



Learning the chemical grammar of biomolecular condensates

Henry R. Kilgore¹✉ and Richard A. Young^{1,2}✉

Biomolecular condensates compartmentalize and regulate assemblies of biomolecules engaged in vital physiological processes in cells. Specific proteins and nucleic acids engaged in shared functions occur in any one kind of condensate, suggesting that these compartments have distinct chemical specificities. Indeed, some small-molecule drugs concentrate in specific condensates due to chemical properties engendered by particular amino acids in the proteins in those condensates. Here we argue that the chemical properties that govern molecular interactions between a small molecule and biomolecules within a condensate can be ascertained for both the small molecule and the biomolecules. We propose that learning this 'chemical grammar', the rules describing the chemical features of small molecules that engender attraction or repulsion by the physicochemical environment of a specific condensate, should enable design of drugs with improved efficacy and reduced toxicity.

Biomolecular condensates are dynamic assemblies that typically are not enclosed by membranes, and they compartmentalize and concentrate biomolecules involved in shared regulatory processes in cells. Early cytologists observed the largest and most stable of these compartments over a century ago: the nucleolus and Cajal bodies^{1–5}. There is now evidence that myriad cellular processes, including DNA repair, transcription, signaling, ribosome biogenesis, synaptic transmission, innate immune recognition and others involve condensates^{6–15}. Condensates thus help segregate subsets of the billions of molecules in a cell into compartments with specific functions.

Transient and multivalent noncovalent interactions among biopolymers drive the formation and behavior of condensates^{16–19}. The nature of these interactions is diverse and driven by contributions from dispersion, electrostatic, ionic and electron-donating interactions alongside the hydrophobic effect. Weak multivalent interactions are thought to drive formation of condensates through phase separation^{16–18,20–24}, although additional mechanisms have been proposed^{25,26}. Biomolecular condensates form and dissolve in the face of external and internal stimuli, and the nature of regulatory stimuli dictates their life-time and size^{6,13,27–35}. These attributes are conferred by each condensate's mesoscopic nature, in which both bulk and microscopic physical and chemical behaviors are important.

Here we discuss the chemical mechanisms that contribute to biomolecular condensate formation and regulation, summarize evidence that small molecules can concentrate in specific condensates and argue that deeper insights into condensate chemistry will lead to improvements in therapeutic molecules (Fig. 1). Condensates formed with different biomolecules have distinct chemical specificities that cause other biomolecules or small molecules to selectively associate with the internal condensate chemical environment. Anticancer drugs have been observed to selectively partition into particular condensates, and there is evidence that the chemical interactions that govern small-molecule–biomolecule interactions can be separated from the biomolecular interactions that govern condensate formation³⁶. Thus, the chemical properties that govern molecular interactions between a small molecule and biomolecules within a condensate can be ascertained for both the small molecule

and biomolecules and can be distinguished as contributing to either the small molecule's interaction with the condensate environment or with its target biomolecule.

Chemical mechanisms in condensate assembly and behavior

Biomolecular condensates generally consist of assemblies of protein and RNA molecules but may also contain DNA and other biomolecules. To gain insights into the chemical mechanisms that contribute to these complex assemblies, investigators have typically studied how purified protein or RNA molecules contribute to condensate assemblies in vitro. For example, purified proteins that can form condensates in vitro have been studied for the contributions of specific amino acids to condensate formation and behavior^{16,17,20,37}. RNA molecules have also been observed to form condensates, doing so by base pairing with one another or through interactions with RNA-binding proteins in vitro^{29,30,38–42}. Thus, diverse features of the complex environment of cellular condensates have often been modeled with simple systems in vitro and, when feasible, these findings have been extended to the more complex condensate environment in cells.

A 'stickers-and-spacers' model, in which polymers contain blocks with strong interaction potentials separated by blocks with weak interaction potentials, has provided a valuable coarse grain approach to describe the features of diverse biomolecules that promote condensate formation and that influence their material properties (Fig. 2a) ^{17,18,20,24,43–48}. The strength of sticker–sticker interaction potentials, their frequency and the patterning of stickers and spacers has been shown with the FET (composed of FUS, EWSR1 and TAF15) family and other proteins to govern the saturation concentration of a condensate^{16–18}. Above this threshold concentration, a system of associative polymers will undergo a phase transition, forming dense and dilute phases. Dense phases are characterized by satisfying the majority of sticker–sticker interactions, and dilute phases are defined by satisfaction of fewer sticker–sticker interactions; these potentials vary with space and in response to different chemical and biological phenomena. This remarkably simple model has proven useful in interpreting experiments with model protein and nucleic acids and predicting the ability of sequences to engender condensates.

¹Whitehead Institute for Biomedical Research, Cambridge, MA, USA. ²Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA.
✉e-mail: hkilgore@wi.mit.edu; young@wi.mit.edu

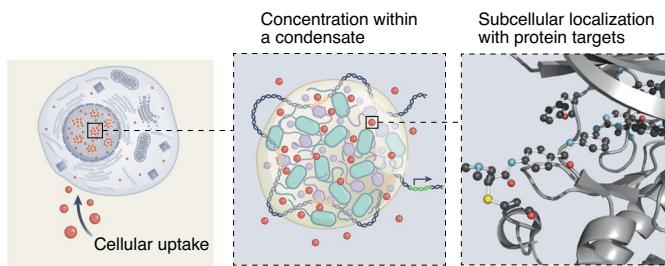


Fig. 1 | Small molecules can concentrate selectively within specific biomolecular condensates and may do so both through interactions with the chemical environment of the condensate and through interactions with target proteins or nucleic acids within the condensate. Condensates are mesoscopic bodies, and interaction of a molecule with the bulk solvation and microscopic chemical environments is determined by a condensate's chemical specificity. This includes both specific interactions with target binding sites and interactions with other parts of biomolecules not constituting the target site that may contribute to the concentration of small molecules in a condensate. Left, small molecules (red spheres) that enter cells can be distributed unequally among diverse membrane-bound and non-membrane compartments. Middle, small molecules (red spheres) concentrating in a transcriptional condensate consisting of a transcription apparatus assembled at a DNA locus; this can occur when the interaction of the small molecule with the chemical environment is favored over its interaction with the environment outside the condensate. Right, interaction of the small molecule is depicted with a specific biomolecular target, shown here: THZ1 bound to CDK7 (Protein Data Bank ID 6XD3).

The roles of diverse amino acids in the formation and behavior of simple protein condensates, which have been called the 'molecular grammar' of these condensates, have come predominantly from studies of the 'fused in sarcoma' (FUS) protein (Fig. 2b)^{17,49,50}. FUS forms multimolecular condensates at low micromolar concentrations, driven by transient and multivalent noncovalent interactions between its arginine and tyrosine residues. In diverse proteins, multivalent noncovalent interactions between acidic, basic and aromatic amino acids have been described that contribute to condensate formation and behavior^{17,18,20,21,23,38,43,44,50–52}. In FUS, spacer regions composed predominantly of glycine, serine and glutamine residues occur between the blocks of strongly interacting groups. Spacer regions with a higher content of glycine residues yielded a more dynamic and liquid-like condensate, while a more solid-like condensate was observed with a higher glutamine content (Fig. 2b). Analysis of how and where these substitutions resulted in different physical properties was facilitated by the stickers-and-spacers framework^{17,24,38,44}. Thus, FET family proteins have provided useful systems for interrogating how condensate behavior is produced by their amino acid sequences and compositions^{17,38}.

Condensates can also be generated by interactions between structured domains that are bridged by a ligand, enabling a cell to achieve thermodynamic control over phase separation. A theoretical framework put forth by Wyman and Gill⁵³ described how protein ligands may exert control over phase separation in a process termed polyphasic linkage, and recent studies have provided additional experimental confirmation of their predictions^{33,54,55}. Bivalent small molecules that have the effect of enhancing or reducing phase separation may prove to be a creative application of this concept^{56,57}.

Evidence for chemical specificities in condensates

The chemical milieu of condensates has been proposed to be a chemically distinct microenvironment where certain protein and RNA molecules are densely concentrated and where these biomolecules will together solvate and enrich for specific sets of additional

molecules while excluding others^{58–63}. Thus, a condensate that contains proteins and nucleic acids engaged in a specific function would be expected to have a microenvironment with chemical features that are distinct from those of condensates with different biomolecules engaged in other functions. Condensates involved in different functions can be visualized in cells by imaging proteins that are specific to these bodies (Fig. 3a). These condensates exhibit a variety of properties that are consistent with the notion that they harbor different physicochemical environments; for example, some form and dissolve in short time frames (for example, transcriptional condensates), whereas others remain assembled for much longer times (for example, nucleoli)^{34,64}.

Selective partitioning of biomolecules. Does the collection of biomolecules that distinguish one type of condensate from another create substantially different chemical microenvironments? The selective condensate partitioning behavior of chemically modified biomolecules provides one line of evidence consistent with this possibility. The chemical modification of a protein molecule can cause that molecule to change its condensate partitioning behavior, that is, to exit one condensate and enter another with different components^{28,52}. The behavior of RNA polymerase II during two different stages of transcription, initiation and elongation, provides an example of this type of modification-dependent condensate partitioning (Fig. 3b)^{28,65,66}. RNA polymerase II can be recruited into transcriptional condensates (compartments marked by the presence of the mediator complex subunit 1 (MED1) protein) during transcription initiation in a form that is minimally phosphorylated. The heptapeptide repeat domain of the largest subunit of RNA polymerase becomes hyperphosphorylated during the transition to elongation, reducing the enzyme's affinity for transcriptional condensates and increasing its affinity for condensates containing the RNA-splicing apparatus. This provides a form of spatiotemporal regulation of transcription, in which the apparatus involved in initiation of RNA synthesis and that involved in RNA processing occur in a slightly different space and time.

What chemical features drive the change in condensate partitioning of the polymerase molecule described above? The 'interaction potentials' of stickers and spacers can be altered and the behavior of the condensate assembly can be changed with amino acid modification, a common event in biological regulation^{17,23,24,27,28,42,49,50,55,67}. For example, the effective interaction potential of a phosphotyrosine and an arginine residue will be significantly greater than that between a tyrosine and an arginine (Fig. 3c). This variation arises because the ionic bond between the phosphate anion and the arginine cation is an order of magnitude greater than the cation–π and hydrogen bonding interactions present in the unphosphorylated case (Fig. 3c). Phosphorylation also endows a greater capacity to engage in hydrogen bonding interactions on a phosphotyrosine residue, drastically influencing its hydrophilicity and likely the local chemical structure of water and inorganic ions. Biological regulatory modifications that alter the pK_a of a side chain, modulate electrostatic surface potential, ablate hydrogen bonding capacity, alter the hydrophobicity of chemical groups or alter the flexibility of amino acids and nucleobases will directly influence the effective interaction potential of a spacer or sticker block.

There are other models that can account for selective partitioning of biomolecules such as that observed with RNA polymerase II modification. It is possible, for example, that proteins resident in one condensate have high-affinity binding sites for unmodified polymerase molecules, whereas proteins resident in the other have high-affinity sites for the modified enzyme. In the case of RNA polymerase II, however, experiments have shown that the heptapeptide repeat domain subjected to phosphorylation exhibits modification-dependent partitioning into simple condensates consisting of proteins representative of transcriptional and splicing

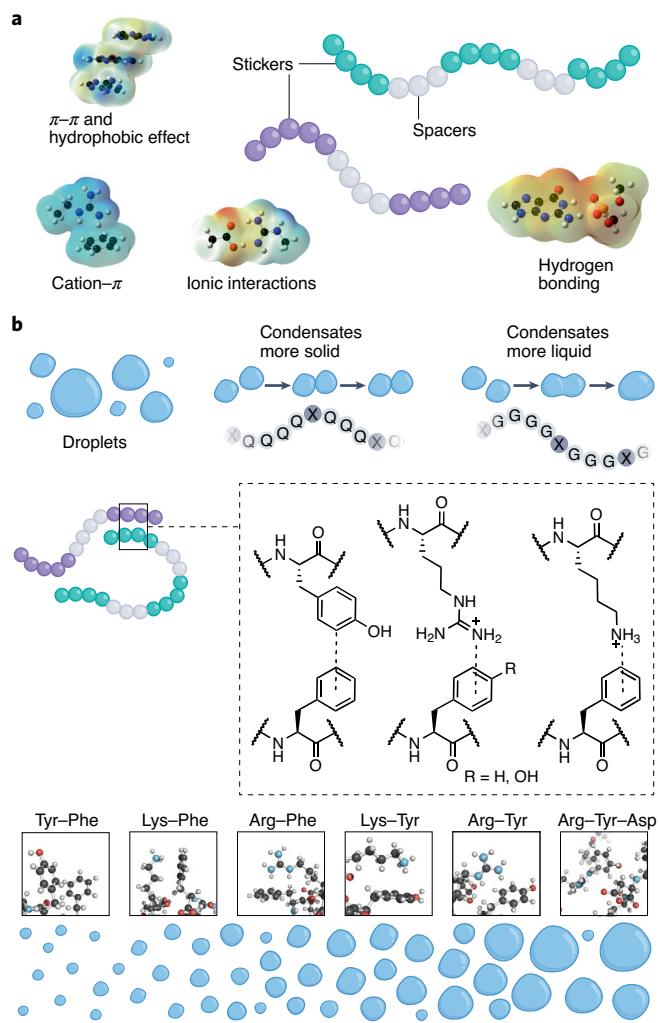


Fig. 2 | Chemical mechanisms in biomolecular condensate assembly.

a, The stickers-and-spacers model suggests how polymers may associate to form condensates; amino acid side chains, nucleobases and folded domains may produce interactions that qualify as stickers or spacers. Examples^{17,18,20,21,23,38,44,45,51–53} of noncovalent interactions mediating sticker interactions are displayed with electrostatic surface potentials, computed at the M06-2x/6-311+g(d,p) level of theory¹⁰. **b**, Condensates form droplets due to the association of biopolymers through noncovalent interactions between sticker regions, which are separated by spacer regions in the same biopolymer. In the model system FUS, glutamine-rich spacers produced more solid-like behavior than the liquid-like character of glycine-rich spacer regions^{18,44}. Noncovalent interactions between amino acids with π -systems and other π -systems or cationic amino acids influence the formation and dissolution of FUS and other protein condensates by creating 'sticker' domains^{17,18,20,21,23,38,43,44,50–52}. These interactions are a consequence of a molecular grammar, rules by which specific amino acids influence condensate formation and behavior¹⁷.

condensates that do not have strong binding interactions with the heptapeptide repeat domain (Fig. 3b). Thus, differences in condensate chemical environments can be exploited by cells to evolve regulatory mechanisms that involve selective partitioning of biomolecules due to chemical modification.

Selective partitioning of small molecules. Diverse small-molecule drugs have been observed to concentrate in biomolecular condensates. Mitoxantrone, a chemotherapy used to treat some forms of

cancer, was long ago observed to concentrate within the nucleolus^{68,69}. Additional anticancer drugs, such as cisplatin and tamoxifen, have now been demonstrated to concentrate in transcriptional condensates and act on DNA or protein targets that occur within those condensates³⁶. This selective concentrating behavior in transcriptional condensates is not dependent on an interaction with the ultimate target of these drugs (DNA for cisplatin and the estrogen receptor (ER) for tamoxifen) but rather appears to be due to interactions with specific amino acid residues in the MED1 protein, a key coactivator protein that is a defining component of transcriptional condensates^{34,70}. Thus, drug molecules can exploit both condensate partitioning properties and those involved in target engagement to concentrate in the same compartment as their target.

The chemical properties responsible for selective partitioning of small molecules into specific condensates are likely to be the same as those that enable selective partitioning of biomolecules in condensates (Fig. 2a). As with biomolecules, in vitro droplet models have proven powerful for interrogating the chemical partitioning behavior of small molecules within condensates. These assays can be conducted with wild-type and mutant forms of protein molecules to identify amino acids that are essential for the partitioning properties of drugs (Fig. 4a). Such an approach led to the observation that cisplatin partitioning into MED1 condensates depends on aromatic amino acid residues and thus cation- π and $\pi-\pi$ interactions associated with these residues and revealed that the amino acids necessary for MED1 phase separation (conserved serine patches) are different from those necessary for the interactions with cisplatin³⁶ (Fig. 4b). Furthermore, experiments with MED1 in vitro droplets have suggested that structure–activity relationships between small molecules and biomolecules in condensates can be deduced with a limited range of fluorescent probes³⁶. These results suggest that the chemical properties that govern molecular interactions between a small molecule and biomolecules within a condensate can be ascertained for both the small molecule and the biomolecules of interest (Fig. 4c) and can be distinguished as contributing to either the small molecule's interaction with the condensate environment or with its target biomolecule.

New models for drug action and resistance

The targets of many commonly used drugs are now known to occur in condensates; therefore, it might be expected that efficacious drugs can readily access these compartments to engage their targets. Nonetheless, an understanding of the interaction of these drugs with the physicochemical environment of diverse condensates, currently lacking for most drugs, may present opportunities for improved therapeutics. As examples of the insights that can emerge from such understanding, we describe below concepts that have emerged from recent studies of drug–condensate interaction that suggest new models to account for therapeutic efficacy and resistance for widely used antineoplastic drugs.

Enhanced pharmacological specificity and activity. Before the observation of selective partitioning in transcriptional condensates, it was widely assumed that the efficacy of cisplatin was due to random platination of the genome, thus selectively affecting tumor cells because they must repair their damaged DNA before continuing to replicate. However, the evidence that cisplatin concentrates in large transcriptional condensates at driver oncogenes, where it selectively platinate oncogene regulatory DNA, suggests a very different model for the drug's anticancer activities (Fig. 5). In this model, the reason that cisplatin has efficacy against a broad spectrum of cancers is due to the fact that each cancer evolves large and stable transcriptional condensates at its driver oncogenes and that cisplatin concentrates and acts on the DNA in these condensates, thus ultimately destroying the ability of the transcription apparatus to operate specifically at those oncogenes. By contrast, the transcriptional condensates in

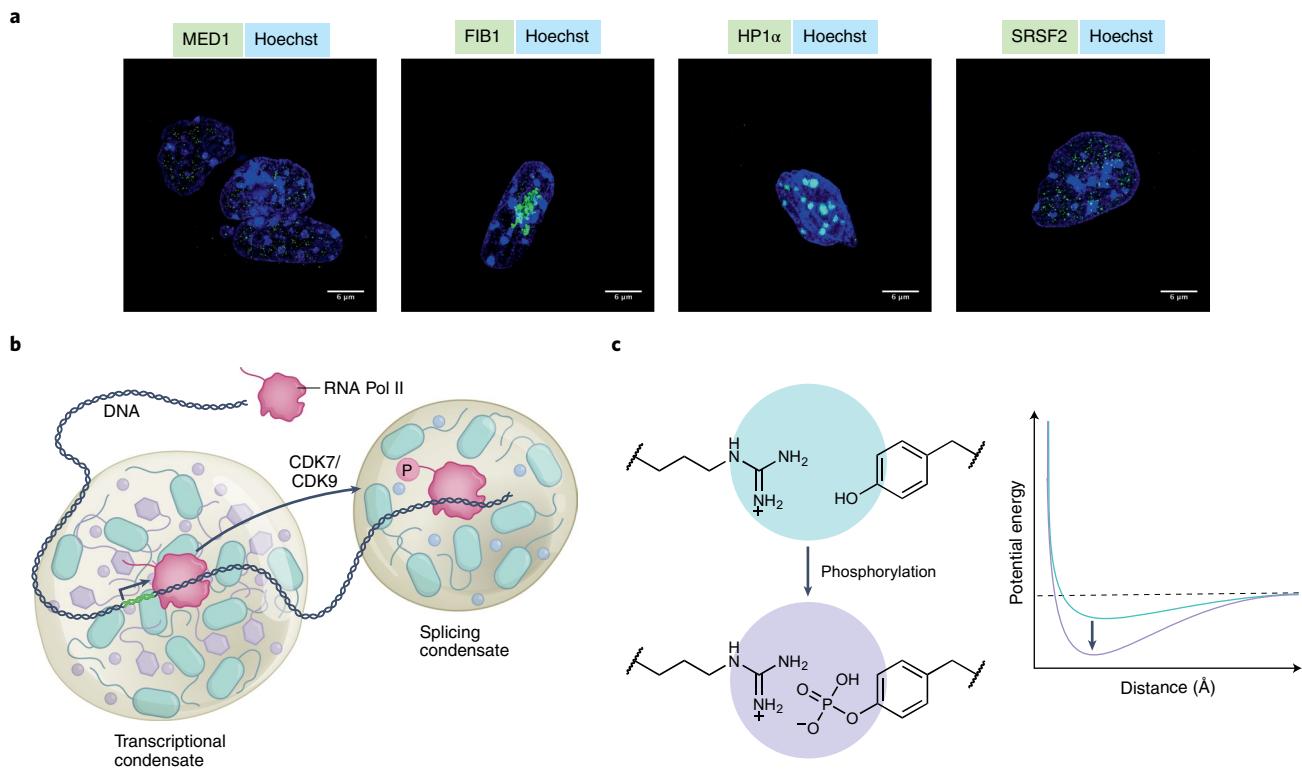


Fig. 3 | Biomolecular condensates composed of different protein components have different chemical environments that engender selective partitioning of biomolecules. **a**, Condensates involved in different functions can be visualized in cells by imaging proteins that are specific to these bodies (for example, MED1 in transcriptional activation, FIB1 in nucleolar ribosome biosynthesis, HP1 α in heterochromatic gene silencing and SRSF2 in RNA splicing). Images of murine embryonic stem cells with green fluorescent protein-tagged proteins (green) and Hoechst staining (blue) acquired with a DeltaVision OM super-resolution microscope. **b**, Model illustrating how chemical modification of a protein molecule can cause that molecule to change its condensate partitioning behavior. The RNA polymerase (Pol) II C-terminal domain becomes hyperphosphorylated during the transition to elongation, reducing the enzyme's affinity for transcriptional condensates and increasing its affinity for splicing condensates. The RNA polymerase II C-terminal domain kinases CDK7 and CDK9 play well-established regulatory roles in transcription²⁸. **c**, Energy diagram showing how post-translational modifications such as phosphorylation can alter the strength of noncovalent interactions among amino acid side chains. For example, phosphorylation of tyrosine residues will increase the interaction potential with an arginine residue.

normal cells are much smaller and have shorter lifetimes and thus accumulate less drug^{70–74} (Fig. 5).

There are now many examples of small-molecule drugs and tool compounds that selectively partition into specific condensates in the absence of their defined target proteins. This includes the drugs cisplatin, mitoxantrone and tamoxifen as well as chemical probes that bind the transcriptional cofactors BRD4 and CDK7 (ref. ³⁶). Before these observations, there was a conundrum: BRD4 and CDK7 are present at all active genes and necessary for their transcription, yet BRD4 and CDK7 inhibitors selectively disrupted the tumor-specific oncogenes that engendered oncogenic properties of these diverse cancer cells^{72–74}. Again, the facts that most cancers evolve large and stable transcriptional condensates at their driver oncogenes and that these BRD4 and CDK7 inhibitors are selectively concentrated in such condensates together explain why these inhibitors have oncogene-selective inhibitory activities and why they have far less deleterious effects in normal cells, which have smaller and more transient transcriptional condensates (Fig. 5).

New mechanisms of drug resistance. Tamoxifen is an anti-estrogen drug that is highly effective in the treatment of ER-positive breast cancer. Tamoxifen resistance can be conferred by ER mutations that reduce drug affinity, as might be expected, but it can also be conferred by MED1 overexpression, which, until recently, did not have a mechanistic explanation^{75,76}. We found that ER partitions selectively

into MED1-containing transcriptional condensates in a manner that is dependent on its binding to estradiol, but, when tamoxifen is present, the drug partitions selectively into the same transcriptional condensates and competes for ER binding with estradiol; tamoxifen binding leads to eviction of ER from the transcriptional condensate³⁶. MED1 overexpression was found to cause an expansion of the volume of transcriptional condensates, thereby diluting tamoxifen in the condensate and rendering tamoxifen less efficient in evicting ER from the condensate. These results suggest that misregulation of genes, a hallmark of cancer, can lead to condensate alterations that contribute to drug resistance in cancer cells.

Chemical grammar and condensate compartments

We suggest that learning the ‘chemical grammar’ of molecules with respect to condensates, which we define as the rules describing the chemical features of molecules that engender attraction to or repulsion by the physicochemical environment of a specific condensate, should enable design of small-molecule drugs with three types of condensate-associated properties. It should be possible to endow small-molecule drugs with chemical properties that (1) concentrate these molecules to higher levels in condensates where their targets occur and lower levels in condensates where toxic effects might be obtained (Fig. 6a), (2) modulate the phase behavior of specific condensates (Fig. 6b) and (3) modify the material properties of condensates (Fig. 6c,d). Some small molecules may impact

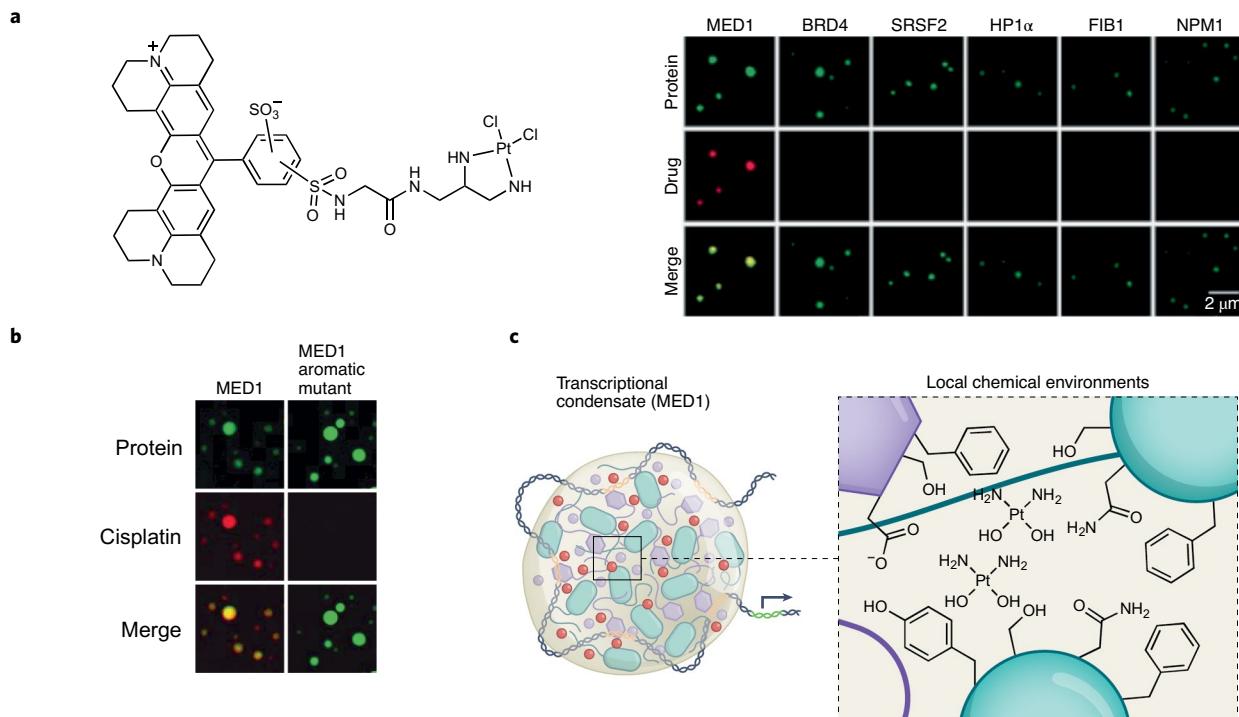


Fig. 4 | Chemical specificity for small molecules in condensates. **a**, Cisplatin–Texas Red selectively concentrates in specific condensates in droplet partitioning assays³⁶. NPM1, nucleophosmin 1. **b**, Changes in amino acid composition of a condensate-forming protein can abrogate small-molecule partitioning behavior without affecting condensate formation. The ability of droplets formed by the MED1 protein of the mediator complex to concentrate cisplatin is reduced with the replacement of aromatic residues with alanine residues³⁶. **c**, Model depicting how specific local chemical environments within a condensate may influence cisplatin partitioning and concentration within a transcriptional condensate.

more than one of these properties; therefore, these are not necessarily mutually exclusive.

Condensate partitioning. It should be possible to engineer small molecules that not only bind selectively to their target proteins but also selectively concentrate in the particular condensates where their targets occur, thereby improving their pharmacological efficacy. In cases in which the targets of small molecules involve families of proteins with similar sequences, such as kinases, and in which such proteins reside in different condensates, some degree of target specificity might be obtained through selective condensate partitioning. Such an approach would benefit from an optimal balance between ligand binding and condensate interaction.

It is also possible that small-molecule partitioning into inappropriate condensates contributes to toxicity. For example, a DNA-modifying drug that concentrates in nucleoli will disrupt the process of ribosome biosynthesis. Mitoxantrone concentrates in nucleoli as well as other condensates, and it is possible that its impact on nucleoli contributes to its toxicity, whereas its impact in other condensates contributes to its efficacy⁷⁷.

Although we do not yet know of an instance in which chemical features that influence partitioning have been purposely incorporated into the design of a small-molecule drug, there is evidence supporting the notion that this can be accomplished. Investigators gained insights into the chemical features of small molecules that contribute to selective concentration in MED1 protein condensates by screening a fluorescent probe library of boron-dipyrromethene (BODIPY) dyes diversified with different chemical functional groups³⁶. Derivatization of the common fluorescent scaffold showed that aromatic rings preferentially engender concentration into these condensates, suggesting that π – π or cation– π interactions are among the noncovalent interactions leading to small-molecule partitioning

in this compartment. Thus, a small molecule that targets a biomolecule resident in a MED1-dominated condensate might be modified to include an aromatic ring to enhance its ability to be selectively concentrated in this condensate.

Modulating phase behavior. Condensate dysregulation due to pathogenic mutations in condensate-associated proteins is now thought to contribute to diverse diseases^{9,10,12,14}. In diseases caused by condensate dysregulation, it should be possible to use small molecules to selectively modulate the saturation concentration of specific condensate-forming proteins. In this manner, both noncovalent and covalent small-molecule interactions with proteins have the potential to rescue the phase behavior of a condensate that has become dysregulated due to a pathological mutation.

Solutions of condensate-forming biopolymers will ‘demix’, separating into biopolymer-dense and biopolymer-dilute phases, at their saturation concentration. Noncovalent interactions between side chains and molecules in solution can impact condensate size and dynamics by providing interaction partners^{24,42,55,78–81}. This concept is directly portable to small molecules that interact directly with a protein to change its saturation concentration (Fig. 6b)^{79,80,82}. Specific chemical features have been shown to engage in the transforming of a condensate’s saturation concentration and may consist of a mixture of strong and weak ions, hydrophobic regions and electron-rich π -systems⁸⁰. These attributes are complicit with strong nonspecific interactions with proteins, driving the formation of transient noncovalent interactions between components of proteins important to condensate formation, mechanisms that are reminiscent of the ‘molecular grammar’ of condensate-forming proteins, which are the rules describing how amino acid chemistries influence protein phase separation and condensate material properties¹⁷.

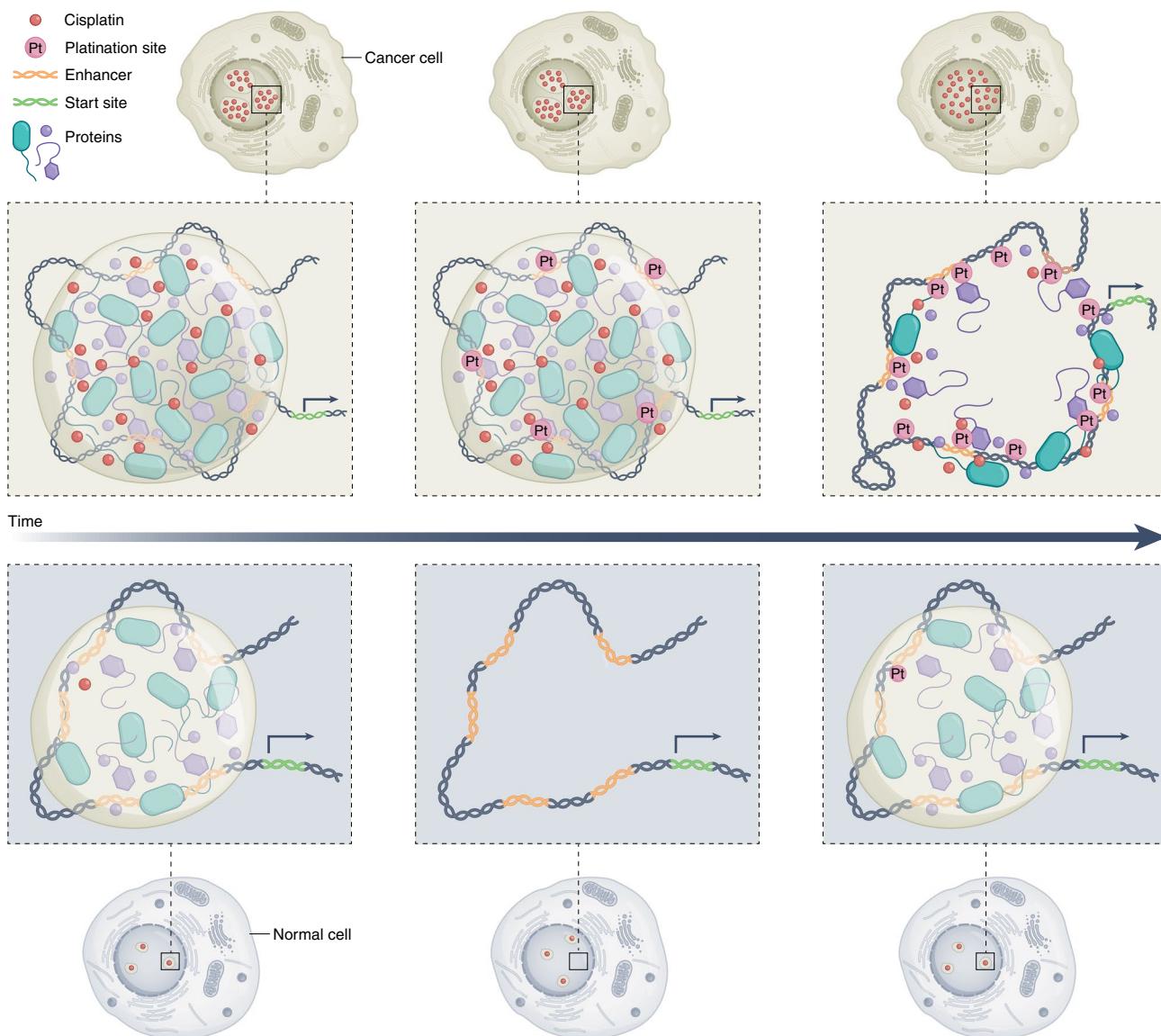


Fig. 5 | Cisplatin concentrates in large transcriptional condensates at driver oncogenes, where it selectively platinates oncogene regulatory DNA³⁶.

Transcriptional condensates have been shown to occur at loci that contain clusters of enhancer regulatory elements called super-enhancers^{34,70}. In normal cells, super-enhancers typically span 5–20 kb of DNA, but in metastatic tumor cells, driver oncogenes acquire super-enhancers that can span as much as 500 kb^{71–74}. The larger super-enhancers are associated with larger amounts of the assembled transcription apparatus and thus larger condensates. Larger transcriptional condensates have longer half-lives and produce more transcription from their associated genes^{33,34}. Thus, the continuous high concentration of cisplatin within the longer-lived oncogenic transcriptional condensates leads to robust DNA platinination at tumor-specific oncogenes, and, ultimately, this permanently disrupts the condensate at the oncogene, leading to tumor cell death. By contrast, the smaller short-lived condensates at normal genes accumulate far less cisplatin and suffer far less DNA damage.

Modulating material properties. Condensates have material properties (viscosity, viscoelasticity, surface tension and diffusivity of molecules within condensates) that result from the chemical properties of the polymers from which they are composed (Fig. 6c), and altered material properties can produce dysregulated condensates in diverse diseases^{7,12,17,83}. Targeting these different physical properties in therapeutic design is an emergent approach for addressing diseases thought to involve condensate dysregulation.

Condensates can be described as liquids, gels or solids, which are terms that reflect the relative diffusivity of molecules within these assemblies. A hallmark of neurodegenerative diseases, which include amyotrophic lateral sclerosis, frontotemporal dementia, Alzheimer's disease and Parkinson's disease, is the formation of protein or RNA aggregates that appear to be dysregulated condensates.

Four well-studied proteins found aggregated in different neurodegenerative diseases (α -synuclein, FUS, tau and TAR DNA-binding protein 43 (TDP-43)) have each been shown to participate in phase separation, and the formation of aggregates is postulated to represent a solid or glass-like pathogenic condensate. In this context, small molecules may prove to have advantages over biomolecules as therapeutics for such neurodegenerative diseases, as they retain a greater capacity to penetrate into disease-state condensates with solid or glass-like material properties.

Condensate assemblies are often employed by viruses to compartmentalize essential life cycle functions, and recent studies have shown that small molecules can have efficacy against viral replication by modulating the properties of viral condensates. Viruses can hijack host cell compartments to provide a protected and metabolite-rich

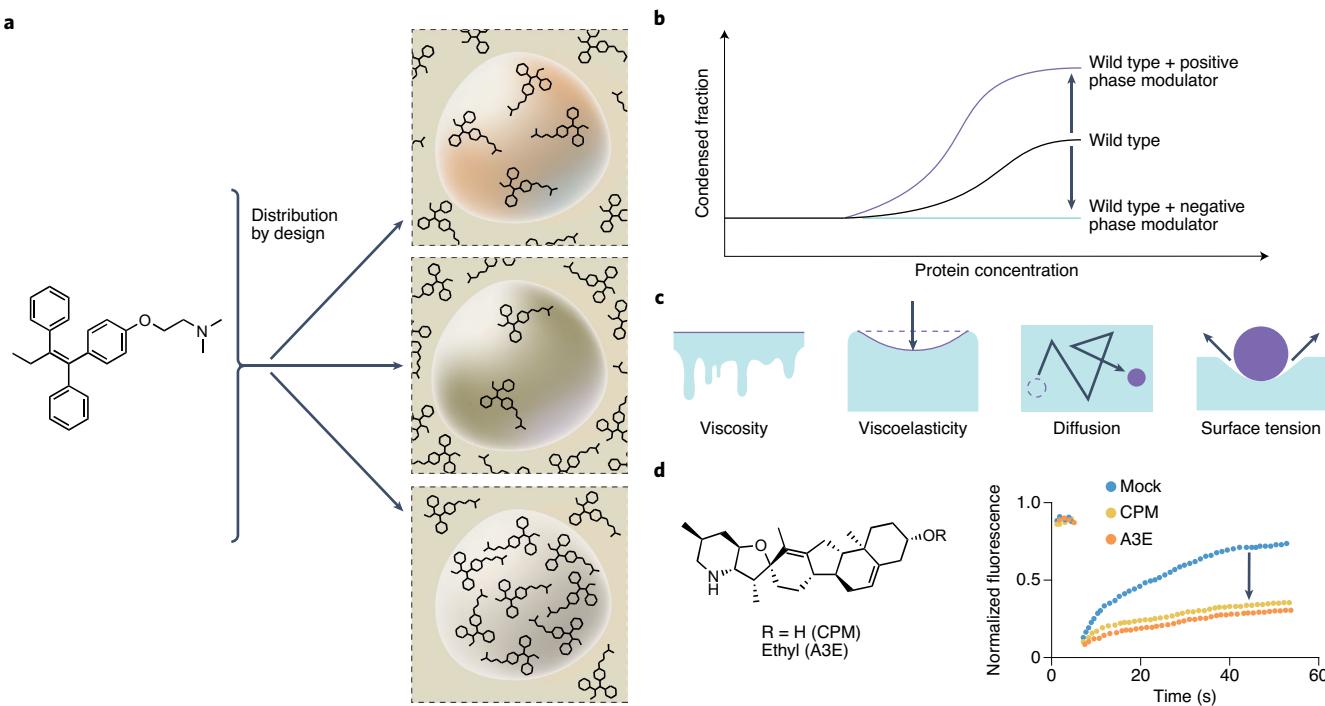


Fig. 6 | Small molecules and peptides may be designed to have chemical properties that interact with and influence condensates to improve therapeutic efficacy. **a**, It should be possible to endow different classes of small molecules and peptides with features that concentrate these molecules in condensates where their targets occur and cause them to avoid partitioning into condensates where toxic effects might transpire. **b**, Alteration of a condensate's condensed fraction and phase behavior with a positive (increasing) or negative (decreasing) phase modulator. **c**, Material properties of condensate may be altered with small molecules; viscosity, viscoelasticity, surface tension and diffusivity of molecules may be augmented, corrected or depleted by drugs. **d**, Cyclopamine is an example of a small molecule that can induce changes in respiratory syncytial virus condensates by hardening (decreased viscoelasticity)⁸⁹. A3E, A-ring, 3'-ethoxy-cyclopamine; CPM, cyclopamine.

microenvironment for replication while minimizing stimulation of innate immune responses^{40,84–88}. Compartmentalization strategies include hijacking host condensates, the formation of phase-separated inclusion bodies of viral components and re-engineering the endoplasmic reticulum. Improved understanding of the properties of viral condensates could lead to new classes of antiviral drugs. Indeed, modification of the viscoelasticity of respiratory syncytial virus condensates by a small molecule has proven possible and can arrest viral replication (Fig. 6d)⁸⁹. New condensate-driven strategies for the development of antiviral drugs may yield much-needed new directions in this classically challenging field of medicinal chemistry.

Simple and cellular condensates. Scientists have a long tradition of studying the complex environment of a cell by using reductionist approaches with purified components in vitro. The simple systems produced in this fashion are more amenable to theoretical and experimental analysis than the more complex systems in living cells, and most of our understanding of the fundamental behaviors of biological molecules has emerged from this approach. Thus, the study of simple homotypic protein condensates will continue to reveal fundamental insights such as the chemical features of small molecules that engender attraction or repulsion to condensate microenvironments and the ability of small molecules to modulate condensate phase behavior and material properties.

Once the fundamentals are established in simple systems, the question of relevance to the more complex living system arises. In this context, the simple condensates produced with a single type of protein are unlikely to have the same internal chemical environment that occurs in a cellular condensate where that protein functions with diverse other molecules in a nonequilibrium environment. This leads us to ask to what extent are insights gained from studies

of simple condensates in vitro predictive of biochemical behaviors of more complex condensates in living cells? The answer is that we do not know, but, based on a limited amount of data with biomolecules and small molecules in studies of transcriptional condensates, observations with simple systems can be predictive of behaviors in vivo^{27,34,36,70}. For example, the selective cisplatin- and tamoxifen-concentrating behaviors of simple MED1 protein condensates extend to condensates formed by the 30-subunit mediator complex in vitro and transcriptional condensates in vivo²⁵.

How is it possible that observations with simple condensate systems can be predictive of behaviors in the much more complex microenvironments of condensates in vivo? Certain proteins in cellular condensates have been proposed to play dominant roles by acting as 'scaffolds' for other 'client' proteins^{15,54}. Proteins that have been proposed to act as scaffolds include MED1 in transcriptional activation, FIB1 in nucleolar ribosome biosynthesis, HP1 α in heterochromatic gene silencing and SRSF2 in RNA splicing. Despite being assemblies of many different biomolecules, it is possible that the internal chemical microenvironment of some condensates is dominated by the chemical features of their scaffolds, and, if so, this could account for the ability of some small molecules to concentrate selectively in both simple condensates containing a scaffold and in the more complex condensate with that scaffold in cells. We suspect, however, that the diverse population of molecules in any one type of cellular condensates creates a chemical microenvironment that is not well replicated in homotypic in vitro condensates.

We imagine that small molecules are not distributed such that they concentrate primarily in a single type of favored condensates but rather are distributed such that they concentrate to different levels in diverse cellular condensates. Furthermore, the physico-chemistry of a cellular condensate can be modulated dynamically

by diverse clients such as proteins, nucleic acids, metabolites and ions, and such modulation is likely to alter small-molecule partitioning. Thus, the chemical grammar of the cellular condensates will be more challenging to discern than that of simple in vitro condensates.

Perspective

We envision a time when molecules can be engineered to selectively enrich in any one type of compartment where a target is contributing to a disease phenotype, thus producing therapeutic molecules with improved efficacy and reduced toxicity. To reach that point and have optimal impact, several important advances are needed. The diverse types of condensate compartments that exist in cells will need to be further cataloged, their components described and their physicochemical properties deduced. The rules describing the chemical features of molecules that engender attraction to or repulsion by the physicochemical environment of a specific condensate, which we call chemical grammar, will need to be learned for these diverse condensates. When disease mutations cause pathological dysfunction by altering the material properties of condensates, it will be useful to obtain a deeper understanding of the means by which small molecules can modify the viscosity, viscoelasticity, surface tension and diffusivity of condensates.

Conceptual and experimental innovation has led to a revolution in our understanding of the compartmental features of cells in the past decade. Conceptual innovations have come from introducing concepts from polymer chemistry and soft matter physics into regulatory biology. We suggest that further conceptual advances will come from treating cells as highly dynamic nonequilibrium environments. Similarly, there have been experimental innovations employing engineered cells and molecules. We suggest that learning condensate chemical grammar will be enhanced and accelerated by combining experimental strategies of chemical biology with modern computational approaches, such as deep learning. This should lead to new insights into the mechanics of how molecules are compartmentalized in and affect the material properties of condensates and ultimately enable the development of more potent therapeutics.

Received: 20 June 2021; Accepted: 21 April 2022;

Published online: 27 June 2022

References

1. Wagner, R. Einige Bemerkungen und Fragen über das Keimbläschen (vesicular germinativa). *Müller's Archiv. Anat. Physiol. Wissenschaft Med.* **1835**, 373–377 (1835). **Wagner was apparently the first to publish a description of the largest and best-characterized condensate, the nucleolus.**
2. Valentin, G. *Repertorium für Anatomie und Physiologie*, Vol. 1 (Verlag von Veit und Comp., 1836).
3. Valentin, G. *Repertorium für Anatomie und Physiologie*, Vol. 4 (Verlag von Veit und Comp., 1839).
4. Cajal, S. R. Y. Un sencillo metodo de coloracion selectiva del reticulo protoplasmatico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados. *Trab. Lab. Invest. Biol.* **2**, 129–221 (1903).
5. Vincent, W. S. The isolation and chemical properties of the nucleoli of starfish oocytes. *Proc. Natl Acad. Sci. USA* **38**, 139–145 (1952).
6. Brangwynne, C. P. et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009). **This pioneering study noted that the behavior of a cellular body was not that of a standard membrane-bound organelle but rather more like liquid droplets that arise by phase separation.**
7. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and challenges in studying liquid–liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 (2019).
8. Case, L. B., Ditlev, J. A. & Rosen, M. K. Regulation of transmembrane signaling by phase separation. *Annu. Rev. Biophys.* **48**, 465–494 (2019).
9. Alberti, S. & Dормann, D. Liquid–liquid phase separation in disease. *Annu. Rev. Genet.* **53**, 171–194 (2019).
10. Boija, A., Klein, I. A. & Young, R. A. Biomolecular condensates and cancer. *Cancer Cell* **39**, 174–192 (2021).
11. Sabari, B. R., Dall'Agnese, A. & Young, R. A. Biomolecular condensates in the nucleus. *Trends Biochem. Sci.* **45**, 961–977 (2020).
12. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
13. Chen, X., Wu, X., Wu, H. & Zhang, M. Phase separation at the synapse. *Nat. Neurosci.* **23**, 301–310 (2020).
14. Nedelsky, N. B. & Taylor, J. P. Bridging biophysics and neurology: aberrant phase transitions in neurodegenerative disease. *Nat. Rev. Neurol.* **15**, 272–286 (2019).
15. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
16. Pak, C. W. et al. Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Mol. Cell* **63**, 72–85 (2016). **This study provides a powerful example of combining theory and experimentation to enhance understanding of the roles of amino acid patterns in protein phase separation.**
17. Wang, J. et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699 (2018). **A study exploring the roles of diverse amino acids in promoting phase separation and influencing the material properties of condensates formed by a class of proteins.**
18. Martin, E. W. et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).
19. Banani, S. F. et al. Composition control of phase-separated bodies. *Cell* **166**, 651–663 (2016).
20. Abyzov, A., Blackledge, M. & Zweckstetter, M. Conformational dynamics of intrinsically disordered proteins regulate biomolecular condensate chemistry. *Chem. Rev.* **122**, 6719–6748 (2022).
21. Greig, J. A. et al. Arginine-enriched mixed-charge domains provide cohesion for nuclear speckle condensation. *Mol. Cell* **77**, 1237–1250 (2020).
22. Case, L. B., Zhang, X., Ditlev, J. A. & Rosen, M. K. Stoichiometry controls activity of phase-separated clusters of actin signaling proteins. *Science* **363**, 1093–1097 (2019).
23. Shrinivas, K. et al. Enhancer features that drive formation of transcriptional condensates. *Mol. Cell* **75**, 549–561 (2019).
24. Choi, J.-M., Holehouse, A. S. & Pappu, R. V. Physical principles underlying the complex biology of intracellular phase transitions. *Annu. Rev. Biophys.* **49**, 107–133 (2020).
25. McSwiggen, D. T., Mir, M., Darzacq, X. & Tjian, R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 (2019).
26. Peng, A. & Weber, S. C. Evidence for and against liquid–liquid phase separation in the nucleus. *Noncoding RNA* **5**, 50 (2019).
27. Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855 (2018).
28. Guo, Y. E. et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 (2019). **This work establishes that post-translational modifications of the transcription apparatus can regulate its partitioning into different condensates involved in RNA synthesis and splicing.**
29. Henninger, J. E. et al. RNA-mediated feedback control of transcriptional condensates. *Cell* **184**, 207–225 (2021).
30. Lu, S. et al. The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nat. Commun.* **12**, 502 (2021).
31. Yu, X. et al. The STING phase-separator suppresses innate immune signalling. *Nat. Cell Biol.* **23**, 330–340 (2021).
32. Su, X. et al. Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352**, 595–599 (2016).
33. Dao, T. P. et al. Mechanistic insights into the enhancement or inhibition of phase separation by polyubiquitin chains of different lengths or linkages. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.11.12.467822> (2021).
34. Cho, W.-K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
35. Cho, W.-K. et al. RNA polymerase II cluster dynamics predict mRNA output in living cells. *eLife* **5**, e13617 (2016).
36. Klein, I. A. et al. Partitioning of cancer therapeutics in nuclear condensates. *Science* **368**, 1386–1392 (2020). **This study showed that selective partitioning and concentration of small molecules within condensates contributes to drug pharmacodynamics, suggesting that further understanding of condensate chemical grammar may facilitate advances in disease therapy.**
37. Schuster, B. S. et al. Identifying sequence perturbations to an intrinsically disordered protein that determine its phase-separation behavior. *Proc. Natl. Acad. Sci. USA* **117**, 11421–11431 (2020).
38. Alshareedah, I., Moosa, M. M., Pham, M., Potocyan, D. A. & Banerjee, P. R. Programmable viscoelasticity in protein–RNA condensates with disordered sticker-spacer polypeptides. *Nat. Commun.* **12**, 6620 (2021).

39. Roden, C. & Gladfelter, A. S. RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.* **22**, 183–195 (2021).

40. Savastano, A., Ibáñez de Opakua, A., Rankovic, A. & Zweckstetter, M. Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. *Nat. Commun.* **11**, 6041 (2020).

41. Guillén-Boixet, J. et al. RNA-induced conformational switching and clustering of G3BP drive stress granule assembly by condensation. *Cell* **181**, 346–361 (2020).

42. Sanders, D. W. et al. Competing protein–RNA interaction networks control multiphase intracellular organization. *Cell* **181**, 306–324 (2020).

43. Boeynaems, S. et al. Spontaneous driving forces give rise to protein–RNA condensates with coexisting phases and complex material properties. *Proc. Natl. Acad. Sci. USA* **116**, 7889–7898 (2019).

44. Holehouse, A. S., Ginell, G. M., Griffith, D. & Böke, E. Clustering of aromatic residues in prion-like domains can tune the formation, state, and organization of biomolecular condensates. *Biochemistry* **60**, 3566–3581 (2021).

45. Kar, M. et al. Glycine-rich peptides from FUS have an intrinsic ability to self-assemble into fibers and networked fibrils. *Biochemistry* **60**, 3213–3222 (2021).

46. Rubinstein, M. & Semenov, A. N. Thermoreversible gelation in solutions of associating polymers. 2. Linear dynamics. *Macromolecules* **31**, 1386–1397 (1998).

47. Tanaka, F. & Ishida, M. Microphase formation in mixtures of associating polymers. *Macromolecules* **30**, 1836–1844 (1997).

48. Leibler, L. Theory of microphase separation in block copolymers. *Macromolecules* **13**, 1602–1617 (1980).

49. Murthy, A. C. et al. Molecular interactions underlying liquid–liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **26**, 637–648 (2019).

50. Qamar, S. et al. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation–π interactions. *Cell* **173**, 720–734 (2018).

51. Kim, T. H. et al. Interaction hot spots for phase separation revealed by NMR studies of a CAPRIN1 condensed phase. *Proc. Natl. Acad. Sci. USA* **118**, e2104897118 (2021).

52. Kim, T. H. et al. Phospho-dependent phase separation of FMRP and CAPRIN1 recapitulates regulation of translation and deadenylation. *Science* **365**, 825–829 (2019).

53. Wyman, J. & Gill, S. J. Ligand-linked phase changes in a biological system: applications to sickle cell hemoglobin. *Proc. Natl. Acad. Sci. USA* **77**, 5239–5242 (1980).

54. Ruff, K. M., Dar, F. & Pappu, R. V. Polyphasic linkage and the impact of ligand binding on the regulation of biomolecular condensates. *Biophys. Rev.* **2**, 021302 (2021).

55. Ruff, K. M., Dar, F. & Pappu, R. V. Ligand effects on phase separation of multivalent macromolecules. *Proc. Natl. Acad. Sci. USA* **118**, e2017184118 (2021).

56. Burslem, G. M. & Crews, C. M. Proteolysis-targeting chimeras as therapeutics and tools for biological discovery. *Cell* **181**, 102–114 (2020).

57. Gerry, C. J. & Schreiber, S. L. Unifying principles of bifunctional, proximity-inducing small molecules. *Nat. Chem. Biol.* **16**, 369–378 (2020).

58. Forman-Kay, J. D., Ditlev, J. A., Nosella, M. L. & Lee, H. O. What are the distinguishing features and size requirements of biomolecular condensates and their implications for RNA-containing condensates? *RNA* **28**, 36–47 (2021).

59. Andersen, J. S. et al. Nucleolar proteome dynamics. *Nature* **433**, 77–83 (2005).

60. Su, X. et al. Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352**, 595–599 (2016).

61. Jain, S. et al. ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell* **164**, 487–498 (2016).

62. Langdon, E. M. & Gladfelter, A. S. A new lens for RNA localization: liquid–liquid phase separation. *Annu. Rev. Microbiol.* **72**, 255–271 (2018).

63. Woodruff, J. B., Hyman, A. A. & Boke, E. Organization and function of non-dynamic biomolecular condensates. *Trends Biochem. Sci.* **43**, 81–94 (2018).

64. Lafontaine, D. L. J., Riback, J. A., Bascetin, R. & Brangwynne, C. P. The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol.* **22**, 165–182 (2021).

65. Boehning, M. et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **25**, 833–840 (2018).

66. Nojima, T. et al. RNA polymerase II phosphorylated on CTD serine 5 interacts with the spliceosome during co-transcriptional splicing. *Mol. Cell* **72**, 369–379 (2018).

67. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).

68. Vibet, S. et al. Differential subcellular distribution of mitoxantrone in relation to chemosensitization in two human breast cancer cell lines. *Drug Metab. Dispos.* **35**, 822–828 (2007).

69. Smith, P. J., Sykes, H. R., Fox, M. E. & Furlong, I. J. Subcellular distribution of the anticancer drug mitoxantrone in human and drug-resistant murine cells analyzed by flow cytometry and confocal microscopy and its relationship to the induction of DNA damage. *Cancer Res.* **52**, 4000–4008 (1992).

70. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).

71. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional addiction in cancer. *Cell* **168**, 629–643 (2017).

72. Filippakopoulos, P. et al. Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).

73. Christensen, C. L. et al. Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor. *Cancer Cell* **26**, 909–922 (2014).

74. Kwiatkowski, N. et al. Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* **511**, 616–620 (2014).

75. Fanning, S. W. et al. Estrogen receptor α somatic mutations Y537S and D538G confer breast cancer endocrine resistance by stabilizing the activating function-2 binding conformation. *eLife* **5**, e12792 (2016).

76. Nagalingam, A. et al. Med1 plays a critical role in the development of tamoxifen resistance. *Carcinogenesis* **33**, 918–930 (2012).

77. Pommier, Y., Leo, E., Zhang, H. & Marchand, C. DNA topoisomerase and their poisoning by anticancer and antibacterial drugs. *Chem. Biol.* **17**, 421–433 (2010).

78. Bielskuté, S. et al. Low amounts of heavy water increase the phase separation propensity of a fragment of the androgen receptor activation domain. *Protein Sci.* **30**, 1427–1437 (2021).

79. Petronilho, E. C. et al. Phase separation of p53 precedes aggregation and is affected by oncogenic mutations and ligands. *Chem. Sci.* **12**, 7334–7349 (2021).

80. Babinchak, M. W. et al. Small molecules as potent biphasic modulators of protein liquid–liquid phase separation. *Nat. Commun.* **11**, 5574 (2020).

81. Krainer, G. et al. Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun.* **12**, 1085 (2021).

82. Fang, M. Y. et al. Small-molecule modulation of TDP-43 recruitment to stress granules prevents persistent TDP-43 accumulation in ALS/FTD. *Neuron* **103**, 802–819 (2019).

83. Alshareedah, I., Kaur, T. & Banerjee, P. R. Methods for characterizing the material properties of biomolecular condensates. In *Methods in Enzymology*, Vol. 646 (ed. Keating, C. D.), Ch. 6, 143–183 (Academic Press, 2021).

84. Wang, S. et al. Targeting liquid–liquid phase separation of SARS-CoV-2 nucleocapsid protein promotes innate antiviral immunity by elevating MAVS activity. *Nat. Cell Biol.* **23**, 718–732 (2021).

85. Hadjadj, J. et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718–724 (2020).

86. Li, Y. et al. SARS-CoV-2 induces double-stranded RNA-mediated innate immune responses in respiratory epithelial-derived cells and cardiomyocytes. *Proc. Natl. Acad. Sci. USA* **118**, e2022643118 (2021).

87. Wu, J. et al. SARS-CoV-2 ORF9b inhibits RIG-I–MAVS antiviral signaling by interrupting K63-linked ubiquitination of NEMO. *Cell Rep.* **34**, 108761 (2021).

88. Zotta, A., Hooftman, A. & O’Neill, L. A. J. SARS-CoV-2 targets MAVS for immune evasion. *Nat. Cell Biol.* **23**, 682–683 (2021).

89. Risso-Ballester, J. et al. A condensate-hardening drug blocks RSV replication in vivo. *Nature* **595**, 596–599 (2021).

Acknowledgements

We thank A. Dallagnese for providing the images presented in Fig. 3a and A. Boija and K. Overholt for helpful discussions. Funding: H.R.K. is supported by a fellowship from the Damon Runyon Cancer Research Foundation (grant number 2458-22). R.A.Y. is supported by NIH grants R01 GM123511 and R01 MH104610, NCI grant CA155258, NSF grant PHY2044895, the 3D Genome Consortium of St. Jude Children’s Research Hospital, and funds from Novo Nordisk.

Author contributions

H.R.K. and R.A.Y. conceived of and wrote this perspective.

Competing interests

R.A.Y. is a founder and shareholder of Syros Pharmaceuticals, Camp4 Therapeutics, Omega Therapeutics and Dewpoint Therapeutics. H.R.K. is a consultant of Dewpoint Therapeutics.

Additional information

Correspondence should be addressed to Henry R. Kilgore or Richard A. Young.

Peer review information *Nature Chemical Biology* thanks Tingting Li and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature America, Inc. 2022