicher, *Ann. Rev. Cell. Dev. Biol.*, 2014, **30**, 39–58.
- 14 S. F. Banani, H. O. Lee, A. A. Hyman and M. K. Rosen, *Nat. Rev. Mol. Cell Biol.*, 2017, **18**, 285–298.
- 15 J. Berry, C. P. Brangwynne and M. Haataja, *Rep. Prog. Phys.*, 2018, **81**, 046601.
- 16 C. A. Weber, D. Zwicker, F. Jülicher and C. F. Lee, *arXiv:1806.09552*, 2018.
- 17 W. M. Aumiller Jr, B. W. Davis and C. D. Keating, in *Int. Rev. Mol. Cell. Bio.*, Elsevier, 2014, vol. 307, pp. 109–149.
- 18 A. A. Hyman and K. Simons, *Science*, 2012, **337**, 1047–1049.
- 19 Y. Shin and C. P. Brangwynne, *Science*, 2017, **357**, eaaf4382.
- 20 E. M. Courchaine, A. Lu and K. M. Neugebauer, *EMBO J.*, 2016, **35**, 1603–1612.
- 21 A. Aguzzi and M. Altmeyer, *Trends Cell Biol.*, 2016, **26**, 547 – 558.
- 22 S. C. Weber, *Curr. Opin. Cell Biol.*, 2017, **46**, 62 – 71.
- 23 T. Mittag and R. Parker, *J. Mol. Biol.*, 2018, **430**, 4636 – 4649.
- 24 M. Ramaswami, J. P. Taylor and R. Parker, *Cell*, 2013, **154**, 727–736.
- 25 A. Bah and J. D. Forman-Kay, *J. Biol. Chem.*, 2016, **291**, 6696–6705.
- 26 F. Chen, P. W. Tillberg and E. S. Boyden, *Science*, 2015, **347**, 543–548.
- 27 D. Lingwood and K. Simons, *Science*, 2010, **327**, 46–50.
- 28 M. S. Steinberg, *Science*, 1963, **141**, 401–408.
- 29 D. Duguay, R. A. Foty and M. S. Steinberg, *Dev. Biol.*, 2003, **253**, 309 – 323.
- 30 T. Baumgart, A. T. Hammond, P. Sengupta, S. T. Hess, D. A. Holowka, B. A. Baird and W. W. Webb, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 3165–3170.
- 31 M. B. Stone, S. A. Shelby, M. F. Nunez, K. Wisser and S. L. Veatch, *eLife*, 2017, **6**, e19891.
- 32 F. A. Heberle and G. W. Feigenson, *CHS Perspect. Biol.*, 2011, a004630.
- 33 R. A. Foty and M. S. Steinberg, *Dev. Biol.*, 2005, **278**, 255 – 263.
- 34 L. Addadi and S. Weiner, *Angew. Chem. Int. Ed.*, 1992, **31**, 153–169.
- 35 W. Halliburton, *J. Physiol.*, 1887, **8**, 133–202.
- 36 F. Straub and G. Feuer, *Biochim. Biophys. Acta*, 1950, **4**, 455–470.
- 37 R. C. Weisenberg, *Science*, 1972, **177**, 1104–1105.
- 38 M. Stefani and C. M. Dobson, *J. Mol. Med.*, 2003, **81**, 678–699.
- 39 C. A. Ross and M. A. Poirier, *Nat. Med.*, 2004, **10**, S10–S17.
- 40 M. S. Forman, J. Q. Trojanowski and V. M.-Y. Lee, *Nat. Med.*, 2004, **10**, 1055–1063.
- 41 J. Hardy and D. J. Selkoe, *Science*, 2002, **297**, 353–356.
- 42 C. E. Shaw, A. Al-Chalabi and N. Leigh, *Curr. Neurol. Neurosci.*, 2001, **1**, 69–76.
- 43 L. P. Rowland and N. A. Shneider, *New Eng. J. Med.*, 2001, **344**, 1688–1700.
- 44 M. Baba, S. Nakajo, P. H. Tu, T. Tomita, K. Nakaya, V. M. Lee, J. Q. Trojanowski and T. Iwatsubo, *Am. J. Pathol.*, 1998, **152**, 879–884.
- 45 G. B. Benedek, *Invest. Ophth. Vis. Sci.*, 1997, **38** **10**, 1911–21.
- 46 Y.-H. Lin, J. D. Forman-Kay and H. S. Chan, *Biochemistry*, 2018, **57**, 2499–2508.
- 47 R. Aris and R. Mah, *Ind. Eng. Chem.*, 1963, **2**, 90–94.

- 48 D. Barlow, *J. Phys. Chem. Solids*, 2013, **74**, 406–409.
- 49 M. Dijkstra, R. van Roij and R. Evans, *Phys. Rev. Lett.*, 1998, **81**, 2268.
- 50 T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert and R. Riek, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 17342–17347.
- 51 A. McPherson, *Methods*, 2004, **34**, 254–265.
- 52 A. McPherson and J. A. Gavira, *Acta Cryst.*, 2014, **70**, 2–20.
- 53 I. M. Hodge, *Science*, 1995, **267**, 1945–1947.
- 54 M. Utz, P. G. Debenedetti and F. H. Stillinger, *Phys. Rev. Lett.*, 2000, **84**, 1471.
- 55 L. Cipelletti, S. Manley, R. Ball and D. Weitz, *Phys. Rev. Lett.*, 2000, **84**, 2275.
- 56 K. A. Burke, A. M. Janke, C. L. Rhine and N. L. Fawzi, *Mol. Cell*, 2015, **60**, 231 – 241.
- 57 T. J. Nott, E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowitz, T. D. Craggs, D. P. Bazett-Jones, T. Pawson, J. D. Forman-Kay and A. J. Baldwin, *Mol. Cell*, 2015, **57**, 936 – 947.
- 58 R. P. Sear and J. A. Cuesta, *Phys. Rev. Lett.*, 2003, **91**, 245701.
- 59 H. Falahati and E. Wieschaus, *Proc. Natl. Acad. Sci. USA*, 2017, **114**, 1335–1340.
- 60 H. Jiang, S. Wang, Y. Huang, X. He, H. Cui, X. Zhu and Y. Zheng, *Cell*, 2015, **163**, 108 – 122.
- 61 J. M. Prausnitz, R. N. Lichtenthaler and E. G. de Azevedo, *Molecular thermodynamics of fluid-phase equilibria*, Prentice-Hall PTR, Upper Saddle River, NJ, 1998.
- 62 J. N. Glosli and F. H. Ree, *Phys. Rev. Lett.*, 1999, **82**, 4659.
- 63 Y. Katayama, T. Mizutani, W. Utsumi, O. Shimomura, M. Yamakata and K.-i. Funakoshi, *Nature*, 2000, **403**, 170.
- 64 J. C. Palmer, F. Martelli, Y. Liu, R. Car, A. Z. Panagiotopoulos and P. G. Debenedetti, *Nature*, 2014, **510**, 385–388.
- 65 J. C. Palmer, A. Haji-Akbari, R. S. Singh, F. Martelli, R. Car, A. Z. Panagiotopoulos and P. G. Debenedetti, *J. Chem. Phys.*, 2018, **148**, 137101.
- 66 S. Minomura and H. Drickamer, *J. Phys. Chem. Solids*, 1962, 451–456.
- 67 P. Bolhuis and D. Frenkel, *Phys. Rev. Lett.*, 1994, **72**, 2211.
- 68 K. Jacobs, D. Zaziski, E. C. Scher, A. B. Herhold and A. P. Alivisatos, *Science*, 2001, **293**, 1803–1806.
- 69 K. Winkel, M. S. Elsaesser, E. Mayer and T. Loerting, *J. Chem. Phys.*, 2008, **128**, 044510.
- 70 A. Haji-Akbari, M. Engel and S. C. Glotzer, *J. Chem. Phys.*, 2011, **135**, 194101.
- 71 A. Haji-Akbari, M. Engel and S. C. Glotzer, *Phys. Rev. Lett.*, 2011, **107**, 215702.
- 72 Y. Peng, F. Wang, Z. Wang, A. M. Alsayed, Z. Zhang, A. G. Yodh and Y. Han, *Nat. Mater.*, 2015, **14**, 101.
- 73 M. Avrami, *J. Chem. Phys.*, 1939, **7**, 1103–1112.
- 74 P. W. Voorhees, *J. Stat. Phys.*, 1985, **38**, 231–252.
- 75 M. Volmer and H. Flood, *Z. Phys. Chem.*, 1934, **170**, 273–285.
- 76 R. Becker and W. Döring, *Ann. d. Physik*, 1935, **416**, 719–752.
- 77 D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergström and H. Cölfen, *Chem. Soc. Rev.*, 2014, **43**, 2348–2371.
- 78 D. Zahn, *ChemPhysChem*, 2015, **16**, 2069–2075.
- 79 H. Falahati, B. Pelham-Webb, S. Blythe and E. Wieschaus, *Curr. Biol.*, 2016, **26**, 277 – 285.
- 80 D. Turnbull, *J. Chem. Phys.*, 1950, **18**, 198–203.
- 81 A. Haji-Akbari and P. G. Debenedetti, *J. Chem. Phys.*, 2017, **147**, 060901.
- 82 J. W. Cahn and J. E. Hilliard, *J. Chem. Phys.*, 1958, **28**, 258–267.
- 83 J. W. Cahn, *J. Chem. Phys.*, 1959, **30**, 1121–1124.
- 84 J. W. Cahn and J. E. Hilliard, *J. Chem. Phys.*, 1959, **31**, 688–699.
- 85 J. S. Huang, W. I. Goldburg and A. W. Bjerkaas, *Phys. Rev. Lett.*, 1974, **32**, 921.
- 86 M. R. Mruzik, F. F. Abraham and G. Pound, *J. Chem. Phys.*, 1978, **69**, 3462–3467.
- 87 F. F. Abraham, S. Koch and R. C. Desai, *Phys. Rev. Lett.*, 1982, **49**, 923.
- 88 S. Koch, R. C. Desai and F. F. Abraham, *Phys. Rev. A*, 1983, **27**, 2152.
- 89 Y.-H. Lin, J. Song, J. D. Forman-Kay and H. S. Chan, *J. Mol. Liq.*, 2017, **228**, 176–193.
- 90 A. S. Holehouse and R. V. Pappu, *Biochemistry*, 2018, **57**, 2415–2423.
- 91 M. Machyna, P. Heyn and K. M. Neugebauer, *WIREs: RNA*, 2013, **4**, 17–34.
- 92 Z. Nizami, S. Deryusheva and J. G. Gall, *CHS Perspect. Biol.*, 2010, **2**, a000653.
- 93 Y. R. Li, O. D. King, J. Shorter and A. D. Gitler, *J. Cell Biol.*, 2013, **201**, 361–372.
- 94 P. Heyn, H. Salmonowicz, J. Rodenfels and K. M. Neugebauer, *RNA Biol.*, 2016, **0**, 1–9.
- 95 E. M. Langdon, Y. Qiu, A. Ghanbari Niaki, G. A. McLaughlin, C. Weidmann, T. M. Gerbich, J. A. Smith, J. M. Crutchley, C. M. Termini, K. M. Weeks, S. Myong and A. S. Gladfelter, *Science*, 2018, eaar7432.
- 96 L. Recher, J. Whitescarver and L. Briggs, *J. Ultra Mol. Struct. R.*, 1969, **29**, 1–14.
- 97 M. Feric, N. Vaidya, T. S. Harmon, D. M. Mitrea, L. Zhu, T. M. Richardson, R. W. Kriwacki, R. V. Pappu and C. P. Brangwynne, *Cell*, 2016, **165**, 1686 – 1697.
- 98 Y.-H. Lin, J. P. Brady, J. D. Forman-Kay and H. S. Chan, *New J. Phys.*, 2017, **19**, 115003.
- 99 G. Pianese, *Beitrag zur Histologie und Aetiologie des Carcinoms*, G. Fischer, 1896, vol. 1.
- 100 D. Ruggero, *Sci. Signal.*, 2012, **5**, pe38.
- 101 A. Patel, H. O. Lee, L. Jawerth, S. Maharana, M. Jahnel, M. Y. Hein, S. Stoynov, J. Mahamid, S. Saha, T. M. Franzmann, A. Pozniakovski, I. Poser, N. Maghelli, L. A. Royer, M. Weigert, E. W. Myers, S. Grill, D. Drechsel, A. A. Hyman and S. Alberti, *Cell*, 2015, **162**, 1066 – 1077.

- 102 C. Ishimoto, *Phys. Rev. Lett.*, 1977, **39**, 474–477.
- 103 J. A. Thomson, P. Schurtenberger, G. M. Thurston and G. B. Benedek, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 7079–7083.
- 104 F. G. Quiroz and A. Chilkoti, *Nat. Mater.*, 2015, **14**, 1164–1171.
- 105 S. F. Banani, A. M. Rice, W. B. Peeples, Y. Lin, S. Jain, R. Parker and M. K. Rosen, *Cell*, 2016, **166**, 651–663.
- 106 C. W. Pak, M. Kosno, A. S. Holehouse, S. B. Padrick, A. Mittal, R. Ali, A. A. Yunus, D. R. Liu, R. V. Pappu and M. K. Rosen, *Mol. Cell*, 2016, **63**, 72–85.
- 107 H. Zhang, S. Elbaum-Garfinkle, E. M. Langdon, N. Taylor, P. Occhipinti, A. A. Bridges, C. P. Brangwynne and A. S. Gelfander, *Mol. Cell*, 2015, **60**, 220–230.
- 108 W. M. Aumiller, F. Pir Cakmak, B. W. Davis and C. D. Keating, *Langmuir*, 2016, **32**, 10042–10053.
- 109 Y. Lin, S. L. Currie and M. K. Rosen, *J. Biol. Chem.*, 2017, **292**, 19110–19120.
- 110 H.-R. Li, W.-C. Chiang, P.-C. Chou, W.-J. Wang and J.-r. Huang, *J. Mol. Biol.*, 2018, **293**, 6090–6098.
- 111 T. Murakami, S. Qamar, J. Q. Lin, G. S. K. Schierle, E. Rees, A. Miyashita, A. R. Costa, R. B. Dodd, F. T. Chan, C. H. Michel, D. Kronenberg-Versteeg, Y. Li, S.-P. Yang, Y. Wakutani, W. Meadows, R. R. Ferry, L. Dong, G. G. Tartaglia, G. Favrin, W.-L. Lin, D. W. Dickson, M. Zhen, D. Ron, G. Schmitt-Ulms, P. E. Fraser, N. A. Shneider, C. Holt, M. Vendruscolo, C. F. Kaminski and P. S. George-Hyslop, *Neuron*, 2015, **88**, 678 – 690.
- 112 A. E. Conicella, G. H. Zerze, J. Mittal and N. L. Fawzi, *Structure*, 2016, **24**, 1537–1549.
- 113 A. Mollieux, J. Temirov, J. Lee, M. Coughlin, A. P. Kanagaraj, H. J. Kim, T. Mittag and J. P. Taylor, *Cell*, 2015, **163**, 123 – 133.
- 114 W. M. Aumiller Jr and C. D. Keating, *Nat. Chem.*, 2016, **8**, 129–137.
- 115 Z. Monahan, V. H. Ryan, A. M. Janke, K. A. Burke, S. N. Rhoads, G. H. Zerze, R. O’Meally, G. L. Dignon, A. E. Conicella, W. Zheng, R. B. Best, R. N. Cole, J. Mittal, F. Shewmaker and N. L. Fawzi, *EMBO J.*, 2017, **36**, 2951–2967.
- 116 T. P. Dao, R.-M. Kolaitis, H. J. Kim, K. O’Donovan, B. Martyniak, E. Colicino, H. Hehnl, J. P. Taylor and C. A. Castañeda, *Mol. Cell*, 2018, **69**, 965 – 978.e6.
- 117 D. Milovanovic, Y. Wu, X. Bian and P. De Camilli, *Science*, 2018, **361**, 604–607.
- 118 C. G. Toyota, M. D. Davis, A. M. Cosman and M. D. Hebert, *Chromosoma*, 2010, **119**, 205–215.
- 119 D. Hernandez-Verdun, *Nucleus*, 2011, **2**, 189–194.
- 120 J. T. Wang, J. Smith, B.-C. Chen, H. Schmidt, D. Rasoloson, A. Paix, B. G. Lambrus, D. Calidas, E. Betzig and G. Seydoux, *eLife*, 2014, **3**, e04591.
- 121 N. L. Kedersha, M. Gupta, W. Li, I. Miller and P. Anderson, *J. Cell Biol.*, 1999, **147**, 1431–1442.
- 122 X. Su, J. A. Ditlev, E. Hui, W. Xing, S. Banjade, J. Okrut, D. S. King, J. Taunton, M. K. Rosen and R. D. Vale, *Science*, 2016, **352**, 595–599.
- 123 A. Jain and R. D. Vale, *Nature*, 2017, **546**, 243–247.
- 124 P. Tompa, *Trends Biochem. Sci.*, 2012, **37**, 509 – 516.
- 125 M. Zeng, X. Chen, D. Guan, J. Xu, H. Wu, P. Tong and M. Zhang, *Cell*, 2018, **174**, 1172 – 1187.e16.
- 126 R. Hancock, *J. Struct. Biol.*, 2004, **146**, 281 – 290.
- 127 A. B. Singer and J. G. Gall, *Nucleus*, 2011, **2**, 403–409.
- 128 L.-P. Bergeron-Sandoval and S. W. Michnick, *J. Mol. Biol.*, 2018, **430**, 4754 – 4761.
- 129 D. M. Mitrea, B. Chandra, M. C. Ferrolino, E. B. Gibbs, M. Tolbert, M. R. White and R. W. Kriwacki, *J. Mol. Biol.*, 2018, **430**, 4773 – 4805.
- 130 C. P. Brangwynne, T. J. Mitchison and A. A. Hyman, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 4334–4339.
- 131 J. Kim, K. Y. Han, N. Khanna, T. Ha and A. S. Belmont, *bioRxiv*, 2018, 347955.
- 132 Z. F. Nizami and J. G. Gall, *Chromosome Res.*, 2012, **20**, 953–969.
- 133 T. M. Franzmann, M. Jahnel, A. Pozniakovskiy, J. Mahamid, A. S. Holehouse, E. Nüske, D. Richter, W. Baumeister, S. W. Grill, R. V. Pappu, A. A. Hyman and S. Alberti, *Science*, 2018, **359**, eaao5654.
- 134 P. P. Gopal, J. J. Nirschl, E. Klinman and E. L. F. Holzbaur, *Proc. Natl. Acad. Sci. USA*, 2017, **114**, E2466–E2475.
- 135 R. D. Phair and T. Misteli, *Nature*, 2000, **404**, 604–609.
- 136 A. R. Strom, A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq and G. H. Karpen, *Nature*, 2017, **547**, 241–245.
- 137 S. Jain, J. R. Wheeler, R. W. Walters, A. Agrawal, A. Barsic and R. Parker, *Cell*, 2016, **164**, 487–498.
- 138 A. Putnam, M. Cassani, J. Smith and G. Seydoux, *bioRxiv: 10.1101/245878*, 2018.
- 139 F. Wippich, B. Bodenmiller, M. G. Trajkovska, S. Wanka, R. Aebersold and L. Pelkmans, *Cell*, 2013, **152**, 791–805.
- 140 J. Berry, S. C. Weber, N. Vaidya, M. Haataja and C. P. Brangwynne, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, E5237–E5245.
- 141 Y. Shin, J. Berry, N. Pannucci, M. P. Haataja, J. E. Toettcher and C. P. Brangwynne, *Cell*, 2017, **168**, 159–171.
- 142 T. E. Kaiser, R. V. Intine and M. Dundr, *Science*, 2008, **322**, 1713–1717.
- 143 V. Badalassi, H. Ceniceros and S. Banerjee, *J. Comput. Phys.*, 2003, **190**, 371–397.
- 144 E. B. Nauman and D. Q. He, *Chem. Eng. Sci.*, 2001, **56**, 1999–2018.
- 145 S. Ren and I. Hamley, *Macromolecules*, 2001, **34**, 116–126.
- 146 G. H. Fredrickson, V. Ganesan and F. Drolet, *Macromolecules*, 2002, **35**, 16–39.
- 147 G. Fredrickson, *The equilibrium theory of inhomogeneous polymers*, Oxford University Press on Demand, 2006, vol. 134.
- 148 D. Zwicker, M. Decker, S. Jaensch, A. A. Hyman and F. Jülicher, *Proc. Natl. Acad. Sci. USA*, 2014, **111**, E2636–E2645.
- 149 S. Mao, D. Kuldinow, M. P. Haataja and A. Kosmrlj, 2018.
- 150 Y.-H. Lin, J. D. Forman-Kay and H. S. Chan, *Phys. Rev. Lett.*, 2016, **117**, 178101.

- 151 V. Y. Borue and I. Y. Erukhimovich, *Macromolecules*, 1988, **21**, 3240–3249.
- 152 P. González-Mozuelos and M. O. De La Cruz, *J. Chem. Phys.*, 1994, **100**, 507–517.
- 153 Y.-H. Lin and H. S. Chan, *Biophys. J.*, 2017, **112**, 2043–2046.
- 154 B. J. Alder and T. E. Wainwright, *J. Chem. Phys.*, 1959, **31**, 459–466.
- 155 B. J. Alder and T. E. Wainwright, *J. Chem. Phys.*, 1957, **27**, 1208–1209.
- 156 A. Panagiotopoulos, N. Quirke, M. Stapleton and D. Tildesley, *Mol. Phys.*, 1988, **63**, 527–545.
- 157 J. I. Siepmann and D. Frenkel, *Mol. Phys.*, 1992, **75**, 59–70.
- 158 W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657–1666.
- 159 A. D. MacKerell Jr, N. Banavali and N. Fopolle, *Biopolymers*, 2000, **56**, 257–265.
- 160 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
- 161 C. Vega and J. L. Abascal, *Phys. Chem. Chem. Phys.*, 2011, **13**, 19663–19688.
- 162 S. J. Plimpton, *J. Comp. Phys.*, 1995, **117**, 1–19.
- 163 J. A. Anderson, C. D. Lorenz and A. Travesset, *J. Comp. Phys.*, 2008, **227**, 5342–5359.
- 164 T. D. Nguyen, C. L. Phillips, J. A. Anderson and S. C. Glotzer, *Comput. Phys. Comm.*, 2011, **182**, 2307–2313.
- 165 S. Izvekov and G. A. Voth, *J. Phys. Chem. B*, 2005, **109**, 2469–2473.
- 166 M. S. Shell, *J. Chem. Phys.*, 2008, **129**, 144108.
- 167 A. Vitalis and R. V. Pappu, *J. Comput. Chem.*, 2009, **30**, 673–699.
- 168 G. Torrie and J. Valleau, *J. Comp. Phys.*, 1977, **23**, 187–199.
- 169 F. Wang and D. Landau, *Phys. Rev. Lett.*, 2001, **86**, 2050.
- 170 M. S. Shell, P. G. Debenedetti and A. Z. Panagiotopoulos, *Phys. Rev. E*, 2002, **66**, 056703.
- 171 A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 12562–12566.
- 172 A. Barducci, G. Bussi and M. Parrinello, *Phys. Rev. Lett.*, 2008, **100**, 020603.
- 173 P. G. Bolhuis, D. Chandler, C. Dellago and P. L. Geissler, *Ann. Rev. Phys. Chem.*, 2002, **53**, 291–318.
- 174 R. J. Allen, P. B. Warren and P. R. Ten Wolde, *Phys. Rev. Lett.*, 2005, **94**, 018104.
- 175 R. J. Allen, D. Frenkel and P. R. ten Wolde, *J. Chem. Phys.*, 2006, **124**, 024102.
- 176 A. Haji-Akbari, *J. Chem. Phys.*, 2018, **149**, 072303.
- 177 T. S. van Erp, D. Moroni and P. G. Bolhuis, *J. Chem. Phys.*, 2003, **118**, 7762–7774.
- 178 T. S. van Erp and P. G. Bolhuis, *J. Comp. Phys.*, 2005, **205**, 157–181.
- 179 R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.*, 1986, **57**, 2607.
- 180 D. J. Earl and M. W. Deem, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3910–3916.
- 181 N. Rathore, M. Chopra and J. J. de Pablo, *J. Chem. Phys.*, 2005, **122**, 024111.
- 182 W. Kauzmann, in *Adv. Protein Chem.*, Elsevier, 1959, vol. 14, pp. 1–63.
- 183 H. F. Fisher, *Proc. Natl. Acad. Sci. USA*, 1964, **51**, 1285–1291.
- 184 K. A. Dill, *Biochemistry*, 1985, **24**, 1501–1509.
- 185 V. N. Uversky, J. R. Gillespie and A. L. Fink, *Proteins: Struct. Funct. Genet.*, 2000, **41**, 415–427.
- 186 R. Van Der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones *et al.*, *Chem. Rev.*, 2014, **114**, 6589–6631.
- 187 P. Romero, Z. Obradovic and A. K. Dunker, *Appl. Bioinform.*, 2004, **3**, 105–113.
- 188 Z. Dosztányi, V. Csizmok, P. Tompa and I. Simon, *Bioinformatics*, 2005, **21**, 3433–3434.
- 189 Z. Dosztányi, B. Mészáros and I. Simon, *Brief. Bioinform.*, 2009, **11**, 225–243.
- 190 B. Monastyrskyy, K. Fidelis, J. Moulton, A. Tramontano and A. Kryshtafovych, *Proteins: Struct. Funct. Genet.*, 2011, **79**, 107–118.
- 191 B. Monastyrskyy, A. Kryshtafovych, J. Moulton, A. Tramontano and K. Fidelis, *Proteins: Struct. Funct. Genet.*, 2014, **82**, 127–137.
- 192 R. B. Best, *Curr. Opin. Struc. Biol.*, 2017, **42**, 147–154.
- 193 P. S. Nerenberg, B. Jo, C. So, A. Tripathy and T. Head-Gordon, *J. Phys. Chem. B*, 2012, **116**, 4524–4534.
- 194 R. B. Best, W. Zheng and J. Mittal, *J. Chem. Theory Comput.*, 2014, **10**, 5113–5124.
- 195 J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller and A. D. MacKerell Jr, *Nat. Methods*, 2016, **14**, 71.
- 196 S. Piana, A. G. Donchev, P. Robustelli and D. E. Shaw, *J. Phys. Chem. B*, 2015, **119**, 5113–5123.
- 197 D. Mercadante, S. Milles, G. Fuertes, D. I. Svergun, E. A. Lemke and F. Gräter, *J. Phys. Chem. B*, 2015, **119**, 7975–7984.
- 198 G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best and J. Mittal, *PLoS Comput. Biol.*, 2018, **14**, e1005941.
- 199 A. K. Dunker, C. J. Oldfield, J. Meng, P. Romero, J. Y. Yang, J. W. Chen, V. Vacic, Z. Obradovic and V. N. Uversky, *BMC Genomics*, 2008, **9**, S1.
- 200 H. S. Ashbaugh and H. W. Hatch, *J. Am. Chem. Soc.*, 2008, **130**, 9536–9542.
- 201 C. M. Miller, Y. C. Kim and J. Mittal, *Biophys. J.*, 2016, **111**, 28–37.
- 202 P. Das, S. Matysiak and J. Mittal, *ACS Cent. Sci.*, 2018, **4**, 534–542.
- 203 A. H. Mao, S. L. Crick, A. Vitalis, C. L. Chicoine and R. V. Pappu, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 8183–8188.
- 204 T. Firman and K. Ghosh, *J. Chem. Phys.*, 2018, **148**, 123305.
- 205 M. Sickmeier, J. A. Hamilton, T. LeGall, V. Vacic, M. S. Cortese, A. Tantos, B. Szabo, P. Tompa, J. Chen, V. N. Uversky and Z. O. A. K. Dunker, *Nucleic Acids Res.*, 2006, **35**, D786–D793.

- 206 S. Das, A. Eisen, Y.-H. Lin and H. S. Chan, *J. Phys. Chem. B*, 2018, **122**, 5418–5431.
- 207 S. Das, A. Amin, Y.-H. Lin and H. S. Chan, *Phys. Chem. Chem. Phys.*, 2018, **20**, 28558–28574.
- 208 M. Sadqi, L. J. Lapidus and V. Muñoz, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 12117–12122.
- 209 D. Nettels, S. Müller-Späth, F. Küster, H. Hofmann, D. Haenni, S. Rügger, L. Reymond, A. Hoffmann, J. Kubelka, B. Heinz *et al.*, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 20740–20745.
- 210 T. D. Langridge, M. J. Tarver and S. T. Whitten, *Proteins: Struct. Funct. Genet.*, 2014, **82**, 668–678.
- 211 I. C. Sanchez, *Macromolecules*, 1979, **12**, 980–988.
- 212 R. Wuttke, H. Hofmann, D. Nettels, M. B. Borgia, J. Mittal, R. B. Best and B. Schuler, *Proc. Natl. Acad. Sci. USA*, 2014, **111**, 5213–5218.
- 213 G. H. Zerze, R. B. Best and J. Mittal, *J. Phys. Chem. B*, 2015, **119**, 14622–14630.
- 214 A. Hicks and H.-X. Zhou, *J. Chem. Phys.*, 2018, **149**, 072313.
- 215 G. H. Zerze, J. Mittal and R. B. Best, *Phys. Rev. Lett.*, 2016, **116**, 068102.
- 216 Z. A. Levine and J.-E. Shea, *Curr. Opin. Struc. Biol.*, 2017, **43**, 95–103.
- 217 J. Mittal, T. H. Yoo, G. Georgiou and T. M. Truskett, *J. Phys. Chem. B*, 2012, **117**, 118–124.
- 218 G. H. Zerze, C. M. Miller, D. Granata and J. Mittal, *J. Chem. Theory Comput.*, 2015, **11**, 2776–2782.
- 219 G. H. Zerze and J. Mittal, *J. Phys. Chem. B*, 2015, **119**, 15583–15592.
- 220 T. Förster, *Ann. Phys.*, 1948, **437**, 55–75.
- 221 G. H. Zerze, R. B. Best and J. Mittal, *Biophys. J.*, 2014, **107**, 1654–1660.
- 222 W. W. Smith, P.-Y. Ho and C. S. O'Hern, *Phys. Rev. E*, 2014, **90**, 042709.
- 223 W. Zheng, A. Borgia, K. Buholzer, A. Grishaev, B. Schuler and R. B. Best, *J. Am. Chem. Soc.*, 2016, **138**, 11702–11713.
- 224 T. Y. Yoo, J.-M. Choi, W. Conway, C.-H. Yu, R. V. Pappu and D. J. Needleman, *eLife*, 2018, **7**, e36392.
- 225 M. Li, T. Sun, F. Jin, D. Yu and Z. Liu, *Mol. Biosyst.*, 2016, **12**, 2932–2940.
- 226 W. Zheng, G. H. Zerze, A. Borgia, J. Mittal, B. Schuler and R. B. Best, *J. Chem. Phys.*, 2018, **148**, 123329.
- 227 Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.*, 1999, **314**, 141–151.
- 228 P. R. ten Wolde and D. Frenkel, *Science*, 1997, **277**, 1975–1978.
- 229 J. P. Doye, A. A. Louis, I.-C. Lin, L. R. Allen, E. G. Noya, A. W. Wilber, H. C. Kok and R. Lyus, *Phys. Chem. Chem. Phys.*, 2007, **9**, 2197–2205.
- 230 C. Gögelein, G. Nägele, R. Tuinier, T. Gibaud, A. Stradner and P. Schurtenberger, *J. Chem. Phys.*, 2008, **129**, 08B615.
- 231 Z. Zhang and S. C. Glotzer, *Nano Lett.*, 2004, **4**, 1407–1413.
- 232 E. Bianchi, J. Largo, P. Tartaglia, E. Zaccarelli and F. Sciortino, *Phys. Rev. Lett.*, 2006, **97**, 168301.
- 233 F. Romano and F. Sciortino, *Nat. Mater.*, 2011, **10**, 171.
- 234 F. Smallenburg and F. Sciortino, *Nat. Phys.*, 2013, **9**, 554.
- 235 W. M. Jacobs, D. W. Oxtoby and D. Frenkel, *J. Chem. Phys.*, 2014, **140**, 05B626.
- 236 A. J. Simon, V. Ramasubramani, J. Glaser, A. Pothukuchy, J. Gerberich, J. Leggere, B. R. Morrow, J. Golihar, C. Jung, S. C. Glotzer, D. W. Taylor and A. D. Ellington, *bioRxiv*, 2018, 323261.
- 237 W. M. Jacobs and D. Frenkel, *Biophys. J.*, 2017, **112**, 683–691.
- 238 D. Adams, *Mol. Phys.*, 1974, **28**, 1241–1252.
- 239 A. Ghavami, E. van der Giessen and P. R. Onck, *J. Chem. Theory Comput.*, 2012, **9**, 432–440.
- 240 E. Rudisill and P. Cummings, *Mol. Phys.*, 1989, **68**, 629–635.
- 241 B. Smit and D. Frenkel, *Mol. Phys.*, 1991, **74**, 35–39.
- 242 P. Debye and E. Hückel, *Phys. Z.*, 1923, **24**, 185–206.
- 243 S. De Leeuw, *Proc. R. Soc. London, Ser. A*, 1980, **373**, 27.
- 244 F. J. Blas, L. G. MacDowell, E. de Miguel and G. Jackson, *J. Chem. Phys.*, 2008, **129**, 144703.
- 245 K. S. Sillmore, M. P. Howard and A. Z. Panagiotopoulos, *Mol. Phys.*, 2017, **115**, 320–327.
- 246 G. L. Dignon, W. Zheng, R. B. Best, Y. C. Kim and J. Mittal, *Proc. Natl. Acad. Sci. USA*, 2018, **115**, 9929–9934.
- 247 J. R. Buchan, R.-M. Kolaitis, J. P. Taylor and R. Parker, *Cell*, 2013, **153**, 1461 – 1474.
- 248 M. Loschi, C. C. Leishman, N. Berardone and G. L. Boccaccio, *J. Cell Sci.*, 2009, **122**, 3973–3982.
- 249 M. G. Thomas, M. Loschi, M. A. Desbats and G. L. Boccaccio, *Cell. Signal.*, 2011, **23**, 324 – 334.
- 250 M. Perez-Pepe, A. J. Fernández-Alvarez and G. L. Boccaccio, *Biochemistry*, 2018, **57**, 2488–2498.
- 251 M. Dunder and T. Misteli, *CHS Perspect. Biol.*, 2010, **2**, a000711.
- 252 S. P. Shevtsov and M. Dunder, *Nat. Cell. Biol.*, 2011, **13**, 167–173.
- 253 G. H. Karpen, J. E. Schaefer and C. D. Laird, *Genes & Dev.*, 1988, **2**, 1745–1763.
- 254 A. Grob, C. Collieran and B. McStay, *Genes & Dev.*, 2014, **28**, 220–230.
- 255 H. R. Salzler, D. C. Tatomer, P. Y. Malek, S. L. McDaniel, A. N. Orlando, W. F. Marzluff and R. J. Duronio, *Dev. Cell*, 2014, **24**, 623–634.
- 256 S. C. Takatori and J. F. Brady, *Curr. Opin. Colloid. In.*, 2016, **21**, 24 – 33.
- 257 C. F. Lee, C. P. Brangwynne, J. Gharakhani, A. A. Hyman and F. Jülicher, *Phys. Rev. Lett.*, 2013, **111**, 088101.
- 258 D. Zwicker, R. Seyboldt, C. A. Weber, A. A. Hyman and F. Jülicher, *Nat. Phys.*, 2017, **13**, 408.
- 259 C. A. Weber, C. F. Lee and F. Jülicher, *New J. Phys.*, 2017, **19**, 053021.
- 260 J. D. Wurtz and C. F. Lee, *New J. Phys.*, 2018, **20**, 045008.
- 261 D. Zwicker, J. Baumgart, S. Redemann, T. Müller-Reichert, A. A. Hyman and F. Jülicher, *Phys. Rev. Lett.*, 2018, **121**,

- 158102.
- 262 S. Elbaum-Garfinkle, Y. Kim, K. Szczepaniak, C. C.-H. Chen, C. R. Eckmann, S. Myong and C. P. Brangwynne, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, 7189–7194.
- 263 T. J. Nott, T. D. Craggs and A. J. Baldwin, *Nat. Chem.*, 2016, **8**, 569.
- 264 K. Gaffney and H. Chapman, *Science*, 2007, **316**, 1444–1448.
- 265 H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall *et al.*, *Nature*, 2011, **470**, 73.
- 266 A. H. Zewail, *Science*, 2010, **328**, 187–193.
- 267 K. A. Dill, *Biochemistry*, 1990, **29**, 7133–7155.
- 268 D. Chandler, *Nature*, 2005, **437**, 640.
- 269 A. J. Patel, P. Varilly, S. N. Jamadagni, H. Acharya, S. Garde and D. Chandler, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 17678–17683.
- 270 S. Sharma and P. G. Debenedetti, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 4365–4370.
- 271 Y. E. Altabet and P. G. Debenedetti, *J. Chem. Phys.*, 2014, **141**, 18C531.
- 272 R. C. Remsing, E. Xi, S. Vembanur, S. Sharma, P. G. Debenedetti, S. Garde and A. J. Patel, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, 8181–8186.
- 273 Y. E. Altabet, A. Haji-Akbari and P. G. Debenedetti, *Proc. Natl. Acad. Sci. USA*, 2017, **114**, E2548–E2555.
- 274 T. Li, D. Donadio, L. M. Ghiringhelli and G. Galli, *Nat. Mater.*, 2009, **8**, 726–730.
- 275 T. Li, D. Donadio and G. Galli, *Nat. Comm.*, 2013, **4**, 1887.
- 276 A. Haji-Akbari, R. S. DeFever, S. Sarupria and P. G. Debenedetti, *Phys. Chem. Chem. Phys.*, 2014, **16**, 25916–25927.
- 277 A. Haji-Akbari and P. G. Debenedetti, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, 10582.
- 278 R. Cabriolu and T. Li, *Phys. Rev. E*, 2015, **91**, 052402.
- 279 M. M. Gianetti, A. Haji-Akbari, M. P. Longinotti and P. G. Debenedetti, *Phys. Chem. Chem. Phys.*, 2016, **18**, 4102–4111.
- 280 A. Haji-Akbari and P. G. Debenedetti, *Proc. Natl. Acad. Sci. USA*, 2017, **114**, 3316–3321.
- 281 D. Bracha, M. T. Walls, M.-T. Wei, L. Zhu, M. Kurian, J. L. Avalos, J. E. Toettcher and C. P. Brangwynne, *Cell*, 2018, **175**, 1467–1480.
- 282 D. L. Chen, C. J. Gerdtts and R. F. Ismagilov, *J. Am. Chem. Soc.*, 2005, **127**, 9672–9673.
- 283 C. J. Gerdtts, V. Tereshko, M. K. Yadav, I. Dementieva, F. Collart, A. Joachimiak, R. C. Stevens, P. Kuhn, A. Kossiakoff and R. F. Ismagilov, *Angew. Chem. Int. Ed*, 2006, **118**, 8336–8340.
- 284 J.-u. Shim, G. Cristobal, D. R. Link, T. Thorsen and S. Fraden, *Cryst. Growth Des.*, 2007, **7**, 2192–2194.
- 285 G. M. Whitesides, *Nature*, 2006, **442**, 368.