

Getting RNA and Protein in Phase

Stephanie C. Weber¹ and Clifford P. Brangwynne^{1,*}

¹Princeton University, Department of Chemical and Biological Engineering, Princeton, NJ 08544, USA

*Correspondence: cbrangwy@princeton.edu

DOI 10.1016/j.cell.2012.05.022

Nonmembrane-bound organelles such as RNA granules behave like dynamic droplets, but the molecular details of their assembly are poorly understood. Several recent papers identify structural features that drive granule assembly, shedding light on how phase transitions functionally organize the cell and may lead to pathological protein aggregation.

Nonmembrane-Bound Intracellular Granules

Membrane-bound organelles, such as the endoplasmic reticulum and Golgi apparatus, are the classical units of intracellular organization. These structures divide the cell into functionally distinct compartments, ensuring that high concentrations of the right molecules localize in the right place at the right time. Over the last few decades, another important class of intracellular structures has emerged: organelles that are not bound by a membrane. Instead, these structures self-assemble from a cytoplasmic or nucleoplasmic pool of soluble components, forming a type of aggregate. However, unlike the irreversible protein aggregates seen in neurodegenerative diseases such as Alzheimer's, these physiological assemblies remain highly dynamic, constantly exchanging subunits with a soluble pool. The rules governing the assembly of nonmembrane-bound organelles have been difficult to discern due to their complex composition, typically consisting of dozens of different molecules.

Nonmembrane-bound organelles often contain both protein and RNA and are variously called ribonucleoprotein (RNP) bodies or RNA granules. A wide variety of RNA granules have been described, including processing bodies, neuronal granules and germ granules in the cytoplasm (Anderson and Kedersha, 2006) and Cajal bodies, nucleoli, and PML bodies in the nucleus (Mao et al., 2011). These granules play a role in many processes involving RNA metabolism, including storage, splicing, decapping, and degradation. In addition to RNA granules, there are protein-only granules, such as the purinosome, a multienzyme body that can facilitate cellular purine biosynthesis (An et al., 2008).

Intracellular Granules as Liquid Droplets

Recent observations of germ (P) granules in *C. elegans* (Brangwynne et al., 2009) have suggested a mechanism for how nonmembrane-bound organelles assemble molecular components into coherent structures while simultaneously facilitating dynamic molecular reactions within. In the early embryo, P granules asymmetrically localize to the posterior, where they are exclusively inherited by the progenitor germ cell that forms upon cell division. This process was found to rely on a spatiotemporally controlled transition from a soluble phase, in which RNA and protein components are dispersed throughout the cytoplasm, to a condensed phase, in which these components are concentrated in the P granule. The condensed P granule exhibits characteristic liquid droplet behavior. For example,

induced cytoplasmic flows cause individual droplets to drip off of the nuclear envelope, and two small droplets fuse into a larger droplet upon contact, like raindrops on a car windshield. P granule localization thus appears to be governed by a classic liquid phase transition. The nucleolus behaves similarly (Brangwynne et al., 2011), and liquid phase transitions have thus been suggested to be a general biophysical mechanism underlying granule assembly.

Phase Transitions in Cell Biology

Phase transitions are ubiquitous in nature. The most familiar examples are warm water vapor condensing on a cool bathroom mirror or lake water freezing into a sheet of ice. Gaseous molecules in water vapor rarely interact with one another. When they condense into liquid water, the molecules form transient hydrogen bonds that are constantly reshuffled by thermal fluctuations. Upon freezing, the water molecules crystallize into an ordered lattice with stable hydrogen bonds holding neighboring molecules firmly in place. In this example, temperature serves as the main control parameter determining the phase of the system—gas, liquid, or solid (Figure 1A). However, it is the molecular properties of water—particularly the hydrogen bonding enabled by the dipole moment between electronegative oxygen and electropositive hydrogen—that define the rules of assembly of its condensed phases.

Like water, macromolecules also undergo phase transitions. For example, soluble proteins can condense into crystalline solids, commonly used for X-ray crystallography. Solutions of purified protein can also condense into droplet phases (Vekilov, 2010). In vivo, such liquid phase separation may be important for structuring the lipid membrane. The most obvious intracellular phase transitions involve cytoskeletal proteins. Actin and tubulin can rapidly transition between a soluble “gas-like” state in which monomers rarely interact with each other and a crystalline solid state, in which adjacent monomers stably interact within filaments. Unlike water vapor in which molecules are spaced far apart, soluble monomers exist in a crowded cytoplasm where the frequent contact with other types of molecules can also influence their tendency to phase separate. A key difference from phase transitions in nonbiological matter is that these cytoskeletal phase transitions involve nucleotide hydrolysis, one of numerous nonequilibrium processes that enable biological control of filament assembly.

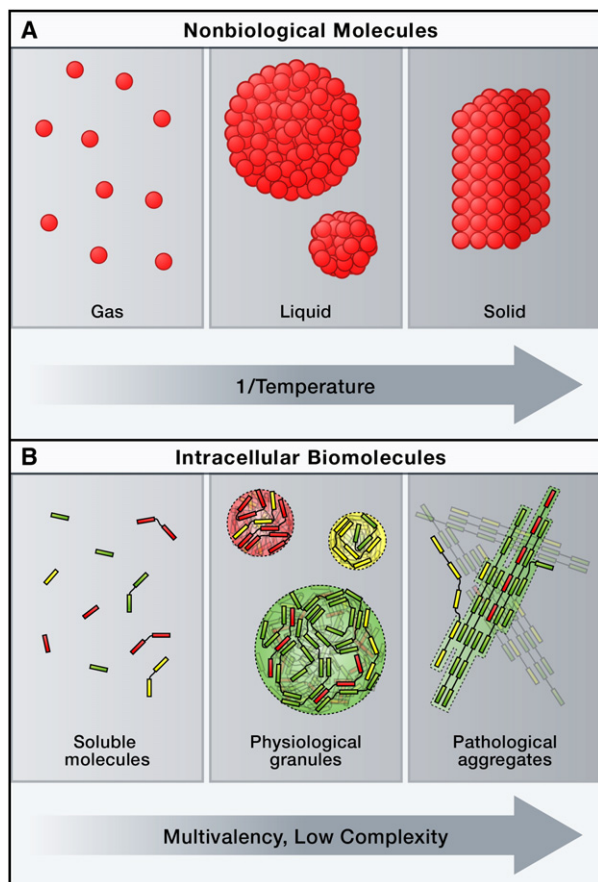


Figure 1. Intracellular Biomolecules Undergo Phase Transitions

(A) Nonbiological molecules can exist in different phases: gas, with few intermolecular interactions; liquid, with many transient intermolecular interactions; and solid, with stable bonds. Transitions between these phases are driven by external parameters such as temperature.

(B) The nucleus and cytoplasm appear to be organized by similar phase transitions. Multivalent interaction domains and low-complexity sequences drive soluble RNA and protein molecules to transiently associate with each other, forming liquid-like granules in the cell. ATP-dependent biological activity could enable the cell to dynamically control the assembly and disassembly of these phases. Intriguingly, structural characterization suggests that the same intermolecular interactions underlying granule assembly can lead to the formation of stable amyloid aggregates seen in diseases such as Alzheimer's.

The Importance of Being Repetitive

If soluble RNA and protein molecules undergo a phase transition to form liquid droplets, what are the molecular properties that define this transition? RNA-binding domains typically bind short stretches of their cognate RNAs with relatively weak affinity yet occur within proteins in multiple copies that serve to increase the total binding affinity (Lunde et al., 2007). The assembly of a dynamic liquid phase could be driven by multiple weak interactions that are strong enough to bring molecules together but not so strong that they arrest the dynamics within.

A recent paper by Rosen and colleagues (Li et al., 2012) demonstrates that weak multivalent binding is capable of assembling liquid phase droplets. Their study focuses on NCK and NWASP, which, together with nephrin, modulate the activity of the actin-nucleating Arp2/3 complex. This interaction is

mediated by three SRC homology 3 (SH3) repeats in NCK that bind to the proline-rich motif (PRM) ligand in NWASP.

Li et al. found that some solutions of purified repeats of the SH3 domain (SH3_n, in which n = 1–5) and its PRM ligand (PRM_n) became opalescent at high concentration. PRM_n + SH3_n precipitated from solution into highly dynamic droplets, concentrating ~100-fold. As expected for liquids, on contact, two droplets fused into a single larger spherical droplet. The transition between the dilute, soluble phase and the condensed droplet phase was sensitively dependent on the number of domain repeats (n). PRM₃ + SH3₃ never assembled into droplets, even at the highest concentration tested. However, higher valency constructs (n = 5) readily formed liquid droplets at concentrations comparable to intracellular protein concentrations (1–10 μM). Similar droplets were observed for other multimerized binding pairs, including the RNA-binding protein PTB and an RNA oligonucleotide, demonstrating that the tendency for multivalent binding domains to phase separate into liquid-like droplets is a general phenomenon, at least in vitro.

Li et al. further demonstrated that multivalent binding interactions can indeed drive structural assembly in living cells. When GFP-tagged PRM₅ and mCherry-tagged SH3₅ were coexpressed in HeLa cells, they colocalized in spherical puncta. As with many endogenous intracellular granules, photobleaching experiments revealed that these synthetic droplets were highly dynamic, turning over PRM₅ and SH3₅ components in less than 20 s. The presence of repeated interaction domains is widely seen in the proteome, particularly among RNA-binding proteins, suggesting that multivalency may be a ubiquitous driving force for droplet condensation.

One potential function of nonmembrane-bound organelles appears to be increasing the local concentration of reactants, thereby accelerating rate-limiting steps in catalytic or assembly processes. For example, Cajal bodies are essential in zebrafish embryogenesis, when RNA metabolism must be particularly rapid (Strzelecka et al., 2010). To determine whether their droplets could function as reaction crucibles, Li et al. monitored in vitro actin assembly facilitated by an N-WASP mutant that can promote the phase transition into droplets but cannot directly stimulate actin assembly. Upon increasing the protein concentration beyond the phase transition boundary, the rate of actin assembly increased ~3-fold. Intriguingly, the work of Li et al. highlights the possibility that one kind of phase transition—in this case, PRM-SH3-mediated droplet condensation—can facilitate another kind: the crystallization of actin monomers into filaments.

Islands of Low Complexity in a Complex Sea

Li et al. identify multivalency as an important molecular feature underlying phase separation. But it is unlikely that this is the whole story. Indeed, the molecular composition of *C. elegans* P granules suggests another important property. GLH DEAD box RNA helicases are constitutive P granule components that contain repeats of the hydrophobic amino acid phenylalanine (F) adjacent to the flexible amino acid glycine (G), followed by a short stretch of relatively hydrophilic residues. These FG repeats, which are natively unfolded, serve as self-association domains in many nuclear pore proteins. When expressed ectopically in *C. elegans*, a single GLH-1 FG repeat domain remains

soluble, diffusely localized throughout the cytoplasm (Updike et al., 2011). However, three tandem FG repeat domains readily form granules, demonstrating that disordered domains are also capable of assembling droplets when multimerized.

Two recent papers from McKnight and colleagues provide strong evidence that such disordered domains play an important role in intracellular phase transitions more generally (Han et al., 2012; Kato et al., 2012). Their discovery hinged on a small molecule, 5-aryl-isoxazole-3-carboxamide. Within minutes of adding a biotinylated form of the isoxazole (b-isox) to cell lysates, the authors observed a white precipitate. Mass spectrometry analysis indicated that the precipitate was highly enriched for RNA-binding proteins (RBPs), specifically those associated with a wide variety of RNA granules. For example, fused in sarcoma (FUS), an abundant RBP found in neuronal granules was precipitated from all cell types studied.

A striking feature of the b-isox-precipitated proteins was the prevalence of low-complexity sequences (LCSs), which have low amino acid diversity and are naturally disordered. These LCSs contained short repetitive motifs that include large hydrophobic amino acids, such as the 27 repeats of the tripeptide [G/S]Y[G/S] found in FUS. Using a series of recombinant GFP fusion proteins, the authors demonstrate that the LCSs of seven candidate proteins, including FUS, are necessary and sufficient to partition into the b-isox precipitate. Furthermore, these sequences are capable of forming hydrogels, reminiscent of those assembled by FG-repeat-containing nuclear pore components (Frey et al., 2006).

The condensed gel phases formed by LCSs *in vitro* thus appear structurally distinct from the more liquid-like droplets described by Li et al. Indeed, electron micrographs of LCS hydrogels reveal well-defined filaments. Intriguingly, their X-ray diffraction patterns are characteristic of cross- β structures, a hallmark of amyloid fibrils found in Alzheimer's and prion diseases. An extended network of hydrogen bonds makes such pathological amyloid fibrils extremely stable and long-lived. However, despite their structural signatures, the filaments observed by Kato et al. are relatively unstable. Unlike fibers of the yeast prion Sup35, which resist SDS solubilization, FUS fibers readily depolymerize upon exposure to even very low concentrations of SDS. Interestingly, hydrogels assembled from the FG repeats of nuclear pore components were recently shown to contain amyloid-like interactions, possibly contributing to a permeability barrier without causing aggregation (Ader et al., 2010). Cross- β contacts are thus present in both static pathological fibrils and dynamic physiological assemblies.

Kato et al. propose that polymerization of LCSs into amyloid-like fibers is the molecular principle driving RNA granule assembly. However, the assembly of these fibers requires extreme, nonphysiological conditions, and it remains unclear whether such fibers normally assemble *in vivo*. Indeed, hydrogels were formed by incubating high concentrations (>100 μ M) of purified protein at low temperature for as long as 1 week. The resulting gels hold the shape of their tube, reflecting very slow molecular relaxation times (hours to days). It is difficult to reconcile these timescales with the more liquid-like behavior *in vivo* and the fast molecular exchange (seconds to minutes) observed between granules and the nucleocytoplasm. More-

over, although filaments are clearly visible in pure LCS hydrogels, light microscopic images of the b-isox precipitate from cell lysates show an amorphous aggregate (no high-resolution EM data are shown). Finally, b-isox precipitates solubilize upon warming to 37°C, suggesting that the cross- β interactions likely driving this precipitation become weak and transient at physiological conditions. This may explain why similar amyloid-like fibers have not been observed in EM images of intracellular RNA granules (Gall et al., 1999).

Given these questions regarding *in vivo* fiber formation, we propose an alternative interpretation (Figure 1B). Just as a network of hydrogen bonds gives rise to solid ice crystals at low temperatures, cross- β contacts can induce polymerization of amyloid-like filaments *in vitro*, but these fibers likely represent an extreme state not normally present in the cell. Instead, under physiological conditions, transient interactions between LCSs could drive the assembly of dynamic droplets as transient hydrogen bonds give rise to the cohesive properties of liquid water. Amyloid pathogenesis may therefore represent an aberrant phase transition from a liquid droplet into a solid fibril. Consistent with this hypothesis, a disordered fluid phase has been proposed to function as a metastable intermediate promoting *de novo* assembly of amyloid fibers (Serio et al., 2000). The particularly slow dynamics reported by Kato et al. for *in vitro* FUS hydrogels could reflect the close proximity of FUS to the phase boundary between physiological and pathological states. Indeed, cytoplasmic aggregations of mutant FUS have been observed in neural tissue from familial ALS patients (Kwiatkowski et al., 2009).

RNA: Driving Assembly or along for the Ride?

In addition to LCS-enriched proteins, thousands of mRNAs coprecipitate with b-isox. Han et al. propose an appealing two-component model for mRNA recruitment to intracellular granules: RBPs bind to mRNA via an RNA-binding domain (RBD) and phase separate by the interaction of their LCSs. Indeed, an RNA hairpin is retained in a FUS hydrogel only when incubated with a chimeric protein containing both the LCSs of FUS and an RBD, but not either domain separately. Though this experiment provides an engineered proof of principle, it remains to be seen whether two domains are required for targeting RNA species *in vivo*.

Given the findings of Kato et al. and Han et al., a study in this issue of *Cell* may provide further insight into the role of RNA in phase transitions. Castello et al. use UV crosslinking to identify proteins that directly bind to polyadenylated mRNAs in proliferating HeLa cells (Castello et al., 2012). They take a comprehensive, unbiased approach to define the human "mRNA interactome," identifying nearly 500 proteins with novel putative RNA-binding activity. Detailed analyses reveal sequence and structural properties of RBPs that are similar to those identified by Kato and colleagues in b-isox precipitates, including an enrichment of LCSs. Interestingly, tyrosine (Y) is overrepresented in the interactome compared with the proteome as a whole and is often flanked by G/S residues. Because Kato et al. identified the [G/S]Y[G/S] motif as important for granule assembly, this observation suggests that many RBPs may be capable of promoting such phase transitions.

It is not yet clear whether these sequence features are found in the RBD itself or whether they comprise a distinct domain. The latter case would support Han's two-component model for recruitment of mRNAs to condensed phases. However, the former suggests a bifunctional model, in which LCSs bind mRNA and simultaneously promote phase separation. Indeed, the RGG box, a classic RBD that is often interspersed with conserved hydrophobic residues, including the GYG motif (Burd and Dreyfuss, 1994), could be a bifunctional domain. In addition, Castello and colleagues also note that RNA-binding motifs occur repetitively, which is consistent with Li et al.'s finding that a multivalent RBP and RNA oligo can phase separate into liquid-like droplets. Interestingly, Dundr and colleagues showed that tandemly tethered RNA transcripts can nucleate assembly of nuclear bodies in vivo (Shevtsov and Dundr, 2011), demonstrating that RNA can indeed drive granule assembly in cells.

Tuning Assembly for Biological Function

Many molecularly distinct granules coexist within a single cell—sometimes even in contact with one another (Gall et al., 1999)—and yet they remain distinct. If LCSs provide a common mechanism for phase separation, as Han, Kato, and colleagues suggest, then how are functionally distinct granules assembled and separately maintained?

Insight into this question could come from the differential retention of proteins in FUS hydrogels. GFP can permeate a preformed FUS hydrogel; however, it rapidly diffuses out following buffer exchange. By contrast, a GFP-FUS fusion protein remains associated for at least 1 hr after washing, indicating a strong homotypic interaction. Intermediate retention times were observed for heterotypic binding by other LCS proteins. This may suggest a potential mechanism for molecular sorting: the energetics of homotypic and heterotypic LCS interactions could promote phase separation into different granules of distinct composition, coexisting within the cytoplasm (Figure 1B).

Finally, how does a cell spatially and temporally regulate granule assembly and composition? Both Li et al. and Han et al. demonstrate that the phosphorylation state of a protein domain can change its ability to assemble into condensed phases. In the case of nephrin, tyrosine phosphorylation favors condensation, shifting the phase boundary to lower concentrations. Conversely, phosphorylation of serine residues in the LCSs of FUS decreases its retention in FUS hydrogels in a dose-dependent manner. Thus, depending on the binding interactions, phosphorylation—or other posttranslational modifications—could be used to shift the equilibrium between the soluble and condensed phases.

Outlook

These recent papers highlight important molecular features—multivalency and low complexity—that appear capable of driving phase transitions in the cell. Together, these results elucidate a mechanistic framework for understanding the assembly and biophysical nature of the dynamic RNA/protein droplets widely observed in vivo. Multimeric domains or stretches of repetitive motifs determine the “stickiness” of weak intermolecular interactions by controlling their number and effective lifetime. The degree of multimerization or low-complexity content thus gives

rise to a spectrum of intermolecular affinities, which result in different phases: gas-like soluble molecules, liquid-like droplets, and solid-like pathological aggregates (Figure 1B).

Posttranslational modifications and other ATP-dependent processes serve as biological parameters controlling the assembly of different intracellular phases, similar to the way in which temperature controls nonbiological phase transitions (Figure 1). For example, the activity of ATP-dependent RNA helicases—which are prevalent in RNA granules and often contain LCSs—could regulate the degree of molecular interactions that drive phase separation. Such biological activity could also be used to tune the timescales of molecular dynamics within RNA droplets, perhaps determining whether they function in RNA storage (high viscosity, slower dynamics) or processing (low viscosity, faster dynamics) (Brangwynne et al., 2011).

These studies provide a molecular foundation for intracellular phase transitions as a unifying principle underlying the spatio-temporal organization of the cell. An exciting future challenge will be elucidating the precise structural and energetic rules that give rise to functionally distinct granules and determining how these physicochemical driving forces can inadvertently lead to pathological aggregation.

REFERENCES

- Ader, C., Frey, S., Maas, W., Schmidt, H.B., Görlich, D., and Baldus, M. (2010). *Proc. Natl. Acad. Sci. USA* 107, 6281–6285.
- An, S., Kumar, R., Sheets, E.D., and Benkovic, S.J. (2008). *Science* 320, 103–106.
- Anderson, P., and Kedersha, N. (2006). *J. Cell Biol.* 172, 803–808.
- Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoegge, C., Gharakhani, J., Jülicher, F., and Hyman, A.A. (2009). *Science* 324, 1729–1732.
- Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). *Proc. Natl. Acad. Sci. USA* 108, 4334–4339.
- Burd, C.G., and Dreyfuss, G. (1994). *Science* 265, 615–621.
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Stein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, et al. (2012). *Cell* 149, this issue, 1393–1406.
- Frey, S., Richter, R.P., and Görlich, D. (2006). *Science* 314, 815–817.
- Gall, J.G., Bellini, M., Wu, Z., and Murphy, C. (1999). *Mol. Biol. Cell* 10, 4385–4402.
- Han, T.W., Kato, M., Xie, S., Wu, L.C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G., and McKnight, S.L. (2012). *Cell* 149, 768–779.
- Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). *Cell* 149, 753–767.
- Kwiatkowski, T.J., Jr., Bosco, D.A., Leclerc, A.L., Tamrazian, E., Vanderburg, C.R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E.J., Munsat, T., et al. (2009). *Science* 323, 1205–1208.
- Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F., et al. (2012). *Nature* 483, 336–340.
- Lunde, B.M., Moore, C., and Varani, G. (2007). *Nat. Rev. Mol. Cell Biol.* 8, 479–490.
- Mao, Y.S., Zhang, B., and Spector, D.L. (2011). *Trends Genet.* 27, 295–306.
- Serio, T.R., Cashikar, A.G., Kowal, A.S., Sawicki, G.J., Moslehi, J.J., Serpell, L., Arnsdorf, M.F., and Lindquist, S.L. (2000). *Science* 289, 1317–1321.
- Shevtsov, S.P., and Dundr, M. (2011). *Nat. Cell Biol.* 13, 167–173.
- Strzelecka, M., Trowitzsch, S., Weber, G., Lüthmann, R., Oates, A.C., and Neugebauer, K.M. (2010). *Nat. Struct. Mol. Biol.* 17, 403–409.
- Updike, D.L., Hachey, S.J., Kreher, J., and Strome, S. (2011). *J. Cell Biol.* 192, 939–948.
- Vekilov, P.G. (2010). *Soft Matter* 6, 5254–5272.