HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2023 November 06.

Published in final edited form as:

Mol Cell. 2023 April 06; 83(7): 1016–1021. doi:10.1016/j.molcel.2023.03.014.

Transitions in the framework of condensate biology

Geraldine Seydoux,
Johns Hopkins University, USA

Mingjie Zhang,
Southern University of Science and Technology, China

Julie D. Forman-Kay,
The Hospital for Sick Children, Toronto, Canada and University of Toronto, Canada

Brian McStay,
University of Galway, Ireland

Kathy Fange Liu,
University of Pennsylvania, USA

Pilong Li
Tsinghua University, China

Abstract

As phase separation is found in an increasing variety of biological contexts, additional challenges have arisen in understanding the underlying principles of condensate formation and function. We spoke with researchers across disciplines about their views on the ever-changing landscape of biomolecular condensates.

Functional compartments or incidental condensates?

The observation (https://doi.org/10.1126/science.1172046) that P granules in *C. elegans* behave like liquid droplets dissolving and growing in a sea of cytoplasm captured the imaginations of many and provided a strong theoretical framework for the assembly of cellular compartments not limited by membranes. The field of “biomolecular condensates” was initially concerned with whether phase separation could account for the many granules, bodies, foci, speckles, and puncta described by cell biologists over the last century. We now understand that not all assemblies visible by light microscopy arise by phase separation, but this preoccupation (is my granule a condensate?) and the confusion over terminology (what is a condensate?) have overshadowed an even greater question looming over the field: what is the function of condensates?

First, some definitions. Although the term “condensate” has been used broadly, I prefer a narrower definition for “condensate” as a phase-separated assembly. Phase separation

DECLARATION OF INTERESTS

Geraldine Seydoux is a member of the scientific advisory board of Dewpoint Therapeutics, Inc. Pilong Li is a member of the advisory board of *Molecular Cell*.

describes the de-mixing of interacting molecules in solution. No energy input is required—when a sufficiently high concentration (c_{sat}) is reached, the energy from favorable binding interactions exceeds the entropic cost of de-mixing and, like a switch, the molecules redistribute between two phases.

The assumption has been that phase separation creates “organelles”: by concentrating molecules, new biochemistry is created in the condensates. Whether this assumption always holds in the cellular environment, however, is not clear. Because phase separation is greatly stimulated by multivalency (https://doi.org/10.1038/nature10879), large complexes—for example, RNA molecules bound by multiple proteins—are particularly prone to condensation. As the concentration of complexes rises in response to cellular needs, they may saturate and form condensates. If c_{sat} is exceeded only by a small margin, the proportion of complexes that join the condensates may be minimal. The condensates will appear bright by microscopy, but the majority of active complexes will still be in the dilute phase. “Incidental condensates” provide no new function and could have evolved to be tolerated by cells as inevitable byproducts of cellular activity.

The hope, of course, is that evolution has taken advantage of condensation to augment what can be achieved by soluble complexes. Clearly that is the case for P granules: their liquid properties help polarize (https://doi.org/10.1126/science.abg7071) mRNAs and proteins in embryos. Distinguishing between the contributions of the condensed and the dilute phases, however, can be challenging. Mutations that disrupt binding surfaces will block condensate formation but also run the risk of breaking up soluble complexes. Mutating intrinsically disordered regions (IDRs) has become a shortcut assumed to only affect phase separation, but that assumption needs to be validated since IDRs can also participate in sequence-specific interactions that build soluble complexes. A big challenge ahead is to learn how to recognize sequences that have been selected by evolution *specifically* to enhance phase separation.

From basic principles to specificity

The field of phase separation research is vibrant but still very young. There are many pressing issues, and I raise two here.

Basic principles of condensate formation:

The overarching rule for forming a biological condensate is that molecules interact with each other in a multivalent fashion, leading to formation of interconnected molecular networks that spontaneously phase separate. However, we do not have a theoretical framework to accurately describe the exact parameters that determine which molecules or molecular systems can phase separate, how these molecules phase separate, or how phase separation of such molecular systems is regulated, etc. Currently, the field has been borrowing existing theoretical frameworks from classical dilution solution chemistry and that of the polymer chemistry to qualitatively or semi-quantitatively describe biological phase separation. A biological condensate is a soft matter, so theories for dilution solutions are not sufficient. Biopolymers are overall more heterogeneous than chemical polymers. Additionally, cellular biological condensates are invariably composed of many different

types of molecules with very different abundances. Thus, existing phase separation theories derived from polymer science are not adequate for biological condensates. Cross-field collaborations are needed to develop theoretical frameworks for biological condensates. On one hand, we need to attract more experts from polymer chemistry, soft matter physics, physical chemistry, etc. to continue developing theories underlying biological condensate formation. On the other hand, new methods need to be developed to study grammars governing molecules in the formation of condensates and to investigate properties of molecules in condensates (e.g., motion behavior, network properties, and conformational heterogeneity). Combining insights gained from theoretical studies and deeper experimental characterization of biological condensates holds huge promise for guiding future phase separation research.

Specificity, specificity, and specificity:

The hallmarks of biological processes are diversity and specificity. Numerous different biological condensates with incredibly diverse but specific functions have been identified (or suggested). This immediately raises the question of how the molecular components in each biological condensate are determined. At present, the prevailing concept in the field is that formation of biological condensates is driven by intrinsically disordered sequences in proteins (or IDPs for short). There is no doubt that IDPs are critical for numerous biological condensates. However, IDPs interact with each other with very weak affinities, at least when IDPs are in the liquid phase. Such low-affinity interactions suggest that the formation of productive molecular networks at physiologically relevant concentrations is low. Thus, the specific condensate assembly is also low given that the governing principle for forming biological condensates is multivalent interaction-mediated large molecular network formation. Numerous proteins contain more than one specific and strong (with respect to IDP-based interactions) protein interaction domain or motif, and many proteins can further multimerize. Thus, many unique biological condensates may form via specific and multivalent protein-protein (or protein-nucleic acid) interactions. Such specific biological condensates (e.g., the postsynaptic density condensate in neuronal synapses; https://doi.org/10.1016/j.cell.2018.06.047) are highly stable and yet can be specifically regulated by various biological signals. It is of note that many of these multi-domain/motif proteins also contain disordered sequences, which are ideal to further promote phase separation of the molecular networks assembled by specific interactions. Thus, a fertile direction for future research is to delineate how strong and specific molecular interactions and IDP-based interactions act together to form functional biological condensates.

Protein condensate physics underlying biology

There has been an explosion of interest in the role of biomolecular condensates for organizing and regulating biological functions, including generating biomaterials. It is useful to consider that biology exploits whatever physics/physical chemistry it has available to enable fine-tuning of cellular organization and regulation of complex functions. Obvious examples are the phase separation and phase transitions of lipids to form and remodel biological membranes. More recently, we have found that lipids are not unusual and that proteins, nucleic acids, and other macromolecules can also undergo phase separation. The

physics underlying structural stabilization and dynamic exchange in folded proteins and nucleic acids to provide a mechanistic basis for function has been well-recognized. But only recently has that same physics been seen to underlie multivalent dynamically exchanging interactions of IDRs and less-stable nucleic acid structures in the context of their functional roles, including driving phase separation.

Indeed, bioinformatics studies show that about a quarter of residues in eukaryotic proteomes are in IDRs that do not have positional conservation diagnostic of adopting a folded structure, but, as reported for yeast IDRs, do have significant conservation of bulk sequence features, many of which are associated with phase separation (https://doi.org/10.7554/eLife.46883). IDR phase-separation predictive algorithms suggest that about a third of proteins in the human proteome contain IDRs that can drive phase separation. Besides these, many folded protein domains contribute multivalent interactions that drive phase separation, indicating a large role for phase separation in biology.

Computational approaches are making a big impact in understanding the physical interactions driving IDR phase separation, including bioinformatics, theory, coarse-grain models, and all-atom simulations. Experimental studies by nuclear magnetic resonance (NMR) spectroscopy are also providing molecular details of the dynamic interactions within select condensed protein phases. Together, these point to the importance of aromatics, other pi-containing groups, and charges, with many other interactions involved, including hydrogen bonds and other backbone interactions, hydrophobic groups, and transient kinked beta and helical structures (i.e., essentially all physical interactions contribute).

The physical interactions within condensed phases generate unique material properties and solvent environments, which are tightly associated with their biological functions. The chemical groups of componentIDRs contribute significantly to condensate solvent environments given their high concentrations. Such solvents can impact folded protein and nucleic acid structure. NMR is increasingly being used to characterize how condensate environments affect structural changes of folded proteins, domain organization, and dynamic regulatory elements within folded proteins, helping us understand how these unique solvents regulate function. Similar to the way that structural studies of isolated proteins provide key insights into molecular bases of function, future computational and NMR studies of simplified *in vitro* model systems will help elucidate mechanisms by which condensates function in biology, complementing cellular imaging data with increased resolution.

At the nucleolus: A driver or a consequence?

The phase separation (PS) model for subcellular organization proposes that multiple low-affinity non-specific interactions—often involving intrinsically disordered domains of constituent proteins—act as drivers for biogenesis of membraneless compartments. A general criticism of PS is that it downplays active processes, focusing instead on the behavior of carefully selected constituent proteins (https://doi.org/10.15252/embj.2021109952). This is most evident in relation to nucleoli, the sites of ribosome biogenesis.

It is now common to see, among the opening sentences of relevant publications, statements such as “nucleoli form through liquid-liquid phase separation.” Higher eukaryotic nucleoli are composed of three recognisable components that reflect the stages of ribosome biogenesis. Fibrillar centers (FCs) comprise ribosomal genes and transcription machinery. Transcription takes place at the interface between FCs and surrounding dense fibrillar components (DFCs). Modifications of nascent pre-rRNA and early pre-rRNA processing occur in the DFC. Late processing and subunit assembly take place in the granular component (GC). Based primarily on *in vitro* observations with cherry-picked components, PS advocates describe FCs, DFCs, and the GC as distinct liquid phases. Such statements fail to appreciate that nucleoli are the most metabolically active domains in the nucleus, forming around nucleolar organiser regions (NORs), the most heavily transcribed regions of the genome. Ribosome biogenesis requires many hundreds of proteins and ribonucleoprotein complexes and involves a diverse array of highly specific interactions. Nucleolar proteins typically exist in either latent or operative states, with PS studies likely reporting on a nucleolar localized but latent state of the target protein (https://doi.org/10.1101/gad.349748.122). It seems more likely to me that phase-separated latent nucleolar proteins overlay a ribosome biogenesis production line, scaffolded by rDNA and pre-rRNAs. If so, PS is a *consequence*, rather than a *driver*, of nucleolar formation.

A single human nucleolus can involve as many as ten chromosomally distinct NORs. Such nucleoli are typically highly irregular in shape and not spherical as predicted by pure liquid-like behavior. Examination of their internal organization reveals that individual NORs and their derived products occupy distinct subnucleolar territories (https://doi.org/10.1101/gad.348234.121). These combined observations are most compatible with a model whereby PS is subservient to scaffolded ribosome biogenesis. The removal of scaffolding by transcriptional inhibition results in altered nucleolar morphology seemingly dominated by PS, highlighting the role of scaffolding in constraining PS.

The condensates underlying sexual dimorphism

Liquid-liquid phase separation (LLPS) underlies the formation of membrane-less compartments in cells. These compartments concentrate on a specific collection of proteins and nucleic acids and subsequently influence a myriad of biological events including but not limited to epigenetic regulation, transcription, and RNA metabolism. Dysregulation of these condensates has been observed in a wide array of disorders, from developmental defects to age-related diseases. Emerging evidence shows that these biomolecular condensates may be regulated in a sex-specific manner, may have different contents depending on the chromosomal sex of the cell, and in turn, may influence cellular functions and disease pathways in a sex-biased manner. These potential sex differences likely arise from sex-chromosome-encoded homologous proteins. Although largely nonhomologous, the human sex chromosomes contain seventeen pairs of homologous proteins, in which the X-chromosome-encoded homolog is between 85% and 98% identical to the Y-chromosome-encoded homolog. Despite their homology, certain X-Y protein pairs have markedly different phase separation propensities and enzymatic activities. For instance, the Y-chromosome-encoded DEAD-box RNA helicase DDX3Y is significantly more prone to LLPS and is a significantly slower ATPase than its X-chromosome-encoded homolog

DDX3X. This is due to differences in their amino acid sequences, which are primarily clustered at their N-terminal IDRs. These distinct LLPS propensities further contribute to the different translation repression strengths and RNA constituents of DDX3X-positive stress granules versus DDX3Y-positive stress granules. The differences between DDX3X and DDX3Y at the condensation and enzymatic functions exemplify a fundamental sexual dimorphism of functional phenotypes at the molecular level that derives directly from genetic differences between the X and Y chromosomes. Similarly, UTX and UTY, a pair of sex-chromosome-encoded homologous histone demethylases, are also distinct in their LLPS and enzymatic activities. UTY is significantly less catalytically active than UTX. This is largely due to a single amino acid substitution (Ile1267 in UTX and Pro1214 in UTY). Like DDX3Y, UTY is more prone to LLPS than UTX and forms less-dynamic condensates; these differences are driven by amino acid sequence differences in the core IDRs of both proteins. The stronger LLPS propensity and less-dynamic property of UTY condensates cause UTY to be a weaker tumor suppressor than UTX. DDX3X/DDX3Y and UTX/UTY are evolutionarily conserved and expressed at the protein level in both reproductive and non-reproductive tissues. Thus, the differences between the Y chromosome homologs and their X chromosome partners may introduce sex distinctions in a wide range of biological functions via their distinct LLPS-mediated and enzymatic activity-involved roles. Based on the evidence presented here, I propose the existence of sex-specific protein-RNA condensates that influence gene regulation at the epigenetic and RNA levels.

Condensate properties: Assembly, size, and future outlooks

The functions of many cellular condensates are likely multifaceted no matter how specific or how general their nomenclatures imply. Ultimately, condensation provides cells with an efficient way to manage their vast assortment of processes, including biochemical reactions and dynamic material storage. For any specific type of condensate, its various functions are inevitably coupled with its initial assembly, subsequent architectural changes, and eventual disassembly. Both condensate assembly and disassembly require triggering signals. The signals and the scaffolds of condensates are unlikely to be isotropic in cells. Hence cellular condensates, normal or abnormal, should not be randomly distributed in cells; instead, they should have biased localization determined by their triggering signals. For condensates with unknown localization, such as destruction condensates of β-catenin and stress granules, dissection of the location(s) of a type of condensate as well as the mechanism via which their localization is achieved are crucial for understanding their functions.

For functional specificity, condensates orchestrate a myriad of related clients directly or indirectly via dynamic networks formed by scaffold(s). The intricate interplay between these components is dictated by their relative stoichiometry and physiochemical properties, which unavoidably imposes a minimal condensate size. That said, there is a misconception that a minimal condensate size is required to qualify as LLPS. Analogously, we do not need to know the sizes of dewdrops to judge whether they form at night. Often, a wet atmosphere and a plummeting temperature suffice. By the same token, LLPS is a well-defined physical process and hence whether it is happening can be determined by intensive parameters including environmental conditions and scaffold concentrations but not by extensive parameters such as the sizes of the resulting condensates.

In recent years, our knowledge of the functions of LLPS and biomolecular condensates has expanded to various horizons in all kingdoms of life. In no particular order, these include (1) phase separation in environmental sensing and stress responses in plants and animals, (2) distinct condensate formation under physiological versus pathological conditions, (3) condensates derived from undruggable (but very valuable) therapeutic targets, (4) studies of cross-talk or coregulation of multiple condensates, and (5) communication between condensates with membrane-enclosed organelles. We are also drilling down deeper to better understand well-known condensates such as the nucleolus and stress granules.

Conceivably, all these known biological functions of condensates may continue receiving attention for years to come. In addition, the roles of phase separation in shaping the cellular landscape will likely be revealed in a more comprehensive way: more cellular molecules (including proteins), RNAs, DNAs, lipids, carbohydrates, or even metabolites will be found functioning in condensates. More phase-separated condensates will likely be found in all stages of life from fertilization and early development to aging. Meanwhile, insights acquired in the current stage of mechanistic exploration will likely benefit future empirical applications such as promoting agriculture production and combating diseases. Moreover, multivalency-driven LLPS should not only be a mechanism to explain the various (dys)functions of condensates in cells, but also a powerful tool to manipulate cellular activities. In time, multiple LLPS-enabled techniques may be invented for studying biology and beyond.

Recent technological advances have helped deepen our understanding of functional phase separation. These techniques include, but are not limited to, imaging such as cryogenic electron tomography and super-resolution fluorescence imaging, biochemical strategies such as microfluidic and sorting technologies, various biophysical methods to measure rheological properties of condensates, and physical approaches to manipulate condensates such as optical tweezers. Nonetheless, application of advanced technologies to study biomolecular condensates, especially those in cells or even *in vivo*, is at a very early stage. A lot of work will be required in the years to come.

Biographies

Geraldine Seydoux

Mingjie Zhang

Julie D. Forman-Kay

Brian McStay

Kathy Fange Liu

Pilong Li