The molecular language of membraneless organelles

Edward Gomes¹ and James Shorter¹*.  

From the ¹Department of Biochemistry and Biophysics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104. U.S.A.

Running title: Molecular language of membraneless organelles

*Correspondence: jshorter@pennmedicine.upenn.edu

Keywords: subcellular organelle, organelle, RNA-binding protein, chaperone, liquid-liquid phase separation, nucleolus, stress granule, Cajal body, amyotrophic lateral sclerosis (ALS), disaggregate

Abstract
Eukaryotic cells organize their intracellular components into organelles that can be membrane-bound or membraneless. A large number of membraneless organelles, including nucleoli, Cajal bodies, P-bodies, and stress granules, exist as liquid droplets within the cell and arise from the condensation of cellular material in a process termed liquid-liquid phase separation (LLPS). Beyond a mere organizational tool, concentrating cellular components into membraneless organelles tunes biochemical reactions and improves cellular fitness during stress. In this review, we provide an overview of the molecular underpinnings of the formation and regulation of these membraneless organelles. This molecular understanding explains emergent properties of these membraneless organelles and shines new light on neurodegenerative diseases, which may originate from disturbances in LLPS and membraneless organelles.

However, cells also organize components into non-membrane bound organelles, suggesting that Oparin’s coacervate idea deserves a second look (2–4). In fact, many cellular organelles are condensates of protein, nucleic acid, or both. In the nucleus, these include nucleoli, Cajal bodies, nuclear speckles, paraspeckles, histone-locus bodies, nuclear gems, and Promyelocytic leukemia (PML) bodies (5–7). The cytoplasm also contains several membraneless organelles, including P-bodies, stress granules, and germ granules (6,8). In this review, we highlight advances in our understanding of the molecular language of these membraneless organelles with respect to how they form, what functions they serve, what rules regulate them, and how their dysregulation may contribute to human disease.

Membraneless organelles are liquids that organize the cell
Early evidence that membraneless organelles may behave as liquids came from study of the C. elegans germ granule, or P granule. P granules are collections of RNA and RNA-binding proteins (RBPs) that accumulate on the posterior side of the C. elegans zygote before the cell divides into a posterior and anterior cell (9,10). By fluorescently labeling a constitutive P-granule protein, Brangwynne, Hyman, and colleagues discovered that P granules display liquid-like properties: the granules are spherical, fuse with one another, deform under shear stress, have fast internal rearrangement as assessed by recovery after photobleaching, and drip off the surface of the...
nucleus like a liquid (9). These observations led to the conclusion that P granules are liquid droplets inside the cell that form via a process called liquid-liquid phase separation (LLPS) (Fig. 1A). Burgeoning evidence now suggests that a wide range of membraneless structures—from ribonucleoprotein (RNP) granules like the nucleolus to centrosomes and clusters of signaling molecules on membranes (Fig. 1B)—exhibit liquid-like properties and coalesce through a LLPS mechanism (11-18).

**Phase separation and transition: liquids, gels, and crystals**

The example of a salad dressing illustrates a simplified version of LLPS (16). Even after a vigorous shake, the oil and water in the salad dressing separate into a demixed two-phase system that has a lower free energy than the fully mixed state. This type of demixing is often called LLPS or a phase transition. Two types of interactions contribute to the process: the homotypic interactions between two molecules of oil or two molecules of water and the heterotypic interactions between a water and oil molecule. Entropy-driven mixing is disfavored due to the higher strength of the homotypic interactions over the heterotypic interactions, which leads to a phase separated two-state system of lower free energy (16,19).

This simple example of phase separation extends more generally to solutions of polymers, for which the physics of LLPS has been well described (20-23). As polymers, proteins and nucleic acids are subject to the same underlying physics of LLPS (Fig. 1A) (19).

The concept that proteins undergo phase transitions is not novel, especially not to protein crystallographers whose work relies on coaxing proteins into crystals and who often observe gels, aggregations, and phase separation of proteins as side products of the crystallization process. As an example, lysozyme undergoes LLPS, gelation, and crystallization depending on certain conditions of temperature, salt concentration, and protein concentration (24,25). Wang et al. have more recently demonstrated that oligomeric peptides undergo LLPS in vitro, with LLPS stimulated by low temperature, crowding agents such as polyethylene glycol (PEG), and pH close to the isoelectric point (26).

Together, the observations that both peptides and well-folded globular proteins undergo liquid demixing in vitro indicate that many if not all proteins can undergo LLPS in conducive environmental conditions. It is not our intention to trivialize the finding that proteins undergo LLPS but rather to point out that this property extends to all proteins and polymers in general. Indeed, RNAs can also phase separate in vitro (27,28). The critical question then is to understand what makes phase separation biologically consequential and achievable within the cellular environment, which we address in the next two sections.

**Biological consequences of phase separation**

Life has harnessed the ability of proteins and other biopolymers to phase separate into liquids and in some cases, further transition to gels and solids. We highlight some of the emergent biological properties of phase-separated compartments below.

**Organization**

First, LLPS serves as a dynamic organizing principle that enables cells to spatiotemporally compartmentalize specific biochemistry, provide specific infrastructure, or both (29). LLPS enables compartmentalization within a boundary while still allowing for both internal rearrangement and diffusion of biomolecules into and out of the compartment (29). For example, neurons have postsynaptic densities (PSDs), which are protein-rich compartments on the intracellular side of the postsynaptic plasma membrane that undergo remodeling in protein composition in response to long-term potentiation, i.e. the persistent strengthening of synapses due to recent patterns of activity, which underlies learning and memory (30,31). Zheng et al. propose a phase-separation model for the formation and remodeling of PSDs with supporting evidence that two major protein components of the PSD, SynGAP and PSD-95, can form liquid droplets in vitro (30). Because neurons have large surface areas, and consequently a large space for protein diffusion, spatially confining molecules involved in the same biochemical pathways poses a challenge. The formation of a PSD through a phase-separation mechanism allows neurons to
locally concentrate protein without having to globally upregulate protein synthesis.

In a similar vein, neuronal mRNP-granule assembly mediated by LLPS of low-complexity domains of ataxin 2 is critical for long-term memory formation in Drosophila (32,33). This mRNP-granule-driven mechanism of long-term plasticity differs from how another RBP, CPEB/Orb2, underlies long-term potentiation. CPEB/Orb2 forms self-templating amyloid or prion conformers that directly stimulate synaptic mRNA translation (34-36). Thus, different RBPs may function via distinct assembly mechanisms and different material phases to encode long-term memories.

Neurons also utilize LLPS for functional purposes in the tight but dynamic clustering of neurotransmitter-laden synaptic vesicles (SVs) at synapses (37). These clusters serve as a replenishable pool of SVs, which can be rapidly mobilized for exocytosis during periods of heightened synaptic activity. It had remained unclear how SVs could remain motile while being confined in these clusters. It is now suggested that the physiological mediator of SV clustering, synapsin, forms a liquid phase that connects and recruits SVs in these clusters (37). This phase can be rapidly dispersed via synapsin phosphorylation by calcium/calmodulin-dependent protein kinase II (CaMKII) (37), which would emancipate SVs en masse for rapid bursts of exocytosis upon synaptic stimulation.

Phase separation is also implicated in transcription, both at the level of transcriptional activation and repression. In Drosophila polytene cells, for instance, the application of stresses like heat shock induces formation of transcription puffs at the sites of heat-shock protein (Hsp) genes where active transcription occurs (38-40). Studies on the recruitment of proteins, such as RNA polymerase II and topoisomerase, to these sites led Zobeck et al. to originally propose a model in which a porous transcription compartment forms at Hsp gene loci (40). In retrospect, these data support a phase-separation model for transcriptional control (41). A phase-separation model has also been proposed for the clustering of enhancer elements in DNA together with coactivator proteins to form super enhancers (42-45). Moreover, prion-like domains (PrLDs) of transcription factors can cluster into dynamic hubs that stabilize DNA binding and recruit chromatin-remodeling factors and RNA polymerase II (46-49). These hubs can manifest as phase-separated structures at elevated transcription-factor expression levels (46-49). LLPS also functions in transcriptional repression. For example, heterochromatin-mediated gene silencing is driven via compartmentalization of condensed chromatin into phase-separated liquid droplets formed by heterochromatin protein 1α (50,51). The involvement of phase separation in regulating genome architecture and transcriptional output provides an exciting new avenue of research.

Finally, it is important to note that not all membranous organelles are fully liquids; many likely exist along a continuum from more liquid-like to more gel-like, depending on the interaction strength of the constituents (29). On the gel side of the spectrum is an additional example of an organizational role for phase separation: the nuclear-pore complex (NPC). The central channel of the NPC is a gel-like, phase-separated structure that organizes the cell by acting as barrier to diffusion of molecules above 30-40 kDa into or out of the nucleus (52-54). Similar selective-permeability barriers also form at the base of primary cilia (55,56). Thus, depending on the structural, functional, or organizational need, the cell employs phase separation that spans from more dynamic, liquid-like compartments to more static, gel-like compartments. For example, globular S-crystallin proteins of different sizes assemble into a gel of varying density, thereby establishing a refractive-index gradient that forms the parabolic lens of the squid eye (57). At the extreme end of the spectrum, stable solid phases comprised of amyloid or prion conformers are utilized as with CPEB/Orb2 prions in long-term potentiation (35,36), Xvelo amyloids in Balbini bodies that specify germline identity (58), or transient Rim4 amyloids in meiotic control (59-61).

**Tuning reactions**

Membraneless organelles likely tune and accelerate biochemical reactions *in vivo* in a manner akin to how various synthetic chemical reactions can be accelerated in microdroplets *in
vitro (62). The specific microenvironment within the liquid phase may serve to tune reaction rates and biochemical activities inside membraneless organelles. Phase separation can increase the concentration of certain molecules within dense liquid condensates compared to the surrounding solution by as much as two orders of magnitude (13). Given the dependence of reaction rates on reactant concentrations, achieving a locally high concentration of molecules due to phase separation can be a biological mechanism for increasing reaction rates. This prediction has been demonstrated in vitro by using an aqueous two-phase system to concentrate RNA substrate into liquid droplets and measuring the rate of substrate cleavage by a ribozyme (63). Concentrating the RNA and ribozyme into dense liquid droplets increased the reaction rate, suggesting that coacervation inside a cell can have a similar effect (63). Nuclear RNP granules called Cajal bodies provide one such in vivo example. Cajal bodies are the sites of assembly of the U4/U6-U5 snRNP complex, which forms eleven times more efficiently within Cajal Bodies than in the surrounding nucleoplasm (64).

Beyond increasing reaction rates, LLPS may also tune a biochemical process by acting as a filter to regulate which molecules enter a liquid droplet and which molecules stay out. In the case of Cajal bodies, only the fully formed U4/U6-U5 tri-snRNP complex can leave the nuclear body whereas the di-snRNP complex cannot, which enables selective accumulation of a reactant into a confined space (65). A model of membraneless organelles acting as a filter also applies to the partitioning of RNA, which can tune the type of RNA chemistry that occurs in the organelle. RNAs can influence the compositional specificity of intracellular phases, with Langdon et al. showing that RNA structure and RNA-RNA interactions affect which RNAs partition into liquid droplets (66). There is other evidence that the length of RNA affects which RNAs become more concentrated in liquid droplets, with longer RNAs partitioning more effectively into the droplet phase (63). Meanwhile, Nott et al. have discovered that the microenvironment within phase-separated liquid droplets favors melting of double-stranded nucleic acids, stabilization of ssRNA secondary structure, and partitioning of RNA into droplets based on the stability of folding rather than the length (67,68). The discrepancy in the length dependence of RNA maybe because Nott et al. (67) used liquid droplets arising from RBPs for their study whereas Strulson et al. (63) used an aqueous two-phase system. Regardless, these findings present important steps toward understanding the molecular determinants of phase separation, some of which will be discussed later. The length dependence of RNA partitioning into liquid droplets is particularly interesting in light of data that local protein concentration and RNA length alter the binding mode and RNA-remodeling activity of the RNA helicases LAF-1 and DDX3X, both of which partition into membraneless organelles in vivo (69). Examining how tuning protein and RNA partitioning into membraneless organelles can modulate organelle activity will be an important avenue of further research.

Cellular fitness

One of the emergent properties of LLPS is that it is environmentally tunable and can thus play a cytoprotective role by sensing and responding to stress (29). Protein folding within the crowded intracellular environment is a challenge that is accentuated by cellular stresses that may trigger protein misfolding (35). The formation of reversible, phase-separated structures enables cells to store their proteins and RNAs temporarily in a manner that allows their rapid recovery after dissipation of the stress. In yeast, the prion protein, Sup35, acts as a pH sensor and forms liquid condensates that undergo a phase transition to gels in response to stress-induced drop in cytoplasmic pH (70). Upon stress, the formation of these Sup35 gels are protective and allow the yeast to better recover from the stress (70). In a similar manner, yeast polyA-binding protein (Pab1) acts as a sensor for pH and thermal stresses (71). In response to stress, Pab1 releases its bound mRNAs (which enables translation of key stress-response transcripts) and forms reversible, phase-separated hydrogels (71). Complementary findings have been made with another yeast RBP, Pub1 (72). Indeed, in response to thermal stress, yeasts form assemblages of functional proteins held together by weak interactions (73).
These assemblages dissolve when the stress subsides, allowing for recovery of cellular proteins without widespread misfolding or degradation (73). Depending on the type of stress, assemblage dissolution can be spontaneous or may require protein disaggregases, such as Hsp104 (70,72,74). This controlled and reversible phase separation of mature proteins likely represents an adaptive strategy to stress, and contrasts with previous models where stress-induced aggregates were thought to be disordered accumulations of misfolded, denatured proteins (73). Results in yeast also extend to mammalian cells, where stress-induced stalling of translation leads to the condensation of protein and RNA into stress granules, which are dissolved after stress by Hsp110, Hsp70, and Hsp40 (8,75). Importantly, stress granules protect against cellular senescence by sequestering PAI-1, an established promoter of senescence (76). Overall, the ability of cells to form assemblages of proteins and nucleic acids in response to stress appears to be a conserved mechanism for cells to weather deleterious conditions.

Molecular language of phase separation

Although many if not all proteins can undergo phase transitions in vitro, not all proteins do so under physiological conditions. One of the common features of proteins that undergo phase separation in a biologically meaningful manner is the presence of multivalent binding domains, which we discuss below.

Multivalency: the key principle

The overarching property of proteins that phase separate is multivalency in interacting partners. Li et al. demonstrated this important principle by creating model proteins composed of tandem repeats of either a ligand or its binding partner (13). Combining repeats of an SH3 domain and its proline-rich motif (PRM) binding partner readily initiated phase separation of the proteins into liquid droplets (13). Increasing the interaction strength of the proteins by increasing the number of repeats of the two domains led to gelation of the liquid droplets (13). In this system, it is specific multivalent protein-protein interactions that drive phase separation.

Intrinsically-disordered domains

Multivalency can arise from protein-protein interactions between ordered domains (13). However, intrinsically-disordered domains represent another method for achieving multivalency and often contain multiple short-linear motifs (SLiMs) that mediate protein-protein interactions (13). Our understanding of the molecular determinants of phase transitions increased with the discovery that biotinylated isoasole (b-isox) reversibly precipitates many components of RNP granules (77,78). The presence of low complexity, PrLDs and RNA-recognition motifs (RRMs) were common denominators for many of the proteins precipitated, and in the case of TIA-1, the presence of a PrLD was sufficient for precipitation (77). This observation highlighted the importance of disordered regions, especially PrLDs, as a determinant of phase separation.

PrLDs represent a subset of low-complexity domains that show similar amino-acid composition to yeast prion domains (79-85). These domains are enriched in polar, uncharged amino acids, such as asparagine (N), glutamine (Q), tyrosine (Y), and serine (S), as well as glycine (G) (79-85). Yeast prion domains enable certain yeast proteins such as Sup35, Ure2, and Rnc1 to form prions, infectious proteins that usually propagate via self-templating amyloid forms (34). Typically, amyloid fibrils are highly stable cross-β structures, which represent an extreme form of phase separation to solid phases that are difficult to reverse (35). Indeed, specialized protein disaggregases such as Hsp104 or Hsp110, Hsp70, and Hsp40 are typically required to reverse their assembly (86-89). The precise features of prion domains and PrLDs that enable them to form phase-separated liquids, gels, or prions are still being delineated (79,90-95).

In humans, of the 240 genes that encode proteins with a PrLD, a remarkable 72 encoded RBPs (84). These include FUS, TDP-43, TAF15, EWSR1, hnRNP A1, hnRNP A2, and TIA-1, which are components of RNP granules that are heavily implicated in neurodegenerative disease (82), and are precipitated by the b-isox
compound (77). At high protein concentrations, the PrLD mediates the phase transition of FUS and hnRNPA1 into hydrogels in vitro that bind the PrLD of other RNP granule components. This observation led Kato et al. to posit that the ability of low-complexity domains to reversibly form labile amyloid-like states lies at the crux of RNP granule formation (77). Numerous studies have since corroborated the importance of intrinsically-disordered domains, especially of PrLDs, in the formation of phase-separated membraneless organelles (11,96-103). In some cases, deletion of the PrLD of key RBPs (e.g. TIA-1, FUS) completely abrogates the formation of RNP granules (102,104,105). The natural tendency of PrLDs to engage in promiscuous interactions and aggregation promotes phase separation.

Evidence has also emerged that PrLDs may interact with another type of intrinsically disordered domain, termed RGG domains to drive phase separation (106-110). RGG domains are enriched for arginine and glycine residues (111), can bind RNA and phase separate (112,113), and are often found in RBPs with PrLDs (85). Indeed, for FUS and related RBPs, LLPS is elicited effectively via multivalent interactions between PrLD tyrosines and RGG arginines (106-108). These contacts are, in turn, modulated by negatively charged residues (107). Glycines confer liquidity, whereas glutamines and serines elicit gelation (107). Thus, a precise molecular grammar for phase separation by FUS and related RBPs begins to materialize (107).

RNA- and DNA-binding domains

RBPs present a special class of proteins that have biologically relevant phase behaviors. Many membraneless organelles are RNP granules that perform various RNA-processing activities, consist of RBPs and RNA, and assemble via LLPS of RBPs and RNA (8,114). The RBPs within these granules contain multiple multivalent domains including RRM and intrinsically-disordered regions, which work together synergistically to modulate phase behavior (106-108,110,113,115). For many of these RBPs, the purified proteins alone undergo LLPS in vitro (11,101,116,117), and the intrinsically-disordered regions of these proteins are sufficient for droplet formation (99,118).

However, phase separation by the intrinsically-disordered region alone can lack the additional levels of regulation that arise from the presence of additional multivalent domains like RRMs, RGG domains, and oligomerization domains (106,107,110,113,115,119). The ability to bind to multivalent scaffolds, such as DNA and RNA, through RRMs, zinc fingers, or other nucleic-acid binding domains present another common characteristic of proteins that undergo LLPS (29). The role of RNA as a scaffold for phase separation is evident from studies on several RBPs, including FUS (103) and Whi3, a fungal RBP that regulates nuclear division and cell polarity (120). In vitro, RNAs that bind Whi3 promote Whi3 phase separation (120) and encipher RNP granule identity (66). Mutations in the Whi3 RRM that abrogate RNA binding also prevent RNA-stimulated phase transitions of Whi3, suggesting that the RRM enables multiple Whi3 proteins to bind to the same RNA (120).

Oligomerization domains

Protein valency increases with the presence of oligomerization domains. For example, TDP-43, a highly expressed nuclear RBP, contains an N-terminal domain that forms oligomers (117,121-124). Recently, Wang et al. established that polymerization of the N-terminal domain promotes LLPS of TDP-43 in vitro and that a single phosphomimetic mutant in the N-terminal domain can reduce the propensity of TDP-43 to phase separate (117). The ability of oligomerization domains to nucleate a locally high concentration of a protein to promote phase separation has been used by Shin et al. to form optogenetically controlled liquid droplets in vivo (125). Here, the intrinsically-disordered regions of several RBPs are fused to Cry2, a protein that oligomerizes in response to blue light (125). Oligomerization of Cry2 elicited by blue light nucleates intracellular droplets of the fusion proteins (125). Thus, environmentally-responsive oligomerization domains can promote phase separation in response to specific environmental cues.

Weak interactions maintain membraneless organelles in phase
The essential physics of polymer phase separation are well established, and help inform biological phase separation (19). Concentrating molecules into a confined space can carry an energetic cost. Numerous weak interactions work together to counteract the entropic cost for phase separation as well the interfacial free energy cost to create a phase boundary. The molecular interactions found to be important in phase separation include π-π stacking, cation-π interactions, charge-charge interactions, and transient cross-β contacts (Fig. 2).

π-π interactions  
Aromatic residues tyrosine (Y), tryptophan (W) and phenylalanine (F) as well as residues arginine (R), glutamine (Q), asparagine (N), aspartic acid (D), and glutamic acid (E) contain delocalized π electrons in their side chains that can engage in π-π stacking (126). Work to understand the sequence features of phase-separating proteins has uncovered π-π interactions as critical (126). Using a comprehensive mutagenesis approach, for example, Pak et al. uncovered that phase separation of Nephrin intracellular domain depends strongly on the presence of tyrosine residues, as missense mutations to those residues reduced the ability of the protein to form liquid droplets in cells (127). Similar observations have been made with FUS and hnRNPA2 PrLDs in vitro (128,129). The gel-like state of the nuclear-pore complex results from π-π interactions between phenylalanine residues in the FG repeats of nucleoporins (52,53). Remarkably, Vernon et al. have established that long-range π-π contact propensity alone can identify the majority of known phase-separating proteins (126), highlighting the critical role for π-π interactions in LLPS.

Cation-π interactions  
Cation-π interactions occur between positively-charged amino acids lysine and arginine and electron-rich aromatic groups. These interactions have also gained importance as drivers of LLPS (106-108,110,113). For the RNA-helicase, Ddx4, cation-π interactions between FG and RG regions of the protein are drivers of protein phase separation in vitro and in vivo (130). As a caveat, interaction between F and R could also include π-π interactions, which likely contributed to the phase separation of Ddx4 as well. Surprisingly, short-range cation-π interactions are strong enough to overcome long-range charge-charge repulsion and cause two positively charged polymers to coacervate in vitro (131). It is also interesting to note that an emergent property of multiple cation-π and π-π interactions that drive LLPS is the ability to melt nucleic-acid duplexes by disrupting the π-π interactions that maintain them (67).

Charge-charge neutralization  
Charge-charge interactions have also gained attention as important drivers of phase separation. Oppositely charged polymers when brought together can coacervate into liquid droplets through charge neutralization, as has been shown for mixtures of RNA and cationic peptides (4,132). While this may be a simplified artificial system, the phenomenon of long-range, charge-charge interactions driving phase separation has also been observed in proteins in vitro and in vivo (127,130). An emerging theme is that it is not the presence of charged residues per se, but rather the arrangement of charged residues into stretches that is important for phase separation (127,130,133). Working in this manner, clusters of charged residues act akin to a multivalent domain to promote phase separation.

LARKs and transient cross-β contacts  
Several RBPs that undergo LLPS contain PrLDs. In the RBP FUS, for example, a portion of the PrLD forms fibrils in which stretches of amino acids assemble into intermolecular cross-β sheets as typically found in amyloid fibrils (35,95). However, recent crystallographic studies of fibrils formed by short segments of the PrLDs of RBPs that undergo LLPS have uncovered a structural difference compared to classic amyloid fibrils (134-136). While amyloid fibrils tend to have cross-β sheets with interdigitated amino acids that form steric zippers, fibrils formed by short segments of PrLDs of RBPs that undergo LLPS have kinked cross-β sheets termed low-complexity aromatic-rich kinked segments.
(LARKS) (134). These kinked β-sheets are less thermodynamically stable than the β-sheets of amyloid fibrils, and proteins with PrLDs enriched for LARKS are found in membraneless organelles that assemble via LLPS (134). Together these findings suggest that weak, transient cross-β contacts might contribute to LLPS whereas more stable cross-β contacts contribute to pathological amyloidogenesis.

Modulators of phase separation

One of the fundamental principles of a living organism is the ability to adapt to change. Cells must constantly tune their biochemistry in response to environmental cues, and the membraneless organelles within a cell must similarly be responsive to intra- and extra-cellular signals. To regulate phase separation, cells rely on several processes including post-translational modifications and seeding mechanisms.

Post-translational modification (PTMs)

PTMs provide cells with a powerful means to facilitate or antagonize LLPS in response to environmental signals (137). Indeed, SLiMs often mediate protein-protein interactions that drive phase separation and are frequently the target of regulation by PTMs (13,138). PTMs can promote LLPS, for example, by increasing the effective valency of a protein. In the nucleus, there are membraneless organelles called PML bodies for which the PML protein acts as a scaffold (Fig. 1B). SUMOylation of PML is necessary for proper formation of PML bodies because SUMO acts as a binding ligand that recruits other proteins, such as Daxx, into the membraneless organelle (139-141). Similarly, tyrosine phosphorylation of the protein nephrin promotes phase separation of nephrin with the protein NCK because the phosphotyrosine acts as a docking site for NCK (13). Phosphorylation of serines in the FUS PrLD fluidizes FUS droplets (115,142). On the other hand, phosphorylation can promote the disruption of phases as with the dissolution of Rim4 assemblies by Lmc2 (60) and the dissolution of various nuclear membraneless organelles during mitosis by Dyrk3 (143). Likewise, arginine methylation of the RBPs Ddx4, hnRNPA2, and FUS can antagonize phase separation (99,106,113,130). This list of PTMs involved in regulation of phase transitions is by no means exhaustive, but rather represents a small subset of the numerous ways that PTMs can modulate LLPS.

Seeding mechanisms

Phase transitions are concentration-dependent, switch-like phenomena that occur above a certain local critical concentration (29). Cells can promote phase separation through a nucleator that seeds a locally higher concentration of certain biomolecules to reach the necessary critical concentration. We highlight three nucleators below:

RNA

Although RBPs receive a lot of attention in phase separation, RNAs also play key roles in the formation of various membraneless organelles (28,66,105). The function of several membraneless organelles is intimately centered around RNA, such as mRNA decay in P-bodies, mRNA storage in stress granules, mRNA splicing in nuclear speckles, and rRNA synthesis in nucleoli (29). Indeed, RNA acts as a potent, biologically important nucleator of intracellular phase separation. As examples, RNA induces the phase separation of MEG-3, a key scaffold protein in C. elegans P granules (144), stalling of translation during a stress response exposes free mRNAs that act as nucleators for stress granules (28,145), and Men ε/β noncoding (nc) RNAs seed the formation of nuclear paraspeckles (146,147). It is noteworthy that there are many ncRNAs for which biological functions are not well known (148). It is possible that these RNAs regulate phase transition events in cells like the Men ε/β ncRNA.

Poly(ADP ribose)

Besides RNA, the cell also utilizes RNA-like molecules to seed LLPS. One such molecule is Poly(ADP ribose) or PAR, a polymer of ADP-ribose monomers that is involved in the formation of several stress-triggered membraneless organelles (149). For example, PAR recruits transcription factors to the heat-shock protein locus in Drosophila polytene cells in response to heat shock and
recruits DNA-damage repair factors to sites of DNA damage (40,101,150,151). PAR is both a PTM and an RNA-like scaffold that assembles proteins, including FUS and TDP-43, into membraneless organelles via LLPS (101,150,152). PAR is also found in stress granules (153). Thus, PAR has a wide-reaching role in stress-triggered assembly of membraneless organelles.

**Polyphosphates**

In light of the importance of charge-charge interactions and poly-anion seeds like RNA in the molecular language of phase separation, it seems plausible that other polyamions, like polyphosphates for instance, may also act as seeds for phase separation. Cremers et al. have previously elucidated a role for polyphosphates in nucleating amyloids (154), while Racki et al. have uncovered that polyphosphate granules assemble and coalesce during starvation-induced stress response in bacteria (155). The uncanny similarity between RNP-granule biogenesis via LLPS and polyphosphate granule assembly merits further exploration.

**Proline cis-trans isomerization**

A key feature of several phase-separating RBPs is the presence of a PrLD (84), which can often contain sporadic proline residues. Given that proline introduces kinks as a result of its constrained side-chain geometry and given that PrLDs can aggregate into amyloid fibrils, prolines may serve as a natural fluidizer in PrLDs and other low-complexity domains to prevent aberrant aggregation. Proline isomerization may then serve a possible role in regulating phase transitions mediated by low-complexity domains. Indeed, peptidyl-prolyl cis-trans isomerases (PPIases) colocalize with stress granules, bind hydrogels formed from the PrLDs of RBPs, and increase the solvent accessibility of certain residues in hnRNPA2 as it assembles into fibrils (129). Importantly, PPIAs can also function as protein disaggregases with activity against amyloid fibrils (87,156). Thus, proline cis-trans isomerization may be another mechanism by which cells modulate the phase behavior of proteins.

**Aberrant phase transitions in neurodegenerative disease**

A hallmark of several neurodegenerative diseases is aberrant protein aggregation: α-synuclein aggregates in Parkinson’s Disease, amyloid-β and tau in in Alzheimer’s disease, and TDP-43 and the FET family of proteins in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) (35). For RBPs implicated in ALS and FTD, LLPS provides a mechanistic link between normal cellular function and disease phenotypes. FUS, TDP-43, hnRNPA1, and TIA-1 are among the RBPs that are associated with ALS and FTD, which coalesce into membraneless organelles called stress granules (83). Observations of purified FUS, hnRNPA1, and TIA-1 uncovered that these proteins form dynamic liquid droplets in vitro that age over time to become more static, fibrillar aggregates (11,101,157). The conversion from a liquid state to a more aggregated state has been termed an aberrant phase transition (101). The final aggregated form of the protein bears resemblance to the protein aggregates found in patients with ALS and FTD. Fibrillization can occur within the condensed liquid state, suggesting that concentrating these RBPs in membraneless organelles via LLPS as part of normal cellular biology may have the inadvertent effect of triggering protein aggregation over time (11,83,101). Indeed, data that aggregates of these RBPs are immunoreactive for other components of stress granules has provided further evidence that stress granules may be the sites of disease biogenesis (83,158). However, RBPs with PrLDs that are connected to neurodegenerative disease like FUS, TDP-43, TAF15, EWSR1, and hnRNPA1 are intrinsically aggregation prone (81,109,159-161). Thus, pathological aggregation could also be nucleated outside of stress granules. Pathological aggregates could then subsequently sequester specific stress-granule proteins.

Additional evidence connecting aberrant phase transitions to disease comes from analysis of mutations in these RBPs that are associated with hereditary forms of neurodegenerative disease. Disease-associated mutations often exacerbate protein aggregation and alter the
phase behavior of the protein (11,101). For example, ALS and multisystem proteinopathy (MSP) associated mutations in the PrLD of hnRNPA1 and hnRNPA2 increase the amyloidogenicity of these proteins and accelerate fibrillation (81). Additionally, ALS-linked mutations in TDP-43 also promote aggregation and alter TDP-43 phase behavior (118,152,159). The PrLDs of these proteins normally form weak, transient interactions with each other in the liquid droplets. Some disease-associated mutations strengthen the otherwise transient interactions in the PrLD, leading to less dynamic droplets and RNP granules (11,101,157,162). Likewise, the arginine-rich, dipeptide-repeat proteins, poly-PR and poly-GR, produced by repeat-associated non-ATG translation of the ALS/FTD-causing G4C2 repeat expansion of C9orf72 also accelerate aberrant phase transitions of RBPs with PrLDs and perturb the phases of several membraneless organelles (132,163-165). The protein aggregates seen in disease likely represents an end stage phenotype after aberrant phase separation has overwhelmed the cellular machinery that ordinarily reverses these altered phases.

**Counteracting neurodegenerative diseases with knowledge of phase separation**

Neurodegenerative disease like ALS and FTD lack effective therapies. Recent advances in our understanding of how altered phase transitions contribute to these disorders reveal several potential avenues for therapeutics. These include: 1) enhancing the machinery already present inside cells to maintain RNP-granule dynamics; and 2) targeting the factors that recruit RBPs to RNP granules.

The cell has various molecular chaperones that remodel misfolded proteins and contribute to proper maintenance of RNP-granule dynamics (166). Nuclear-import receptors also act as chaperones and dissolves that reverse LLPS and aberrant phase separation of their RBP cargo (106,110,113,119). Small-molecule enhancers of these chaperones or *de novo* designed chaperone proteins with enhanced disaggregate activity thus present promising approaches for targeting neurodegenerative diseases (86,87,167-170).

Targeting the specific factors that recruit neurodegenerative disease-associated RBPs to RNP granules may also therapeutically tune the accumulation of these RBPs inside stress granules. For example, knockdown of Ataxin 2 reduces accumulation of TDP-43 in stress granules and is therapeutic in reducing TDP-43 toxicity in several ALS models (171,172). Additionally, molecular seeds like PAR that nucleate RNP granules can also be potential targets for therapies using anti-sense oligonucleotides or small-molecule inhibitors of specific PAR polymerases (101,152), or methods to upregulate specific PAR glycohydrolases (173). Finally, RNA acts both as a molecular seed in the cell as well as a safeguard against aberrant phase separation in the nucleus where RNA concentration is higher (174). Thus, expression or delivery of certain RNAs that are particularly effective at reducing aberrant protein phase separation may also be therapeutic. Overall, we anticipate that advances in our understanding the molecular language of phase separation will ultimately enhance efforts to combat neurodegenerative diseases.

**Acknowledgments**

We thank Lin Guo, Bede Portz, and Zach March for feedback on the manuscript, and Lin Guo for the images in Figure 1A. J.S. was supported by grants from the NIH (R01GM099836, R21NS090205, and R21NS102687), the MDA, the ALSA, Target ALS, the Life Extension Foundation, Sanofi, the Office of the Assistant Secretary of Defense for Health Affairs, through the Amyotrophic Lateral Sclerosis Research Program under Award No. W81XWH-17-1-0237, and the Robert Packard Center for ALS Research at Johns Hopkins.

**Conflicts of Interest**

The authors declare no conflicts of interest.

**References**

Phases Underlie Nucleolar Subcompartments. Cell 165, 1686-1697


69. Kim, Y., and Myong, S. (2016) RNA Remodeling Activity of DEAD Box Proteins
Tuned by Protein Concentration, RNA Length, and ATP. Mol Cell 63, 865-876


amphicola. Proc Natl Acad Sci U S A 114, E2440-E2449


Figure 1. LLPS phase separation in vitro and in vivo. (A) In a mixture of two types of molecules, LLPS leads to the formation of two phases akin to droplets of oil appearing from a mixture of oil and water. Proteins can undergo a similar phase separation. In this case, the RBP FUS (olive circles) undergoes LLPS upon cleavage of the MBP tag (cyan circles) and forms liquid droplets that are enriched in FUS compared to the surrounding medium. (B) LLPS underpins the biogenesis of a wide array of membraneless organelles within cells. Depicted here is a non-exhaustive list of these organelles.
Figure 2. Critical interactions that drive LLPS. The interactions important in LLPS include cation-π, π-π, electrostatic, and transient cross-β contacts. Proteins that undergo LLPS are enriched for low-complexity disordered regions and multivalent domains. Polymers of ions, such as RNA, may additionally act as scaffolds or molecular seeds for LLPS. The image for transient cross-β contacts and LARKS comes from Hughes et al. (99).
The molecular language of membraneless organelles
Edward Gomes and James Shorter

J. Biol. Chem.  published online July 25, 2018

Access the most updated version of this article at doi: 10.1074/jbc.TM118.001192

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts