Formation of biological condensates via phase separation: Characteristics, analytical methods, and physiological implications

Zhe Feng¹, Xudong Chen¹, Xiandeng Wu¹, and Mingjie Zhang¹,²,*

¹Division of Life Science, State Key Laboratory of Molecular Neuroscience, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

²Center of Systems Biology and Human Health, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

Running title: Biological condensates formation via phase separation

* To whom correspondence should be addressed: Mingjie Zhang; Division of Life Science, State Key Laboratory of Molecular Neuroscience, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China; mzhang@ust.hk; Tel. (852) 2358 8709.

Keywords: phase separation, biological condensates, protein-protein interaction, cell signaling, cell biology, cellular regulation, membraneless organelle, intrinsically disordered protein, scaffold proteins, multivalent interactions

ABSTRACT

Liquid–liquid phase separation (LLPS) facilitates the formation of condensed biological assemblies with well-delineated physical boundaries, but without lipid membrane barriers. LLPS is increasingly recognized as a common mechanism for cells to organize and maintain different cellular compartments in addition to classical membrane-delimited organelles. Membraneless condensates have many distinct features that are not present in membrane-delimited organelles and that are likely indispensable for the viability and function of living cells. Malformation of membraneless condensates is increasingly linked to human diseases. In this review, we summarize commonly used methods to investigate various forms of LLPS occurring both in 3D aqueous solution and on 2D membrane bilayers, such as LLPS condensates arising from intrinsically disordered proteins or structured modular protein domains. We then discuss, in the context of comparisons with membrane-delimited organelles, the potential functional implications of membraneless condensate formation in cells. We close by highlighting some challenges in the field devoted to studying LLPS-mediated membraneless condensate formation.

INTRODUCTION

In eukaryotic cells, reaction components are spatiotemporally compartmentalized such that materials are concentrated,
activities are localized and protected from damaging activities such as proteolysis, changes in pH and undesired covalent modifications. Classical organelles are membrane-enclosed where the lipid bilayer provides a physical barrier to separate their interior contents from the exterior environment. Examples include Golgi apparatus, mitochondria and endoplasmic reticulum (ER). However, many organelles are not membrane-enclosed (often referred to as membraneless compartments in the literature), and such organelles include but not limited to germ granules, stress granules, nucleoli, centrosomes and synapses in neurons. In these membraneless compartments, due to the lack of physical separation, molecules can freely exchange with their counterparts in the surrounding bulk solution. Sharp concentration gradients are maintained between the proteinaceous (and sometimes protein and nucleic acid mixtures) interior and the much more diluted exterior. Reaction machineries can reversibly assemble and disassemble within a short time window, as fast as a few seconds. Reaction constituents can be integrated or removed to control specific activities. Although recognized for many years, the mechanisms governing the formation of membraneless organelles have remained unclear until about 10 years ago. First direct experimental evidence came from the study of P granules in germ cells of Caenorhabditis elegans (1). P granule is a collection of RNA and RNA binding proteins (RBPs) localized at the posterior cortex of a dividing embryo. P granules appear as spherical droplets with liquid-like properties—fuses with one another, deform under shear stress and flow off the surface of nucleus. Fluorescence recovery after photobleaching (FRAP) analysis demonstrated rapid turnover rates of constituent proteins, which is indicative of fast molecular rearrangements. These observations together suggested that P granules form through liquid-liquid phase separation (LLPS), distinct from canonical macromolecular assemblies. Since then the list of membraneless organelles that are organized by LLPS has ever been growing. Nevertheless, early concerns had been raised over the specificity of phase-separated condensates observed in vitro and their biological significance in vivo. Comprehensive studies were followed to show that the concept of phase separation can help to explain the formation and organization of non-membrane bound biomolecular compartments as well as their physical and material properties that cannot be understood with the classical physical chemistry theories for dilute solutions. It now comes to the realization that LLPS might be a general mechanism to drive compartmentalization in the absence of lipid bilayers (2-7), and this has greatly motivated the field to re-investigate mechanisms underlying formation and functional implications of membraneless organelles from new perspectives. However, cautions should be exercised that not all condensed phase properties observed in vitro can be extrapolated to living cells. Rigorous characterizations, both in vitro and in vivo, are required to demonstrate the existence of LLPS of a particular biological system under physiological conditions. Here we first review our understanding about the molecular codes that contribute to LLPS formation. We discuss some of the common techniques for characterization of LLPS. We then discuss the functional implications of LLPS-driven organelle assemblies. Finally, we propose a few new potential research directions inspired by current works on LLPS.

**What is phase separation?**
By definition, phase separation refers to the immiscibility of two solutions whereby they separate into two states. In biological systems, this often leads to a large volume, dilute liquid phase and a small volume, concentrated liquid phase. In biology, phase separation is not unheard of. In protein crystallization, when a crystallization reagent is added to the protein solution oil-like droplets can be observed to separate from the bulk solution. LLPS in this case is indicative of a metastable transition state from which crystals may grow by changing temperature, precipitant concentration, protein concentration etc. Molecules are miscible in solution until reaching their solubility limit. Phase separation happens when the macromolecule/macromolecule or solute/solute interactions are energetically favored over the macromolecule/solute interactions and the gain in free energies is favored over its loss in entropic tendency towards homogenous solution state (6,8,9). A free energy minimum is then reached, but the two phases with different solute concentrations are at the same Gibbs free energy (4). For each molecular system, a phase diagram can be constructed by systematically screening through conditions such as temperature, salt concentration, pH or macromolecular concentration. Phase diagram helps one to identify conditions that promote phase separation and to determine the likelihood of phase separation happening under physiological conditions (Fig. 1A). Phase boundary, which is defined by the binodal line, indicates the boundary that two distinct phases stably co-exist in solution. Outside the binodal curve, molecules stay as homogenous mixtures. Between the binodal and spinodal curves lies a metastable region where liquid demixes via a nucleation process. Within the spinodal zone, spinodal decomposition occurs. In the other words, spontaneous phase separation takes place in the spinodal zone where molecules rapidly transit from a less stable region to a more stable phase separated region bypassing the metastable nucleation zone.

Multivalency is a key determining factor underlying LLPS in biological systems (10). Molecules can undergo inter- or intra-molecular interactions to assemble into oligomers or polymers which tend to have lowered solubility limit and thus more likely to demix with the surrounding solution. In a folded domain protein, multiple binding sites, either for itself or for its binding partners, promote phase separation (Fig. 2). In proteins with higher content of intrinsic disorder, multivalent weakly self-attracting interactions collectively drive phase separation (Fig. 1B-D). Aggregations of misfolded cytosolic or nuclear proteins have been associated with a broad range of neurodegenerative diseases such as Alzheimer’s disease (AD), Parkinson’s disease (PD) and amyotrophic lateral sclerosis (ALS) (11-13). Solid fibrils formed by disordered proteins represent another form of phase separation− sol to solid transition. Pathological aggregation and its close link to brain diseases are discussed in several recent reviews (14-19). For the scope of this review, we focus on recent works on LLPS. Below we discuss the sequence properties coding for LLPS, the biological functions of condensate formation and the technical developments for in vitro characterization of phase separation systems.

Codes for LLPS

Tremendous progresses have been made over the past decade trying to understand the molecular features in common that drive phase separation. In this section, we discuss examples of phase separation promoted by intrinsically disordered sequences or more
specific folded domain/target interactions.

**Multivalency driven by intrinsic disordered sequences**

Intrinsically disordered proteins represent an abundant class of proteins involved in phase separation. Low complexity regions (LCRs) show biased amino acid preferences including Gly, Ser, Asn, Glu, Phe and Tyr. These amino acids often appear in repeats such as RG, FG and YG repeats that are important for forming ribonucleoprotein (RNP) granules in P bodies, P granules, stress granules in cytoplasm or nucleoli and paraspeckles in the nucleus (20-25). The lack of a defined three-dimensional (3D) structure in LCRs favors weakly adhesive interactions that drive phase separation. A good example studied in detail is Fused in Sarcoma (FUS) proteins. Full length FUS proteins have been shown to undergo LLPS at close to physiological concentrations in the presence of crowding reagent or when cooled to 4 °C (26,27). Gel-like state is observed when FUS protein concentration reaches hundreds of μM, well above its physiological concentration. FUS LCR alone can assemble into hydrogels at sub-mM concentration (28), and the hydrogels trap and retain the LC domains of other RBPs such as hnRNPA1, Sup35, TIA1 and TDP43, although to different levels. The liquid droplets of FUS protein can further mature into fibrous aggregates resembling the pathological protein fibers found in ALS patients. Mutations in the prion-like domain, which induce the early onset of ALS, further promote the sol-solid transition. Tremendous progresses have been made in recent years in revealing the emergent sequence determinants in LCR that promote phase separation. Noticeably, the types of interactions critical for phase separation are commonly known to drive protein folding or interactions. We discuss below how these sequence features can drive molecular interactions in new ways.

Intrinsically disordered proteins rich in aromatic residues are favored to form pi-pi stacking interactions that can drive phase separation (25,28-30) (Fig. 1B). Mutation of these aromatic residues to serine can strongly decrease the amount of protein enrichment into condensed phase. Apart from side chain pi-pi interactions, small residues with relatively exposed backbone peptide bonds can also form the so named planar pi interactions (31). Gly, Ser, Thr and Pro residues are indeed frequently found in LCRs of RBPs. RG/RGG repeats are also found in multiple LCR containing proteins such as in the nuage protein DEAD-box helicase 4 (Ddx4) (25), the P granule protein LAF-1 (22) and the neuronal granule protein Fragile X Mental Retardation Protein (FMRP) (32,33). Arginine can form cation-pi interactions with aromatic residues, either intramolecularly or intermolecularly (Fig. 1B). Increase in the number of cation-pi interactions by arginine substitutions in FUS protein, for example, can significantly promote its ability to phase separate and lower the threshold concentration of the sol-gel transition (30,34). Conversely, substitution of arginine and tyrosine or phenylalanine with alanine disrupts cation-pi interactions and consequently the ability to phase separate. Similarly, arginine methylation in Ddx4 (25), hnRNPA2 (35) and FMRP (33) reduces or abolishes their phase separation likely because of the weakened intermolecular cation-pi interactions. Charged residues also contribute to droplet formation both in vitro and in vivo (Fig. 1C). Ddx4, for example, contains blocks of net negative or positive charges, typically 8-10 residues in length with 3-8 charged residues (25). Interestingly, these charge blocks appear in clusters with
alternating positive and negative charge distributions. Removal of such opposite charge patterning inhibited Ddx4 phase separation. Apart from amino acid side chain interactions, phase separation can be driven by secondary structural elements. Examples include a short, evolutionarily conserved helical segment in TDP-43 C-terminal domain involved in intermolecular helical interactions (36) and a 57-residue segment in FUS LCR involved in the formation of cross β-sheets stabilized by hydrogen bonding and pi-pi stacking interactions (37). Recent crystallographic studies of LCRs in FUS, hnRNPA1 and nup98 revealed another type of interactions between secondary structural elements that drive phase separation. These regions are highly abundant in aromatic residues, which are involved in inter- and intra-sheet stabilizations. In addition, the LCRs form kinked β sheets to allow close encountering of the backbones for hydrogen bonding or Van der Waals interactions and subsequently to stabilize the packing of neighboring β sheets. Such regions are therefore referred to as low complexity aromatic-rich kinked segments (LARKS) (38) (Fig. 1D).

Based on current knowledge of relationship between emerging sequence features and phase separation, a spectrum of predictive tools has been developed to enable researchers to identify regions in intrinsically disordered proteins that might be involved in LLPS and to understand the molecular mechanisms behind sol-sol/gel transitions. This has been extensively discussed in a review written by Alberti and colleagues (3).

**Multivalent interactions driven by defined modular protein domains**

Experiments to manipulate the valency of a folded protein have proven an inverse correlation between the number of binding domains or motifs and the saturation concentration above which the system undergoes phase separation. For instance, repeats of SH3 domain bind Pro-rich motifs (PRMs) and phase separate into condensed droplets upon mixture at high concentration (10). The phase boundary (i.e. the threshold concentration for LLPS) is lowered when the number of binding modules increases suggesting that LLPS is strongly dependent on the valency of interactions. Similarly, multivalent nucleic acid/protein interaction systems are known to undergo LLPS both *in vitro* and in cells when certain critical numbers of valency are reached (39,40). There are now many examples of phase separation systems driven by modular domain interactions. One example is the multivalent protein network involving the transmembrane protein nephrin, the adaptor protein NCK and its ligand N-WASP that regulate actin assembly in podocytes of kidney (10,41) (Fig. 2A). NCK contains three SH3 domains, each of which can bind to the six PRMs in N-WASP; two proteins assemble into higher order oligomers that phase separate. This process is accelerated by nephrin addition where phospho-tyrosine (pTyr) residues in nephrin bind to SH2 domains in NCK. The assembled droplets can further recruit Arp2/3 complexes for actin polymerization. An analogous system is observed in T cell receptor signaling (42). LAT, a transmembrane protein for T cell activation, is phosphorylated at multiple Tyr sites that are required for T cell signaling. Addition of Grb2, an adaptor protein, and Sos1, a guanine nucleotide exchange factor for Ras GTPase, caused phase separation of pLAT through interaction between pTyr in LAT and SH2 domain in Grb2 and between SH3 domains in Grb2 and PRMs in Sos1. The number of pTyr residues affects binding valency in the system and consequently the
efficiency of receptor clustering on supported lipid bilayer. In both systems, properties of reconstituted condensates in vitro strongly correlate with the observations in living cells.

In neurons, synapses assemble between axons and dendrites. Beneath the postsynaptic membranes there lies an electron dense layer of material known as the postsynaptic density (PSD). Addition of SynGAP, a negative PSD activity regulator, to the major PSD scaffold protein PSD-95 caused droplet formation (43). Phase separation was completely abrogated if the interaction interface or SynGAP trimer interface was impaired. PSD-95 also clustered with SAPAP, Shank and Homer, which are major PSD scaffold proteins, but with much higher efficiency compared to SynGAP (44) (Fig. 2B). This increased propensity to form droplets is likely because of increasing valency provided by multivalent interaction interfaces among PSD constituents. Importantly, all of the interactions involved in forming the PSD protein network are highly specific and with strong affinities, and these interactions involve well folded protein binding domains (Fig. 2B). Strikingly, the assembly of PSD droplets was dispersed when Homer1a, a monomeric splice variant of Homer1, was added to the pre-assembled condensates (44). This suggests Homer1 oligomerization plays a crucial role in promoting LLPS. Reconstituted PSD condensates can further cluster the cytoplasmic tail of NMDA receptor subunit and nucleate actin polymerization both in solution and on supported lipid bilayer.

The presynaptic active zone is also organized by phase separation (45). As viewed under the electron microscope, active zone comprises densely packed proteins, which organize into a layer of electron dense projection beneath the presynaptic membranes. RIM and RIM-binding protein, two major active zone scaffold proteins, formed liquid droplets upon mixing. RIM-binding protein contains three SH3 domains, each of which binds to Pro-rich motifs in RIM (Fig. 2C). The cytoplasmic tail of voltage-gated Ca\(^{2+}\) channel (NCav) is also enriched into RIM/RIM-binding protein condensates, and this co-clustering significantly lowers the threshold concentration to undergo LLPS. When NCav C-terminal tail is attached to membrane, RIM, RIM-binding protein and NCav co-cluster on supported lipid bilayer, providing a mechanistic explanation to the tight coupling of Ca\(^{2+}\) influx and neurotransmitter release in presynaptic termini.

In addition to these systems, aggregation of Rubisco by the protein CcmM (46) and interaction between the tetravalent RNA binding protein PTB and an RNA oligonucleotide (10,47) are also shown to phase separate through multivalent folded domain interactions. In RBPs, phase separation can be driven by modular domain interactions apart from those weak, self-adhesive interactions. In particular, RGG repeats and RNA recognition motifs (RRMs) are involved in RNA binding. Although binding between individual repeats and RNA is relatively weak, multiple RGG repeats together generate a high affinity interaction and crosslink proteins and RNAs into higher order oligomers (33,36,48-50). Repeats of RRM on RBPs and multiple RRM binding sites on RNA provide further degree of multivalency to promote phase separation. Addition of RNA to hnRNPA1 caused formation of liquid droplets (24). In other cases, RNA is readily recruited to the pre-assembled liquid droplets or hydrogels via charge-charge interactions (33,51-54). It appears that intrinsically disordered regions and high affinity domain interactions can
both contribute to phase separation. In reality, it is likely that both types of multivalency are coupled to promote droplet formation in many systems.

Methods for observation of phase separation

When it comes to work with phase separation in vitro, multiple approaches are generally combined to describe the phenomenon of phase separation, to distinguish it from non-physiologically relevant aggregations or simple binding-induced molecular assemblies, and probably most importantly, to link to its biological functions in cells.

3D solution system

The most intuitive observation of phase separation in a test tube may be the turbidity and opalescence of a solution when components are mixed under certain conditions (protein concentration, stoichiometry, salt concentration, pH, temperature etc.). Such simple sedimentation assay can be used to quantitatively evaluate the fractional distribution of proteins in each phase (see (43) for an example). A transparent “pellet” observed after centrifugation implies a liquid phase rather than aggregates or precipitates. Alternatively, direct measurement of the turbidity (46,55,56) is also helpful for estimating the extent of phase separation. But practically, due to the non-negligible gravity of the condensed phase droplets, researchers should be cautious about the heterogeneity within final solution, and two experimental setups must be considered: 1, a vertical, instead of horizontal, light beam is recommended for absorbance measurement to gain a more reliable result. 2, sample solution needs to be vortexed immediately before the measurement to ensure a homogenous suspension.

Imaging assays are necessary to confirm the liquid-like properties of condensed phase. Differential interference contrast (DIC) imaging is the most straightforward method to depict the coexistence of two (or more) distinctive phases (Fig. 3A). The spherical morphology, fusion upon contact, and droplet fission, as well as the deformation of droplets under shear forces, together demonstrate the liquid-like properties of condensed phase (1,4,6) (Fig. 3B-D). Combining with fluorophore labeling, either colocalization or coexistence of sub-compartments can be visualized (57) (Fig. 3A). But we should always be alerted of potential imaging artifacts and should not merely rely on imaging assays for the following reasons: 1, cross-talk between multiple channels is not easy to be completely blocked, and it will always give an image when high laser power and long exposure time are applied. Thus, we recommend using single fluorophore labeling, if possible, in a given system except for co-localization or other necessary conditions. 2, the conjugation of a fluorophore (or genetically encoded fluorescent tag) may affect the properties of labeled protein, and relatively subtle impact may be augmented in a much more concentrated phase. In addition, overexposure of high percentage labeled protein may cause photobleaching of the fluorophore under continuous laser power, therefore we recommend researchers to dilute the labeled protein with unlabeled one to achieve sparse labeling (usually, ~1% is sufficient). It is helpful to use sedimentation assay with labeled protein to check the potential effect of labeling on its ability to phase separate. 3, it’s hard to accurately judge whether a protein of interest is specifically retained in the condensed phase, especially when the fluorescence contrast to
the surrounding environment is not high enough. Since the matrix pore size (i.e., the void space between the protein network mesh) of condensed phase is deemed to be large enough to accommodate normal protein, as indicated by our recent observation that the dodecameric CaMKII (~550 kDa) can penetrate the reconstituted condensed PSD phase in vitro (44), it is not surprising that even some irrelevant proteins can go through but do not prefer the condensed phase. Nonetheless, fluorescent imaging can provide valuable information with rigorous controls. Recently, we developed an absolute concentration estimation method based on measured fluorescence intensity (44) (Fig. 3E). A standard curve is first constructed by plotting the fluorescence of protein measured at known concentrations. Based on this curve we can then back calculate the exact concentration of components in a condensed phase, even though the concentration in the surrounding dilute phase cannot be precisely estimated due to the very low signal to noise ratio. Protein concentration in the dilute phase may be determined taking into account of its fractional distribution observed in sedimentation assay. Concentration ratio and therefore, volume ratio between condensed and dilute phases can ultimately be estimated (45).

FRAP analysis is increasingly adopted to demonstrate the mobility and dynamics of molecules within liquid droplets (Fig. 3C). Molecules exchange within condensed phase (half-bleach) or exchange between condensed and dilute phases (half-/whole-bleach) can be faithfully captured by FRAP experiments. However, we should be cautious in assessing the fitting of fluorescence recovery curves because it can always give us a result regardless of whether the model is appropriate or not. Fluorescence recovery is dependent on the movement rates of molecules, but this “movement” consists of diffusion (in the dilute phase, condensed phase, and the interface) and interaction (dissociation $k_{off}$ and association $k_{on}$). It is therefore difficult to derive an exact value of either characteristic diffusion rate constant ($\tau$) or diffusion coefficient (D) before figuring out a theoretical model, even though the apparent $\tau$ and D provide certain referential meanings. Besides, a less mobile or immobile fraction may occur, and a second bleach immediately after the plateau of the first bleach is suggested to confirm this. The immobile fraction can be generated from systematic background which is an intrinsic property or produced during imaging time by rapid hardening/ageing. In some systems, the phase droplets are initially fluid, but their viscoelasticity increases over the time and molecules eventually cannot exchange with their counterparts in the surrounding solution. This process is known as hardening/ageing, although the mechanism behind it is currently unknown.

Apart from direct visualization, several techniques have recently been brought into phase separation field to describe the material properties of biological condensates. The isolated droplets make it possible to monitor the material properties of individual phase via atomic force microscopy (AFM), from which the stiffness, viscosity, elasticity, pore size, and other soft material parameters can be quantitatively extracted. This information will provide useful insights into the behavior of biological condensates in cells. For example, AFM measurements of the mechanical properties of PSD droplets indicate that the 6-component PSD condensates is more gel-like comparing to the 2-component PSD condensates reconstituted in vitro (44). It is thus reasonable to speculate that under physiological conditions, PSD may be a more
gel-like structure due to the more complex valences and more crowded environment, which fits well to the previous electron microscopy (EM) observations (58). Measured material properties also provide explanations towards the observed protein dynamics in vitro. The time-dependent hardening indicated by elastic modulus values suggested an aging process of reconstituted PSD condensates, consistent with the observation that PSD constituents demonstrated a time-dependent decreasing of the signal recovery in FRAP analysis (44).

How are molecules organized within the condensed phase? Researchers show great interests towards understanding the atomic details of condensed phase. The intrinsic heterogeneity, highly dynamic properties, and the numerous transient interactions existing in liquid-like phase make it extremely difficult to obtain structural information (59). Nevertheless, protein concentrations within condensed phase remain the same in a given condition (pH, T, salt etc.) (3,44), bringing hopes to acquire some configuration rules from structural studies. A recent study combing cryo-EM and cryo-electron tomography (cryo-ET) to solve the structure of Rubisco-CcmM complex under LLPS condition may give us some inspiration (46). Cryo-ET analysis of clusters of Rubisco complexes revealed that the median nearest-neighbor distance is around 150Å, and the linker region sequesters two complex modules, which makes it possible to solve the complex structure within condensed phase by single particle cryo-EM.

2D membrane system

Signal transduction between cells cannot skip over membranes. Supported lipid bilayer has been a popular working model to mimic cell membrane in vitro for years (60). People have noticed large number of membrane proteins such as adhesion molecules, receptors and channels that are required to be assembled/enriched/clustered together to transmit signals, but conventional protein-protein interactions can hardly elucidate the coupling principle until phase separation came into sights. Reconstitution of transmembrane protein clustering on supported lipid bilayer is important for studying the mechanism of formation and the functional consequences of these microclusters (Fig. 4A). A recent review by Lindsay et al has summarized the significance of LLPS in transmembrane signaling (61). Here we discuss about some applications and their caveats when dealing with supported lipid bilayer in phase separation systems.

To visualize the clustering of membrane proteins on supported lipid bilayer, either Total Internal Reflection Fluorescence (TIRF) or confocal microscopy can be performed considering the membrane thickness is much below the optical microscopy resolution. Traditional imaging methods like colocalization, fusion, dispersion, and FRAP can also be conducted to describe the ensemble behaviors of proteins on supported lipid bilayer. In addition, all the materials are restricted to a single membrane sheet which theoretically accounts for all the signal sources, making it more convenient for quantification. Firstly, fluorescence intensity-based quantification methods as described for analysis of droplets in 3D solution is applicable. Same fluorophore can be conjugated to either lipid or protein, the absolute number of lipids can be calculated by the known coating surface area and lipid head group size. A standard curve of fluorescence intensity with respect to the number of fluorophores can then be plotted. Assuming the illumination property of a
fluorophore remains the same no matter being conjugated to lipid or to protein, protein density can thus be converted by referring to the standard curve generated from the labeled lipid (41,45). An alternative approach is to take advantage of the bilayer membrane to perform single molecule counting. The protein of interest is labeled with two different fluorophores, where one is sparse enough for single molecule counting and the other for experiments. One can then calculate the molecular number from the known concentration ratio of sparse labeled fraction before coating (42,62).

Furthermore, the confinement of membrane protein on supported lipid bilayer allows one to trace proteins at single molecular level (Fig. 4B). STORM provides a method to delineate behaviors of membrane localized protein microclusters (45). Distribution of molecules is directly counted (i.e. bypassing the fluorescence intensity conversion in bulk imaging experiments), and the concentration ratio can be extracted. The dynamics of individual molecules is directly evaluated by tracking their individual movement trajectories instead of depending on the overall fluorescence recovery as in FRAP analysis. Trajectories of single molecules over time can be categorized, and it has been shown that molecules within condensed phase move significantly slower than their counterparts in dilute phase (Fig. 4B). Super resolution imaging is a powerful technique for illustrating features of individual molecules, but deliberations need to be taken for the compatibility of the two systems. For example, the imaging buffer for STORM experiments contains a thiol compound to enable photo-switching. However, when a His-tagged protein attaches to the membrane via interaction with DGS-NTA-Ni^{2+} lipid to mimic its membrane localization, reducing reagent can interfere with membrane attachment. Fortunately, this system is tolerant to sub-dosage of 2-mercaptoethanol to some extent while sufficient amount of photo-switching can still be achieved (45).

Another concern is the requirement of oxygen scavenger system that consists of glucose, glucose oxidase, and catalase. High concentration of glucose (~10%) may affect the ability of the molecular components to phase separate, although the influence is case by case and rigorous controls are required. In addition, the intermediate product, hydrogen peroxide, will oxidize and destroy the lipid membrane if failed to be eliminated by catalase. Therefore, the relative stoichiometry of imaging buffer components and duration of imaging time are crucial to maintain a reduced environment at all times.

Open questions

The application of 2D supported lipid bilayer system for characterization of LLPS is still at the initial stage and in the ascendant. Many important questions remain to be considered to promote the development of this system, and in return, to facilitate the thorough comprehension of this field. In current systems, the coating of membrane proteins mainly relies on NTA-Ni^{2+}-His interaction for simplicity. Clustering of membrane proteins, however, will drag synchronized movement and clustering of NTA-lipids, which may affect the overall membrane properties. Besides, the protein coating efficiency is dependent on the proportion of NTA-lipid (DGS-NTA), but the percentage of DGS-NTA itself will affect membrane fluidity as indicated by FRAP analysis (63). Continuous efforts are thus demanded to overcome this problem. For example, transmembrane proteins might be inserted into supported lipid bilayer independent of NTA-Ni^{2+}-His interaction.
The lipid components we are looking at in current studies are way too simplified comparing to those in natural conditions. Importantly, lipid compositions can change over time and in response to cellular activities, and protein-lipid interactions are often involved in signal transduction and regulation. In addition, lipid itself can undergo phase separation, which represents another fascinating research field (64). What happens if protein phase separation comes across lipid phase separation? Membrane bilayers constituting more close-to-physiology lipid compositions should certainly be taken into consideration over the long haul. The concept of “membraneless compartment” is gradually becoming a consensus, and studies performed with supported lipid bilayer increasingly uncover the relationship between protein condensates in solution and protein clusters on lipid membranes. Early evidences have already shown a direct connection between synaptic vesicle pool and its buffering surrounding—synapsin phase separation in presynaptic termini (56). In the future, it will be interesting to study how membraneless and membrane-bound compartments are coupled using a combination of 3D solution and 2D membrane systems.

Functional implications

We have so far discussed the molecular mechanisms that drive phase separation and how to characterize LLPS in solution and on lipid bilayers in vitro. Studies on in vitro reconstitution systems shed light on the biological significance of phase separation. In this section, we propose a few potential functional implications of having non-membrane enclosed biological condensates in cells.

Compartmentalization without physical barriers

Membrane-mediated molecular confinements guarantee specific proteins/nuclei acids subcellular localization thus allowing distinct functions of each organelle. However, membrane-bound organelles with limited types are insufficient to support diverse cellular processes with multiple functions. Cytoplasm should be further segregated to control each unique chemical reaction without potential disturbances. Functional proteins are often found to have their preferential sites in a cell with a sharp concentration gradient to the neighboring environment. Phase separation among different biomacromolecules facilitates spontaneous formation of different subcellular compartments without the help of lipid membranes. Since phase separation is driven by intrinsic properties of an exact protein and its binding partners, such compartmentalization can be highly specific to its inner components. The transition is achieved in a membrane-independent manner, therefore cells can simultaneously condensate different materials into various compartments, each with defined constitution and function (Fig. 5). Moreover, forming membraneless organelles would be “energy friendly” to cells because lipid biogenesis and membrane identity maintenance consume a huge amount of energy. Phase separation by molecules under physiological conditions is a natural process with no demands for extra energy, thus cells do not need to actively deliver materials towards each condensate against a huge concentration gradient. The membraneless and liquid-like properties allow a newly formed condensate to fuse with another to enlarge its size, a process which could also be energy costly if every condensate is enclosed by membranes (Fig. 5).
Achieve high local concentrations for molecular interactions and rapid chemical reactions

Self-condensation is one of the key features and probably also the most important function of phase separation. Compared to macromolecular complexes formed by traditional interaction mode, a demixed phase enables molecules to be massively enriched into a restricted subcellular region and subsequently to increase their local concentrations by hundreds of folds. This massive increase in concentration brings at least two non-negligible changes towards materials inside the condensates. First, for scaffolding proteins involved in assembly of the entire architecture, weak interactions between molecules, which usually is almost undetectable in aqueous solution, can get dramatically amplified and contribute to the properties of biological condensates (65,66). That could explain why sometimes a single amino acid substitution on a given protein, which hardly changes its behavior in homogenous solution, might severely influence its ability to phase separate. Those previously identified weak interactions should also be re-evaluated taken into the consideration of phase separation, because they may no longer be nonspecific or without any functional implications. Secondly, higher local concentration of enzymes enriched in condensates might show altered activities or kinetics during chemical reactions. If a given enzyme gets concentrated into condensed phase with an open conformation, the active recruitment or exclusion of its substrate determines whether a chemical reaction would get promoted or inhibited (Fig. 5). CaMKII is the most abundant enzyme in synapses, colocalizes with its numerous substrates in PSD and exhibits neuronal activity-dependent translocation into synapse from dendritic shaft (67). Upon kinase activation, one might foresee enhanced phosphorylation of CaMKII substrates to activate downstream signaling pathways. Actin polymerization provides another example of how phase separation can promote reaction kinetics. In nephrin and LAT systems, the amount of actin assembly is dramatically upregulated when the signaling components undergo phase separation (discussed above). It has recently been demonstrated that the increased membrane dwell time of N-WASP, in the condensed phase, promotes its association with the Arp2/3 complex and subsequently the actin polymerization rate compared to homogenous solution state (10,41,68,69). The formation of astral microtubules from centrosome is also promoted when tubulin monomers become massively enriched into a reconstituted centrosome condensate by SPD5 (66). Stress granules provide a contrasting example where protein translation is sequestered by actively “squeezing” mRNAs and some of the translatational machineries from cytoplasm into a densely packed condensates (69). Future experiments should be focused on both accurately measuring the enzyme kinetics and systemically proposing reaction theories in condensed phases.

Allow fast changes of molecules upon signaling

Membrane-enclosed structure shows reduced molecular dynamics because its inner materials are completely segregated by lipid bilayers. Condensates formed by phase separation can confine molecules to a given region but meanwhile allow them to freely exchange with their counterparts in the surrounding solution. This type of molecular dynamics provides a unique feature of phase separation-mediated condensation—to
rapidly rearrange its interior constituents in response to different stimuli. Compositional reorganization within a particular compartment can be accomplished by selectively altering the behavior of a given protein with covalent modifications that favor or disfavor its local environment. For example, Synapsin undergoes phase separation by itself and further cluster synaptic vesicles (56). This condensation can be dissolved upon Synapsin phosphorylation by CaMKII. Arginine methylation on FUS protein did not affect its phase separation ability, but dramatically decreased the hardness of FUS droplets, indicating that post-translational modification could also modulate the material properties of a given condensate (34). Phase separation might also undergo an overall weakening when the key organizer is depleted or competed off by other regulatory molecules. The dispersion of reconstituted PSD phase droplets by an alternatively spliced form of Homer1 provides another good example of biological condensate regulation (44). Since multivalent intermolecular interactions (both strong and weak) are the driving force of phase separation, one could imagine that the condensation process can be extremely sensitive to changes in the outside environment including salt concentration, pH, temperature, redox conditions etc. Tuning biomolecular interactions might bring huge influences on a condensed phase, making each condensed system a perfect biosensor that enables cell to recognize various signals and make rapid responses to them.

Sub-segregation via phase-in-phase, phase-to-phase, or surface coating

Organelles with multiple membrane layers are not commonly used in living cells. Mitochondria and chloroplasts are the only two known systems with double layers of lipid membrane which allow their inner materials to be further segregated to facilitate multistep reactions during respiration and photosynthesis. Sub-segregations within an organelle can provide new isolated regions with distinct functions but meanwhile allow each segregated part to communicate with each other. This smart design might also be adopted by membraneless condensates to support themselves with multiple functions. Sub-segregation can happen when multiple proteins co-cluster into the same condensates with one of them forming a smaller droplet and being totally embedded among the other, a phenomenon termed as phase-in-phase (Fig. 5). For example, three subcompartments (NPM1, FIB1 and POLR1E) of nucleoli in *Xenopus laevis* form distinct and immiscible liquid phases where FIB1 and POLR1E condensed into smaller-sized puncta inside single NPM1 condensate (57). A completely buried phase is isolated by outer layer proteins thus preventing potential dynamic exchange. At the same time, protein concentrations would further increase as the total volume of a droplet gets smaller, which might vastly speed up reactions inside the condensate. When sizes of two sub-segregation become similar, a layer-to-layer structure could form with two droplets sharing a common interface but each exposed to the outside environment. A condensation organized in a phase-to-phase pattern can generate multiple functional interfaces for biomolecular interactions and signaling transductions with specific orientations (Fig. 5). During germline development in *C. elegans*, two P granule proteins ZNFX-1 and WAGO-4 become phase separated from P granule to form an independent liquid phase. This newly formed phase further assembles into tri-condensation with P granule and Mutator foci in a phase-to-phase manner to spatiotemporally regulate epigenetic
inheritance during development (70). The detailed molecular mechanisms for phase-in/to-phase is still not well understood, and it is believed that both the interaction among inner materials and the surface tension of individual droplets may govern the sub-segregation process. Surface coating is another unique type of sub-segregation when some molecules only localize to the surface of a transitioned phase (Fig. 5). Recruitment of particular molecules to droplet surface might change the surface properties of a given condensate and offer it with new functions. In addition, biomolecules, even without the help of transmembrane or membrane-binding domains, can be confined into a 2D system, which dramatically alters their activities. This might be achieved when a given protein has two featured surfaces, one of which favors the inner environment of a condensed phase but the other disfavors and gets excluded. Thus, surface coating is regarded as an equilibrium between protein attraction and exclusion from materials inside the condensate. Surface coating to a selected phase droplet may also affect its material properties. Molecules on the condensate surface can regulate dynamic exchange of its inner materials, influence free diffusion of small molecules, and even alter fusion process of droplets. A recent study reported that EPG-2, a scaffold protein in C. elegans P granule, can specifically decorate the surface of SEPA/PGL-1/-3 droplets and modulate the condensate properties (71).

Direct communications between membraneless and membrane-bound organelles

Membraneless organelles formed by phase separation could communicate with membrane-bound organelles via direct interactions (Fig. 5). Such communication might help to specify the localization of membrane-bound organelles, to reorganize protein distributions on membrane surface and to introduce new functions to organelles. TIS11B, an RNA-binding protein, for example, forms membraneless granules that directly attaches to ER (72). TIS granules specifically retain certain mRNAs and exclude others to enable accurate protein translation in ER. RNA granule is also observed to associate with late endosomes residing close to mitochondria in neuronal axons to regulate local synthesis of axonal proteins (73).

Phase separation and evolution

Previous subsections discuss the functional implications of phase separation in terms of offering a living cell with multiple functions. Here we postulate on possible biological importance of phase separation from the angle of life evolution. It is hard for one to imagine that the earliest form of life directly starts with membrane-bound organelles as lipids are not typical information carriers like nuclei acids. In addition, the biogenesis of lipids depends on other molecules with catalytic activities, such as protein or RNA. The origin of life is believed to depend on RNA because it carries the genetic information and possesses enzymatic activity which many allow them to self-reproduce. RNA is so unique as it forms a long chain with multiple binding sites for other RNAs or proteins, which might also explain why it could always easily undergo LLPS with different RBPs or even by itself. Compartmentalization mediated by phase separation may reveal how proteins and nuclei acids assemble into condensed bioreactors in ocean before the emergence of lipid membranes. Indeed, membraneless condensation is observed not only in higher animals but also in ancient cyanobacteria, indicating a common and conserved
biological process that could be shared by all the living creatures on earth (46). We refer readers to several recent reviews (6,74,75) that also discussed about the emerge of LLP process during the evolution.

Conclusions and perspective

In cells, biomaterials can be organized into membrane-bound or membraneless compartments. Significant progresses have been made over the past decade in understanding the mechanisms underlying the formation and organization of non-membrane enclosed organelles. Many of these systems are driven by LLPS via which collections of molecules demix from the bulk solution/cytoplasm to form biological condensates. Modular domain proteins and intrinsic disorder containing proteins exhibit multivalent intermolecular interactions either via specific, high affinity interactions or weak adhesions that drive phase separation. Analysis of features of molecules involved in LLPS has started to reveal sequence determinants in intrinsically disordered regions that promote phase separation. Although our understanding is still rudimentary, it is clear that certain sequence patterns are heavily involved and can determine the material properties of a condensate. A large collection of methods has been developed to study the dynamics, composition and physical properties of condensed droplets in vitro. We still need more quantitative assays, measurements and descriptions for future phase separation studies. For instance, theoretical work is required for understanding the underlying physical and chemical principles of LLPS. Bioengineering tools may be designed to precisely control phase separations in vitro or in vivo. Another appealing, yet very challenging, idea is to reveal the atomic details of the condensed droplets. In particular, could there be a single structure that might be “solved”? Cryo-EM studies have been conducted on some condensates trying to answer this question, although little success has been achieved so far. This is somewhat expected since many of the phase separations are contributed by intrinsically disordered elements. It is difficult to imagine how the multivalent interactions might be restricted to oligomers of homogenously distributed sizes. Nevertheless, we cannot rule out the possibility that a core structure of defined stoichiometry and conformation might exist among other flexible and heterogeneous structures, especially where the condensate formation is driven by modular domain interactions. To “solve” the atomic structure of phase condensates might be too optimistic at current stages, but it is possible that we might reveal the molecular organizations within the condensates using cryo-ET and cryo-EM techniques. Is there a layered organization within the phase droplets? How do molecules assemble into supramolecular complexes that phase separate from the bulk solution? Results from in vitro characterizations of phase droplets would provide insights into how the macroscopic properties of condensates might contribute to their biological functions in cells. There remains much to answer about the biological condensates. How is the enzyme kinetics regulated in condensed phase? Recent studies using nephrin and LAT systems have shown that increased dwell times of enzymes in condensed phase lead to faster reaction rates (62,68). Is this a general mechanism for other molecular systems? Compared to membrane-bound organelles, LLPS-mediated membraneless structures have their own advantages; but it also brings many potential problems. For example, how to achieve specificity in organization? How to prevent unwanted
fusion or mixing without a membrane barrier? How does sub-segregation happen inside an organelle and how to maintain this specific pattern? Is this physiologically relevant and functionally regulated? Will the concept of phase separation help us to re-examine many diseases that are hard to be explained by physical chemistry principles of dilute solution systems? For instance, alteration in the material properties of many neuronal protein condensates might contribute to neurodegenerative diseases. The concentration-dependence of phase separation might help explain the dosage sensitivity of SynGAP, a negative regulatory protein of PSD assembly, in psychiatric diseases. Answers to these questions will provide us with in-depth insights into mechanisms underlying the formation and regulation of biological condensates in cells and to understand how nature evolves this type of compartmentalization in life. In the end, although the list of non-membrane bound organelles formed by LLPS continues to expand, researchers should always ask themselves — is the LLPS-driven protein condensation observed *in vitro* biologically relevant? If so, what is its contribution to cellular functions? Can the *in vivo* observations be explained by mechanisms other than liquid phase separation (76)? Tailored experiments need to be designed in order to distinguish between these possibilities. Nonetheless, it is assured that LLPS-mediated biological condensate formation is an emerging life science research field with numerous exciting opportunities.

**Acknowledgments**

Work in our laboratory is supported by grants from RGC of Hong Kong (AoE-M09-12 and C6004-17G) and a grant from Simons Foundation for Autism Research (510178). ZF is supported by Mandatory Provident Fund Scheme and is a Junior Fellow of IAS at HKUST. MZ is a Kerry Holdings Professor of Science and a Senior Fellow of IAS at HKUST.

**Competing interests**

The other authors declare that no competing interests exist.
References


proofreading modulate Ras activation by SOS. Science 363, 1098-1103
FIGURE LEGENDS

Figure 1. Types of multivalent interactions driven by intrinsically disordered elements in liquid-liquid phase separation (LLPS) systems.

(A) Phase diagram constructed by varying protein concentration and storage conditions such as buffer reagents and temperature. Solid line depicts the boundary at which molecules reach their solubility limit and become immiscible with the surrounding solution. Grey box highlights confocal image showing the homogeneous solution state of NR2B C-terminal tail (labeled with Alexa Cy3) in the absence of PSD scaffold proteins. Conditions within the spinodal curve (indicated as dashed line) are where spinodal decomposition occurs. Example of fluorescence image, highlighted by green box, showing that the membrane tethered NR2B tail (labeled with Alexa Cy5) formed clusters on supported lipid bilayers upon the addition of major PSD scaffold proteins. Phase separation is only observed in the presence of a nucleation process when conditions lie in between the binodal (indicated as solid line) and spinodal curves. Representative image, highlighted by yellow box, showing the clustered state of NR2B tail (labeled with Alexa Cy3) in 3D solution in the presence of major PSD scaffold proteins (adapted from Ref. 45). Scale bar, 10 μm.

(B) Aromatic residues in intrinsic disorder containing proteins are involved in pi-pi or cation-pi interactions with positively charged residues such as Arg and Lys. RGG repeats are frequently found in low complexity regions (LCR).

(C) Patterned charge distributions to facilitate electrostatic interactions between oppositely charged residues.

(D) Secondary structural elements are involved in multivalent intermolecular interactions, such as the kinked cross β sheets formed by a segment of FUS LCR (PDB code:6BWZ).

Figure 2. Types of multivalent interactions driven by modular domains in LLPS systems.

(A) Interaction network of N-WASP, nephrin and NCK.

(B) Schematic representations showing the network of multivalent interactions involving major PSD proteins. Solid line indicates direct modular domain interactions. Dashed line indicates indirect recruitment of actin filaments via Shank3 and Homer proteins.

(C) Schematic interaction network of presynaptic active zone proteins RIM and RIM-binding protein together with the cytoplasmic tail of the N-type voltage gated Ca^{2+} channel (NCav).

(D) RNA binding proteins, hnRNPA1 for example, binds to RNA triplets via RNA recognition motifs (RRM). Multiple RRM-RNA interactions, albeit of low affinities, together provide multivalency to drive phase separation.

Figure 3. Techniques for characterizing of condensed phase formed in 3D solution.

(A) Differential interference contrast (DIC) (left image) coupled with fluorescence imaging (middle and right images) of the phase droplets, multiple labeling of different components demonstrate their colocalization.

(B-D) Dynamic properties of the condensed phase.

(E) Fluorescence intensity based absolute concentration estimation (adapted from (44,45)). A standard curve of fluorescence intensity to dye concentration is initially generated for calibration. Z direction scanning is performed to determine the proper focal plane for concentration estimation. At each Z stack, the fluorescence intensity distribution is plotted.
Within the Z dimension of selected droplets, average fluorescence intensities are then compared across different layers. In a given system the fluorescence intensity is constant regardless of the droplet size, and therefore the absolute protein concentration within a condensed phase can be calculated from the standard curve.

**Figure 4. Condensed phase formed on 2D supported lipid bilayer.**

(A) Schematic diagram of microdomain formation on 2D supported lipid bilayer. Membrane proteins homogeneously distribute on supported lipid bilayer via the tethering of His-tag to Ni$^{2+}$-NTA decorated lipids. Protein clusters are observed on lipid bilayers after the addition of other components to drive phase separation.

(B) Stochastic optical reconstruction microscopy (STORM) analysis of membrane proteins, the cytoplasmic tail of NCav as an example, on supported lipid bilayer (adapted form (45)). Image captured under total internal reflection fluorescence (TIRF) microscopy mode first sketches the contours of the condensed phase, which turns out to perfectly overlap with the image reconstructed from STORM analysis. Trajectories of individual molecules are followed by single molecular tracking assay, both inside and outside the condensed phase. Direction of movement is marked by gradient color from black to red.

**Figure 5. Biological functions of LLPS-mediated membraneless compartments.**
Figure 2

A

Nephrin

PRM

N-WASP

SH2 SH3

NCK

Nephrin

B

NMDAR

PSD-95

SynGAP

GAP

Shank3

GKAP

Homer

PSD

C

NCav

PRM

PNGY

ZP

PDZ

CSA

PRM

CUB

SH3

PN

SH3

SH3

Actin