HHS Public Access

Author manuscript

Biochemistry. Author manuscript; available in PMC 2023 November 15.

Published in final edited form as:

Biochemistry. 2022 November 15; 61(22): 2443–2455. doi:10.1021/acs.biochem.2c00210.

Principles governing the phase separation of multi-domain proteins

Priyesh Mohanty[*],

Utkarsh Kapoor,

Dinesh Sundaravadivelu Devarajan,

Tien Minh Phan,

Azamat Rizuan,

Jeetain Mittal[*]

Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843

Abstract

A variety of membraneless organelles often referred to as “biological condensates”, play an important role in the regulation of cellular processes such as gene transcription, translation and protein quality control. Based on experimental and theoretical investigations, liquid-liquid phase separation (LLPS) has been proposed as a possible mechanism for the origin of biological condensates. LLPS requires multivalent macromolecules which template the formation of long-range, intermolecular interaction networks and results in the formation of condensates with defined composition and material properties. Multivalent interactions driving LLPS exhibit a wide range of modes from highly stereospecific to non-specific and involve both folded and disordered regions. Multi-domain proteins serve as suitable macromolecules for promoting phase separation and achieving disparate functions due to their potential for multivalent interactions and regulation. Here, we aim to highlight the influence of the domain architecture and inter-domain interactions on the phase separation of multi-domain protein condensates. First, the general principles underlying these interactions are illustrated based on examples of multi-domain proteins which are predominantly associated with nucleic-acid binding and protein quality control, and contain both folded and disordered regions. Next, examples are presented which showcase how LLPS properties of folded and disordered regions can be leveraged to engineer multi-domain constructs that form condensates with desired assembly and functional properties. Finally, we highlight the need for improvements in coarse-grained computational models that can provide molecular-level insights into multi-domain protein condensates in conjunction with experimental efforts.

Introduction

Most proteins in prokaryotes and eukaryotes are composed of two or more domains which can fold independently to adopt unique, three-dimensional structures. An interesting

[*] priyeshm@tamu.edu, jeetain@tamu.edu.

observation is that new proteins which emerge over the course of evolution are often multi-domain in nature as opposed to being single, novel folds[1]. Multi-domain proteins are typically formed via gene duplication, divergence and recombination events[2] giving rise to new combinations and arrangements based on a limited number of pre-existing folds (<10,000). The individual domains are usually connected by linker sequences (30–50 residues) and can exhibit inter-domain flexibility. The presence of additional domains confers new functional capabilities through an increase in the number of potential interaction sites, a phenomenon referred to as multivalency. The acquisition of multivalency leads to an increase in the number of interacting partners and allows multi-domain proteins to connect components of different signaling pathways by acting as interaction "hubs"[2]. New domains also confer novel regulatory capacity by allowing for chemical modifications or non-covalent interactions with other macromolecules, which allows for fine-tuning the cellular response to both internal and external cues.

Nearly 30% of the eukaryotic proteome encodes for intrinsically disordered proteins (IDPs) or regions (IDRs)[3,4] which are attached to folded domains. Unlike their folded counterparts, these protein sequences lack tertiary structure and populate a heterogeneous ensemble of conformations with high solvent accessibility. Over the last twenty years, it has become clear that the sequence diversity and conformational flexibility of IDPs/IDRs imparts versatility in the regulation of cellular pathways through the formation of macromolecular assemblies[5,6]. Importantly, IDPs and IDRs exhibit a capacity for multivalency due to the presence of two or more interacting regions which may be repeats (2 residues), short-linear motifs (3–10 residues) or molecular recognition features (10–40 residues). Recently, IDPs and IDRs have received increased attention due to their role in the formation and maintenance of biomolecular condensates[7–9], which are concentrated assemblies of cellular macromolecules (e.g. proteins, nucleic acids).

Biomolecular condensates, often referred to as "membraneless organelles", are believed to form in many instances via liquid-liquid phase separation (LLPS)[10–12], a demixing process which leads to the formation of a "dense" phase with high macromolecular concentration and a surrounding "dilute" phase. It has been suggested that LLPS may serve as general mechanism for organization of cellular macromolecules to regulate biochemical processes. Formation of "liquid-like" organelles involves a complex interplay between enthalpic and entropic processes.[13] Multivalency of macromolecules plays a critical role in condensate formation, stability and regulation of its material properties.[14] In this regard, multi-domain proteins containing one or more folded domains along with linker and disordered regions can be considered as ideal candidates for modulating LLPS. Not surprisingly, such proteins are widely prevalent in cellular condensates such as germ granules[15], nucleoli[16], heterochromatin[17], stress granules[18], Balbiani bodies[19], Cajal bodies[20] and PML nuclear bodies.[21,22] Moreover, functional dysregulation of these condensates due to cellular stress and aging[23] are implicated in cancer[24,25] and neurodegenerative diseases.[26–28] It is therefore critical to understand the general principles governing the phase separation of disorder-containing, multi-domain proteins and how the modulation of underlying inter-domain interactions may affect the formation, structure and stability of condensates.

Molecular-level interactions implicated in protein phase separation

Both folded and disordered regions can influence the propensity for condensate formation through their involvement in different types of domain-level interactions with varying degree of specificity (Fig. 1). The binding affinity (k_D) of these interactions can be high (<1 μM), moderate (1–100 μM) or weak (>100 μM). Several excellent reviews have previously summarized the individual roles of folded and disordered regions in driving condensate formation.[29–31] The molecular-level interactions driving the phase separation of disordered low complexity domains (LCDs) and function in biomolecular condensates have been extensively studied and summarized in previous articles.[32,33] LCDs may be enriched in charged residues (e.g. DDX4, LAF1-RGG). LCDs enriched in aromatic and polar residues are referred to as “prion-like” due to their similarity to the yeast prion protein (e.g. FUS, hnRNPA1/2 and TDP-43). LCD condensates are weakly stabilized by interactions such as salt-bridges, cation-**π**, **π**-**π**, sp^2-**π**, hydrogen bonds and hydrophobic interactions.

Overall, both specific and non-specific interactions are influenced through stress (changes in pH, salt and temperature), post-translational modifications and interactions with other macromolecules such as nucleic acids and regulatory proteins (e.g., Ubiquitin, SUMO). Recent studies have attempted to address the interplay between folded and disordered regions in mediating LLPS of multi-domain proteins.[8,34–37] In the subsequent sections, we describe various domain-level strategies which underlie the phase behavior of multi-domain proteins and their implications for the design of biomolecular condensates with novel properties. For discussion on the influence of inter-domain interactions on the material and structural properties of biomolecular condensates, we refer the reader to previous review articles.[38,39] To conclude this perspective, we discuss the need for improved coarse-grained computational models to study the properties of multi-domain protein condensates at various length and timescales, and complement experimental efforts.

Homooligomerization domains can act as potent enhancers of phase separation and influence condensate structure

Stereospecific interactions between identical domains leads to homo-oligomerization, giving rise to dimeric and higher-order multimers. Both folded and disordered domains can form homo-oligomers with moderate to high affinity, that greatly enhance the phase separation propensity (Fig. 2). Examples of folded domains that have been shown to undergo homo-oligomerization and form higher-order oligomers are TDP-43 N-terminal domain (NTD) and Ph sterile alpha motif (SAM) domain. Both these domains can form head-to-tail polymeric structures[40] in a concentration-dependent manner as suggested by X-ray crystallography[41,42] and NMR spectroscopy[43]. NTD oligomerization (k_D~2 μM)[44] strongly promoted phase separation of both full-length TDP-43[43] and NTD-RRM1/2[36] construct wherein the disordered, C-terminal domain (CTD) was deleted. NTD-mediated phase separation was shown to be driven by electrostatic interactions and sensitive to salt concentration[36]. Under *in vivo* conditions, TDP-43 NTD-mediated oligomerization was shown to be critical for the formation of membraneless organelles with Hsp70 chaperones called “Anisosomes”[45] which protects RNA-free TDP-43 from pathological aggregation.

Similar to TDP-43 NTD, Ph SAM polymerization was also shown to enhance the formation of mini Ph-DNA condensates[46].

Homo-oligomerization can also occur through IDR regions which exhibit the formation of transient helice as in the case of TDP-43 CTD (aa:321–342) and UBQLN2 STI1-II region (aa:379–462). TDP-43 CTD populates a transient helical structure[47,48] which drives the formation of dimers ($k_D \approx 100$ μM) and higher-order oligomers. Deletion of the conserved helical region led to a ~two-fold increase in c_{sat} of TDP-43 full-length in vitro[49] and adversely affected its ability to bind certain mRNA transcripts and perform autoregulation. Analogously, STI1-II region (aa:379–462) of proteasomal shuttle factor - UBQLN2, drives its dimerization[50] through hydrophobic interactions and deletion of this region abolished LLPS of UBQLN2.

The above-mentioned examples illustrate that oligomerization domains can make substantial contributions to the overall phase separation propensity. It was shown that substitution of TDP-43 NTD for other oligomerization domains (e.g. SOD1, Transthyretin)[36] could also enhance the phase separation of RRM1/2 *in vitro*. Interestingly, Transthyretin which forms a stable tetramer, showed a noticeably greater reduction in c_{sat} compared to SOD1 which formed dimers with similar k_D, implying that the degree of oligomerization (i.e. multivalency) dictates the extent of increase in LLPS capacity. Alternatively, homodimerization domains may increase multivalency in multi-component assemblies such as SGs through additional, heterotypic interactions as observed in the case of NTF2 dimerization domain of the RNA-binding protein - G3BP[51]. These observations establish a direct, positive relationship between the multivalency of the homooligomerization domain and LLPS enhancement. Similarly, co-phase separation of multi-domain protein pairs with multiple heteroligomerization domains (e.g. poly-SUMO/polySIM, polySH3/polyPRM) exhibit an increase in LLPS with increasing multivalency.

Overall, the above examples illustrate that homooligomerization domains may be utilized in conjunction with IDRs to drive robust LLPS at physiological concentrations and maintain interaction specificity within the condensate. Future studies aimed at uncovering the effect of positional variation of the oligomerization domain (central versus terminal location) and the effect of surface properties of monomeric domains (charge, hydrophobicity), may allow for a better understanding of the molecular forces which promote the formation and stabilization of multi-domain protein condensates.

Charge-rich domains can influence phase behavior through non-specific interactions

RNA-binding proteins (RBPs)[52] play an essential role in transcription, RNA processing and stabilization. Several RBPs are associated with the formation and maintenance of liquid-like, membraneless compartments called stress granules (SGs)[53]. These proteins comprise of RNA-binding (RBD) and disordered, prion-like domains (PLDs). Several RBPs exhibit a synergistic interplay between RBD and PLD in the context of phase separation. Isolated PLDs undergo LLPS at or above physiological salt concentrations (>150 mM) and is primarily driven by weak interactions involving polar and aromatic residues.[54–56]

In contrast, LLPS of full-length RBPs including FUS, hnRNPA1 and TDP-43 are favored at low salt (<75 mM)[43,57,58] which implies a dominant role for electrostatic interactions in driving phase separation. The reversal of salt-dependence is linked to the enrichment of charged residues in RBDs.

RBPs belonging to the FET family (FUS, EWSR1 and TAF15)[59] possess a C-terminal RBD with disordered regions enriched in arginine motifs (RG/RGG) interspersed between the RRM and a zinc-finger fold. It was observed that FET proteins underwent robust LLPS at physiological concentrations while RBPs without RGG regions (e.g. hnRNPA1, TDP-43) failed to undergo LLPS under the same conditions without the addition of a crowding agent[34]. Isolated FUS PLD but not RBD, underwent LLPS at significantly higher concentrations (c_{sat} > 50-fold). Moreover, mutation of all tyrosine residues to serine (PLD) or arginine to glycine (RBD) led to a 15-fold increase in c_{sat}. These observations established a key role for arginine/tyrosine-mediated interactions in the phase separation of FET proteins. Altogether, the disordered RGG regions in FET proteins act synergistically with the PLD through non-specific, interactions to drive phase separation under physiological conditions.

Folded RNA recognition motifs (RRMs) within RBDs contain a substantial number of charged residues, some of which directly engage in binding RNA[60]. Recently[35], it was suggested based on biophysical experiments and coarse-grained simulations that electrostatic interactions between RRMs and PLD could explain the phase separation of hnRNPA1 under low-salt conditions. Surface electrostatic potential analysis of hnRNPA1 RRMs indicates the presence of distinct, oppositely-charged surfaces which suggests the possibility of encounter complex formation through these complementary surfaces[35]. The role of long-range electrostatic interactions in driving the formation of non-specific encounter complexes[61–64] during protein-protein association is well-established for folded proteins. Moreover, repulsive interactions between like-charged surfaces of RRMs could prevent the establishment of long-range structural order within condensates and promote “liquid-like” dynamics. Although *in vitro* experiments indicate a weak propensity for isolated RRMs to undergo LLPS, rapid aggregation of the TDP-43 NTD and its linker region was observed in the absence of RRMs[36], pointing towards a role for electrostatic interactions between RRMs in influencing condensate formation. Overall, in the absence of RGG-rich disordered regions, non-specific interactions between charged RRMs can potentially exert an influence on the formation and material properties of RBP condensates.

Domain architecture and inter-domain interactions dictate phase separation propensity, function and regulation

Studies of Ubiquitin-binding shuttle proteins[65] implicated in protein quality control (PQC) - UBQLN2, hHR23B and p62 elucidate how the LLPS propensity of multi-domain proteins is modulated by an intricate network of inter-domain interactions. Shuttle proteins possess at least two types of folded domains: UBL (Ubiquitin-like) and UBA (Ubiquitin-associated) domains along with interspersed, disordered regions. UBA domains can participate in weak interactions with UBL (*cis* and *trans*) and polyubiquitin chains. In addition, interactions

between disordered regions and those between folded and disordered regions are also observed. Overall, these interactions can either enhance or reduce the phase separation propensity of shuttle proteins within the context of a specific multi-domain architecture.

UBQLN2 (624 aa) is a shuttle factor responsible for the trafficking of ubiquitinated proteins from SGs to the 26S proteasome[66]. UBQLN2 comprises of an N-terminal UBL and a C-terminal UBA domain which flank a disordered, low complexity region (aa:109–576). The IDR region comprises of two hydrophobic, STI1-like regions which are enriched in hydrophobic residues, and a proline-rich (PXX) region (aa:491–538). Phase separation assays with truncated constructs[50] determined that STI1-II region (aa:379–462) which drives the dimerization of UBQLN2, is essential for phase separation. The presence of UBL increased c_{sat} while UBA decreased c_{sat}, indicating opposite effects of the two domains with regard to LLPS. UBL and UBA can interact weakly ($k_D \approx 175$ μM) and upon addition of UBL domain in *trans* to UBQLN2-ΔUBL, the LLPS propensity decreases. These observations established that UBL-UBA interactions inhibited LLPS of UBQLN2. Consistent with these observations, the addition of either ubiquitin ($k_D \approx 5$ μM) or K48-linked polyubiquitin chains dissolved UBQLN2 droplets *in vitro* and strongly inhibited LLPS.[50]

NMR titration experiments with isolated fragments uncovered an underlying hierarchy of inter-domain interactions and provide deeper insights into UBQLN2's phase separation propensity.[37,50] UBA was found to enhance LLPS through weak interactions with two IDR regions - STI1-II and Pxx. In addition to UBA, UBL also interacts with STI1-I (aa:176–247) and connecting regions between STI1-I/II (aa:248–378) and Pxx-UBA (aa:555–570), all of which inhibit LLPS and are stronger than UBA-IDR interactions. To conclude, UBL engages in inhibitory interactions with UBA and IDR regions, which counteracts the ability of UBA domain to promote LLPS. In stark contrast to UBQLN2, the shuttle factors - p62 and hHR23B, undergo LLPS only in the presence of polyubiquitin chains. For a detailed discussion related to the effect of polyubiquitin chain linkage and topology on co-phase separation with shuttle factors, we refer the reader to the review by Dao and Castenada.[65]

p62 is associated with the formation of cytoplasmic foci which may serve as precursors of autophagosomes.[67] p62 possesses an N-terminal UBL-like domain (PBX1) which can undergo homo-oligomerization to form a filamentous scaffold required for LLPS.[68] The C-terminal UBA domain acts synergistically with PBX1 and freely engages with polyubiquitin chains which act as LLPS-promoting, multivalent scaffolds. The intrinsically disordered region (aa:246–300) of p62 can also drive phase separation without PBX1 or polyubiquitin, through its interactions with the histone chaperone DAXX.[69] Thus the domain architecture of p62 encodes for alternate mechanisms of co-phase separation which are dependent on the presence of a suitable interaction partner and exploits either folded or disordered domains.

hHR23B regulates the formation of proteasomal foci in the nucleus through its interactions with polyubiquitin chains and proteasomal receptors.[70] In the absence of polyubiquitin chains, hHR23B forms a dimer[71] which is stabilized by Ubl-UBA1/2 interactions and cannot undergo LLPS.[70] Deletion of UBL domain in the presence of polyubiquitin however, leads to aggregation indicating that LLPS-inhibitory interactions between UBL and UBA domains

help to maintain “liquid-like” characteristics. Overall, hHR23B appears to utilize inhibitory UBL-UBA interactions in a manner similar to UBQLN2 to regulate its LLPS propensity.

Taken together, it is evident from LLPS studies of shuttle factors that pre-encoded inhibitory interactions (*cis and trans*) can tune the LLPS of multi-domain proteins in isolation or in the presence of suitable binding partners.

Physical characteristics of linkers influence phase separation ability

In addition to IDRs, multi-domain proteins may also possess short, disordered segments (10–50 aa) termed as linkers. Linkers between the folded domains impart varying degrees of flexibility, a characteristic that could dictate how folded domains interact with each other. Although it is common to treat linkers as passive tethers, the physical nature of the linkers (*i.e.*, its length and sequence composition) has the potential to influence the phase separation in conjunction with its impact on binding affinity and avidity.[72]

Notably, a charge-segregated 50 residue linker connecting the first two SH3 (SRC Homology 3) domains in the adaptor protein - Nck, has been shown to enhance the phase separation of Nck/N-WASP and p-nephrin/Nck/N-WASP complexes.[73] The linker promotes phase separation by enhancing the ability of the Nck to self-associate through electrostatic interactions with the acidic, second SH3 domain. Deletion of the linker or mutations that changed the predominantly basic character of the N-terminal or the highly conserved, central KVKRK motif inhibited phase separation. Interestingly, when this linker was used instead of $(GGS)_4$ linker in the $SUMO_5$-SIM_5 complex, it further enhanced the phase separation propensity of this system. This linker was also observed to bind the GBD (GTPase binding domain) in N-WASP that directly impacts actin assembly on cellular membranes.[74]

System-spanning reversible physical cross-links are common in protein droplets, the state at which the droplets are referred to as gels rather than liquids.[75–77] But, gel-formation (*i.e.*, gelation) can occur without LLPS.[75,78,79] The length and sequence of linkers can dictate whether multi-domain proteins prefer gelation without phase separation or that driven by phase separation. Out of 226 unique linker regions identified among 100 linear multivalent proteins from the non-redundant human proteome, simulations demonstrated that 38% of the linkers behaved like a Flory random coil (FRC) while 30% of the linkers that behaved like a self-avoiding random coil (SARC).[80] Interestingly, gelation driven by phase separation occurred for a mixture of poly-SH3 and poly-PRM proteins with FRC linkers. In contrast, SARC linkers suppressed phase separation and gelation could only occur at high protein concentrations (i.e., gelation without phase separation). This suggests that proteins with tandem repeat domains can be used to design condensates by modifying the properties of linker regions. Furthermore, simulations suggest that “sticky” inter-linker interactions govern coil-to-globule transitions of multivalent proteins within clusters of metastable droplets, resulting in increased cluster density.[81] Such interactions along with linker flexibility, also dictate the cluster/droplet growth by modulating chain reorganization times within them, indicating that linker properties can be tuned to achieve varied phase separation propensities.

Inter-domain and domain-nucleic acid interactions in the formation and regulation of protein-nucleic acid condensates

Protein-nucleic acid interactions can modulate the phase behavior of multi-domain proteins *in vitro* and are essential for the formation and integrity of various ribonucleoprotein (RNP) granules.[18,82–84] These interactions also influence the formation of heterochromatin and transcriptional condensates which control gene expression.[17,85] While most nucleic-acid binding proteins either interact with RNA (RBPs) or DNA (DBPs), several of these (~400) bind to both types of nucleic acids.[86] The various types of protein-nucleic acid interactions include hydrogen bonding, π stacking and electrostatic interactions with the sugar-phosphate backbone.

RBPs may contain one or two RRMs[60,87–89] (e.g. FUS, hnRNPA1/2 and TDP-43) which adopt a characteristic fold and recognize RNA sequences with high affinity (k_D≈nM-μM) and sequence specificity (e.g. UG/AG-rich). RBPs also utilize disordered, RGG/RG motifs[90] which engage in non-specific interactions with the RNA backbone to increase the overall binding affinity. Phase-separating DBPs can bind either double-stranded (ds) or single-stranded (ss) genomic DNA through a variety of folded domains[91] or disordered regions[92] enriched in positively-charged residues (Arg/Lys-rich). The co-phase separation propensity and/or condensate properties of multi-domain, nucleic acid-binding proteins can be effectively modulated by factors such as (i) protein to nucleic acid stoichiometry, (ii) length and composition of the nucleic acid strand, and (iii) disease mutations and/or post-translational modifications (PTMs) within folded and disordered regions.[93]

In mammalian cells, RNA serves as a buffer to modulate the condensation of RBPs.[94] A high RNA-to-protein concentration ratio in the nucleus was shown to maintain RBP solubility, while a lower ratio in the cytoplasm was shown to promote condensate formation. Based on NMR experiments and cell-based assays, a possible mechanism by which RNA could promote RBP solubility was proposed in the case of TDP-43.[95] It was observed that long GU-rich RNA repeats promote cooperative binding of TDP-43 through intermolecular interactions between its tandem RRMs. Specifically, an RRM1/2 intermolecular interface was identified which minimized NTD/CTD-mediated interactions responsible for condensation and subsequent aggregation in the absence of RNA and Hsp70 chaperones.[45] *In vitro* analysis of FUS-RNA interactions revealed that single ALS/FTD mutants at Arginine (R) and Glycine (G) positions cause significant differences in droplet properties compared to FUS wild-type.[96] R mutants near the prion-like LCD and RGG regions led to significantly larger droplets with reduced dynamics. Correspondingly, a reduction in interaction dynamics of R mutants with RNA was detected in smFRET experiments. In contrast, G mutations in prion-like LCD and RGG regions led to rapid aging of droplets despite exhibiting similar interaction dynamics with RNA as wild-type FUS in smFRET experiments. These observations highlight that ALS/FTD mutants which occur in disordered regions can significantly perturb the dynamics of RBP-RNA condensates and lead to pathological aggregates implicated in neurodegeneration. RBP-RNA condensates can also be disrupted by PTMs such as tyrosine phosphorylation in prion-like LCDs[97,98] or lysine acetylation in RRMs.[99]

The α-isoform of human Heterochromatin Protein 1 (HP1α) promotes the formation of heterochromatin domains (DNA compaction) through co-phase separation with chromatin.[100,101] DNA or chromatin-binding and co-phase separation of HP1α is critically dependent on electrostatic interactions via a basic hinge (lysine-rich linker) region[102] between an N-terminal chromodomain (CD) and a C-terminal chromoshadow domain (CSD). The critical concentration of HP1α (~50 μM for 147 bp DNA) required for co-phase separation was largely invariant of DNA concentration (0.125 to 4 μM) and condensates were observed even at HP1α to DNA ratios of 5000:1.[103] Notably, the critical concentration of HP1α reduced by more than an order of magnitude (~3 μM) for a longer DNA molecule (2.7 kbp) due to its higher valency and lies within the physiological range (1–10 μM). The interactions mediated through the disordered N-terminal extension (NTE) lowered the critical concentration of HP1α required for co-phase separation with longer DNA (~50 kbp), likely by promoting higher order oligomerization.[103]

Furthermore, phosphorylation of NTE weakens DNA binding[104] and co-phase separation[105], likely due to electrostatic repulsion. In contrast, the C-terminal extension (CTE) which can bind to the basic hinge[101], increases the critical concentration required for co-phase separation through competition with DNA.[103] The disruption of CSD-mediated dimerization also inhibited DNA-driven LLPS *in vitro*.[101] In the case of *Drosophila* HP1α homolog (HP1a), disruption of CSD dimerization (I191E mutation) led to faster dynamics within phase-separated heterochromatin domains.[100] Overall, a network of intra/inter-molecular interactions involving the disordered NTE, hinge and CTE regions together with CSD-mediated dimerization regulate the co-phase separation propensity and material properties of HP1α-DNA condensates.

Post-translational modifications of histone proteins may exert a substantial influence on heterochromatin formation by directly tuning the interaction affinity between chromatin and multi-domain regulatory partners. For example, it was observed that both wild type and NTE-phosphorylated HP1α show a significantly higher LLPS capacity with histone 3 tri-methylated (H3K9me3) chromatin[105] compared to unmodified chromatin. From a structural viewpoint, LLPS is enhanced due to higher binding affinity of H3K9me3 by the N-terminal CD domain of HP1α[106]. However, the LLPS capacity of HP1α in the presence of H3K9me3 chromatin is much lower compared to DNA. These observations highlight a complex interplay between histone-regulator and DNA-regulator interactions which underlie the formation of heterochromatin.

The Polycomb Group (PcG) protein - Ph, which constitutes a subunit of Polycomb Repressive Complex 1 (PRC1) is implicated in silencing gene expression through large-scale chromatin organization.[107] A mini Ph construct from *Drosophila melanogaster* which contained the N-terminal Helicase domain 1 (HD1), the Phe-Cys-Ser (FCS) zinc-finger domain, and the C-terminal Sterile Alpha Motif (SAM) formed phase-separated condensates in the presence of chromatin or DNA.[46] Deletion of either SAM or HD1/FCS led to abrogation of phase separation, indicating that all three domains are required for co-phase separation. Mass spectrometry-footprinting experiments identified potential lysine residues across all three domains that bind DNA and fully-acetylated mini Ph showed complete loss of DNA binding. As mentioned earlier, SAM polymerization was shown to enhance

mini Ph-DNA phase separation and this occurred through an increase in binding affinity towards DNA.[46] In addition, SAM polymerization also led to slower dynamics of mini Ph and chromatin in the condensed phase. The acidic linker connecting FCS and SAM has been shown to limit SAM polymerization through possible linker-SAM interactions[108] (*in trans*) and may therefore exert a negative influence of Ph-DNA condensation. In conclusion, all three folded domains act cooperatively to facilitate the formation of mini Ph-DNA condensates while SAM-linker interactions can modulate the properties of the condensed phase.

Exploiting folded domains and disordered regions to engineer multi-domain proteins with desirable condensed phase properties

Folded domains and disordered regions which are known to drive the LLPS of multi-domain proteins can be harnessed as modules to design multivalent constructs which undergo LLPS under desirable conditions (e.g. temperature, pH and salt). Further, a combination of phase-separating modules with other types of protein interaction motifs (e.g. protease cleavage sites) or domains (e.g. RNA-recognition, cellulose-binding) can be used to generate condensates with tunable stability and composition. These artificial condensates or organelles can be utilized as microreactors to control the rate of biochemical reactions. For a detailed explanation of how multi-domain protein scaffolds may be exploited to engineer synthetic membraneless organelles, we refer the reader to the review by Bracha et al.[38]

An intuitive approach to designing synthetic condensates is to create a multivalent platform based on tandem repeats of a folded domain or disordered motif as demonstrated in the case of a poly-SH3 construct which phase separates in the presence of a poly-PRM (proline-rich motif) ligand.[72] Interestingly, fusion of FUS prion-like domain to poly-SH3 was shown to lower c_{sat} of the poly-SH3/poly-PRM system by nearly 8-fold.[109] The modulation of poly-SH3/poly-PRM phase separation was dependent on tyrosine residues of FUS LCD and their mutation to non-aromatic residues or phosphorylation of all tyrosines increased c_{sat} compared to wild-type FUS. These observations elucidate how varying the aromatic content of prion-like LCDs attached to folded domains either through mutations or PTMs, can exert a strong influence on the phase separation propensity of engineered multi-domain constructs.

Tandem repeats based on the low complexity, RGG region (aa:1–170, Fig. 3A) of *C. elegans* P-granule RNA helicase, LAF-1 were shown to form droplets with elevated thermal stability and could be genetically engineered to achieve controlled-assembly, and cargo transport.[110] The transition temperatures for $(RGG)_2$ and $(RGG)_3$ were 40 °C and 50 °C compared to the RGG construct which formed droplets only below 15 °C. These differences in transition temperatures allowed for the regulation of droplet assembly in an enzymatically-controlled fashion through the introduction of TEV protease cleavage sites between RGG domains. Further, the composition of the droplets could be regulated through (i) direct attachment of cargo domains to the tandem RGG construct with an additional cleavage site (Fig. 3B), or (ii) via attachment of cargo-recruiting interaction motifs (SYNZIP1/2) to tandem RGG and cargo constructs individually (Fig. 3C). These synthetic, tandem RGG organelles

were shown to be functional in *Xenopus* egg extracts, protocells and in mammalian HEK293 cells. Recently, elastin-like polypeptides (artificial IDRs) which are known to phase-separate[111], were fused to RNA-binding RGG region of the P-granule protein PGL-1 to create synthetic RNP granules.[112] These granules could bind and sequester mRNA within microdroplet-based protocells, thereby suppressing translation. The structure of $(RGG)_2$ condensates could be be altered when mixed with amphiphilic proteins which contain a non-phase separating protein attached to an RGG domain[113]. Mixing $(RGG)_2$ with an MBP-based amphiphilic protein lead to the formation of enveloped condensates containing MBP (Fig. 3D) at the surface of the droplet and RGG in the core. Interestingly, MBP-based amphiphiles acted as surfactants and affected droplet size, with increasing concentrations resulting in smaller droplet radii. In contrast, mixing $(RGG)_2$ with GST-based amphiphiles resulted in coassembly and formed multiphasic structures.

In combination with marine mussel foot protein - Mfp5, LLPS of TDP-43 prion-like LCD (TLCD) and its subsequent liquid-to-solid maturation was utilized to design ultra-strong, underwater adhesives.[114] Mfp5 is also disordered, enriched in lysine/DOPA residues and derives its adhesive properties through LLPS driven by electrostatic and hydrophobic interactions. The low temperature-induced coatings of Mfp5-TLCD driven by LLPS-dependent substrate wetting followed by concentration into uniform, amyloid nanofibers exhibited strong adhesiveness over a wide range of pH and salt concentrations. Alanine-rich, disordered repeats of *Araneus diadematus* spidroin which undergo LLPS were fused to cellulose-binding module (CBM) domains on either side and mixed with cellulose nanofibrils (CNFs) to form protein-cellulose composites with enhanced strength, stiffness and toughness.[115] Overall, both Mfp5-TLCD and spidroin-$(CBM)_2$-CNF constructs demonstrate how the LLPS properties of low-complexity sequences can be exploited in combination with other types of unique polypeptides and macromolecules to design biomimetic materials with improved physico-chemical properties.

Computational modeling approaches to study the phase separation of multi-domain proteins

Evidently, the formation, structuring, and dissolution of biomolecular condensates is a tightly regulated process. It is quintessential to unravel the sequence-dependent molecular driving forces involved in the formation and stabilization of such condensates to be able to regulate their function as needed. However, due to the vast chemical phase space, uncovering such sequence-to-function relationships necessitates a computationally-driven approach.[116] Molecular simulations provide the ability to investigate how the cooperativity between multiple domains affects the phase behavior of proteins (Fig. 4).

Protein simulations with fully atomistic details[54,117,118] and coarse-grained (CG) models (e.g., Martini model[119] with 2–4 heavy atoms per amino acid) provide chemically detailed representations at the single-chain level but they are not computationally efficient to simulate large assemblies such as multi-domain protein condensates.[35,120–122] To simulate large-scale assemblies of multi-domain proteins at a reasonable computational cost, CG models that utilize a coarse representation of amino acids as a single bead offer an

efficient route and have been successfully used to uncover the sequence determinants of phase behavior for IDPs.[123–126] A detailed description of the residue-specific, HPS model and coexistence sampling technique developed in our group for studying the phase behavior of IDPs using molecular dynamics (MD) simulations is provided elsewhere.[127] The HPS model was also extended to assess the effect of post-translational modifications such as phosphorylation and acetylation on LLPS.[128] Furthermore, we also developed a nucleotide-specific, CG model for RNA which was utilized to study LAF-1 RGG/RNA interactions during co-phase separation and structure formation in the condensed phase.[129] Despite efforts in designing residue-level CG models for disordered and folded[130] proteins separately, there are two main challenges that arise in CG simulations of multi-domain proteins: (i) accurate modeling of inter-domain interactions between folded/disordered regions and (ii) lack of secondary structure potentials that allow for the accurate simulation of both globular proteins and flexible IDPs.

Numerous groups have leveraged different CG simulation approaches for proteins with both folded[130] and disordered domains thus far. One approach to simulate multi-domain proteins uses rigid body dynamics by constraining the structure of folded domain as a rigid body.[48,123] Alternatively, the folded domains can also be represented with the Go-like force field[131], allowing increased conformational flexibility compared to rigid constraints. Even though both methods require an experimental structure for folded domains, the usage of rigid body dynamics renders computational efficiency. A new strategy to generalize the CG model for both folded and disordered proteins has been introduced by the Zhang group[132] using a combination of maximum entropy optimization and energy gap constraint to capture the conformational dynamics of both folded and disordered proteins while maintaining the structural stability of the former. However, this force field relies on a protein-specific and non-transferable secondary structure potential. More accurate secondary structure potentials are required for folded domains as well as IDPs which may also exhibit transient secondary structure. Therefore, there is a need to develop transferable CG model(s) that provide a reasonable description of both inter-domain interactions and structural features without the need for system-specific, experimental information.[133] Such transferable models would greatly benefit from rigorous benchmarks performed against biophysical experiments[134,135] (e.g. NMR, SAXS and FRET) which provide valuable information regarding the inter-domain dynamics and interactions of multi-domain proteins.

Challenges and future outlook

Over the last decade, LLPS has emerged as a principal mechanism for the organization of cellular biochemistry and response to changing environmental conditions. Of considerable importance is the emerging link between LLPS and pathological aggregation[136] of various proteins implicated in cancer and neurodegenerative diseases, which provides numerous opportunities for the development of novel therapeutics. Therefore, it is critical to achieve a deeper understanding of the intramolecular conformational transitions which are associated with droplet formation and liquid-to-solid transitions which lead to the formation of aggregates.[137]

As evident from the examples discussed in this perspective, the formation and regulation of multi-domain protein condensates often involves a complex interplay among the constituent domains which may either be synergistic or antagonistic towards LLPS. The stereospecificity and affinity of inter-domain interactions exhibit a wide range of variation and can be tuned through both post-translational modifications and heterotypic interactions (e.g. ubiquitin, nucleic acids). In this regard, biophysical experiments aimed at uncovering the network of inter-domain interactions and the relative contributions of individual domains towards phase separation[37] provide much-needed insights and serve as a foundation for the development of accurate computational models to study LLPS. Furthermore, open questions remain regarding how domain architecture and inter-domain interactions may influence the structural organization and material properties of multi-domain protein condensates.

Future efforts which aim to uncover the underlying mechanisms of multi-domain protein LLPS would benefit greatly from the availability of transferable, CG computational models that afford an accurate yet cost-effective description of structure and dynamics of inter-domain interactions. As evident from studies of IDP/IDRs, a closer synergy between such computational models and biophysical experiments can provide rich and complementary insights into the formation, structure and dynamics of multi-domain protein condensates and their regulation. Such insights have the potential to inspire the utilization of folded and disordered domains in unique ways to design synthetic condensates with desired assembly and material properties.

Acknowledgments

We acknowledge countless discussions with experimental collaborators (Profs. Nicolas Fawzi, Young Chan Kim, Galia Debelouchina, Nicole Francis, Benjamin Schuster, Matthew Good, Kristi Kiick, Sapun Parekh, Alaji Bah) that have contributed to a better understanding of issues outlined in this perspective. Our work on biomolecular phase separation is supported by the National Institutes of Health grants R01NS116176, R01GM136917, and National Science Foundation grant DMR2004796.

References

(1). Levitt M (2009) Nature of the protein universe. Proc. Natl. Acad. Sci 106, 11079–11084. [PubMed: 19541617]

(2). Koonin EV, Wolf YI, and Karev GP (2002) The structure of the protein universe and genome evolution. Nature 420, 218–223. [PubMed: 12432406]

(3). Oldfield CJ, and Dunker AK (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. Annu. Rev. Biochem 83, 553–584. [PubMed: 24606139]

(4). van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, Fuxreiter M, Gough J, Gsponer J, Jones DT, Kim PM, Kriwacki RW, Oldfield CJ, Pappu RV, Tompa P, Uversky VN, Wright PE, and Babu MM (2014) Classification of Intrinsically Disordered Regions and Proteins. Chem. Rev 114, 6589–6631. [PubMed: 24773235]

(5). Fung HYJ, Birol M, and Rhoades E (2018) IDPs in macromolecular complexes: the roles of multivalent interactions in diverse assemblies. Curr. Opin. Struct. Biol 49, 36–43. [PubMed: 29306779]

(6). Latysheva NS, Flock T, Weatheritt RJ, Chavali S, and Babu MM (2015) How do disordered regions achieve comparable functions to structured domains? Protein Sci. 24, 909–922. [PubMed: 25752799]

(7). Uversky VN (2017) Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. Curr. Opin. Struct. Biol 44, 18–30. [PubMed: 27838525]

(8). Protter DSW, Rao BS, Van Treeck B, Lin Y, Mizoue L, Rosen MK, and Parker R (2018) Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. Cell Rep. 22, 1401–1412. [PubMed: 29425497]

(9). Abyzov A, Blackledge M, and Zweckstetter M (2022) Conformational Dynamics of Intrinsically Disordered Proteins Regulate Biomolecular Condensate Chemistry. Chem. Rev 122, 6719–6748. [PubMed: 35179885]

(10). Hyman AA, Weber CA, and Jülicher F (2014) Liquid-Liquid Phase Separation in Biology. Annu. Rev. Cell Dev. Biol 30, 39–58. [PubMed: 25288112]

(11). Mitrea DM, and Kriwacki RW (2016) Phase separation in biology; functional organization of a higher order. Cell Commun. Signal 14, 1. [PubMed: 26727894]

(12). Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Van Den Bosch L, Tompa P, and Fuxreiter M (2018) Protein Phase Separation: A New Phase in Cell Biology. Trends Cell Biol. 28, 420–435. [PubMed: 29602697]

(13). Brangwynne CP, Tompa P, and Pappu RV (2015) Polymer physics of intracellular phase transitions. Nat. Phys 11, 899–904.

(14). Banani SF, Lee HO, Hyman AA, and Rosen MK (2017) Biomolecular condensates: organizers of cellular biochemistry. Nat. Rev. Mol. Cell Biol 18, 285–298. [PubMed: 28225081]

(15). Brangwynne CP, Eckmann CR, Courson DS, Rybarska A, Hoege C, Gharakhani J, Jülicher F, and Hyman AA (2009) Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. Science (80-.). 324, 1729–1732.

(16). Brangwynne CP, Mitchison TJ, and Hyman AA (2011) Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc. Natl. Acad. Sci 108, 4334–4339. [PubMed: 21368180]

(17). Larson AG, and Narlikar GJ (2018) The Role of Phase Separation in Heterochromatin Formation, Function, and Regulation. Biochemistry 57, 2540–2548. [PubMed: 29644850]

(18). Protter DSW, and Parker R (2016) Principles and Properties of Stress Granules. Trends Cell Biol. 26, 668–679. [PubMed: 27289443]

(19). Boke E, Ruer M, Wühr M, Coughlin M, Lemaitre R, Gygi SP, Alberti S, Drechsel D, Hyman AA, and Mitchison TJ (2016) Amyloid-like Self-Assembly of a Cellular Compartment. Cell 166, 637–650. [PubMed: 27471966]

(20). Machyna M, Heyn P, and Neugebauer KM (2013) Cajal bodies: Where form meets function. Wiley Interdiscip. Rev. RNA 4, 17–34. [PubMed: 23042601]

(21). Bernardi R, and Pandolfi PP (2007) Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat. Rev. Mol. Cell Biol 8, 1006–1016. [PubMed: 17928811]

(22). Lallemand-Breitenbach V, and de Thé H (2018) PML nuclear bodies: from architecture to function. Curr. Opin. Cell Biol 52, 154–161. [PubMed: 29723661]

(23). Alberti S, and Hyman AA (2021) Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. Nat. Rev. Mol. Cell Biol 22, 196–213. [PubMed: 33510441]

(24). Boija A, Klein IA, and Young RA (2021) Biomolecular Condensates and Cancer. Cancer Cell 39, 174–192. [PubMed: 33417833]

(25). Taniue K, and Akimitsu N (2022) Aberrant phase separation and cancer. FEBS J. 289, 17–39. [PubMed: 33583140]

(26). Elbaum-Garfinkle S (2019) Matter over mind: Liquid phase separation and neurodegeneration. J. Biol. Chem 294, 7160–7168. [PubMed: 30914480]

(27). Babinchak WM, Haider R, Dumm BK, Sarkar P, Surewicz K, Choi J-K, and Surewicz WK (2019) The role of liquid–liquid phase separation in aggregation of the TDP-43 low-complexity domain. J. Biol. Chem 294, 6306–6317. [PubMed: 30814253]

(28). Zbinden A, Pérez-Berlanga M, De Rossi P, and Polymenidou M (2020) Phase Separation and Neurodegenerative Diseases: A Disturbance in the Force. Dev. Cell 55, 45–68. [PubMed: 33049211]

(29). Mittag T, and Parker R (2018) Multiple Modes of Protein–Protein Interactions Promote RNP Granule Assembly. J. Mol. Biol 430, 4636–4649. [PubMed: 30099026]

(30). Gomes E, and Shorter J (2019) The molecular language of membraneless organelles. J. Biol. Chem 294, 7115–7127. [PubMed: 30045872]

(31). Dignon GL, Best RB, and Mittal J (2020) Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. Annu. Rev. Phys. Chem 71, 53–75. [PubMed: 32312191]

(32). Martin EW, and Mittag T (2018) Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. Biochemistry 57, 2478–2487. [PubMed: 29517898]

(33). Franzmann TM, and Alberti S (2019) Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. J. Biol. Chem 294, 7128–7136. [PubMed: 29921587]

(34). Wang J, Choi JM, Holehouse AS, Lee HO, Zhang X, Jahnel M, Maharana S, Lemaitre R, Pozniakovsky A, Drechsel D, Poser I, Pappu RV, Alberti S, and Hyman AA (2018) A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell 174, 688–699.e16. [PubMed: 29961577]

(35). Martin EW, Thomasen FE, Milkovic NM, Cuneo MJ, Grace CR, Nourse A, Lindorff-Larsen K, and Mittag T (2021) Interplay of folded domains and the disordered low-complexity domain in mediating hnRNPA1 phase separation. Nucleic Acids Res. 49, 2931–2945. [PubMed: 33577679]

(36). Carter GC, Hsiung C-H, Simpson L, Yang H, and Zhang X (2021) N-terminal Domain of TDP43 Enhances Liquid-Liquid Phase Separation of Globular Proteins. J. Mol. Biol 433, 166948. [PubMed: 33744316]

(37). Zheng T, Galagedera SKK, and Castañeda CA (2021) Previously uncharacterized interactions between the folded and intrinsically disordered domains impart asymmetric effects on UBQLN2 phase separation. Protein Sci. 30, 1467–1481. [PubMed: 34029402]

(38). Bracha D, Walls MT, and Brangwynne CP (2019) Probing and engineering liquid-phase organelles. Nat. Biotechnol 37, 1435–1445. [PubMed: 31792412]

(39). Schuster BS, Regy RM, Dolan EM, Kanchi Ranganath A, Jovic N, Khare SD, Shi Z, and Mittal J (2021) Biomolecular Condensates: Sequence Determinants of Phase Separation, Microstructural Organization, Enzymatic Activity, and Material Properties. J. Phys. Chem. B 125, 3441–3451. [PubMed: 33661634]

(40). Bienz M (2020) Head-to-Tail Polymerization in the Assembly of Biomolecular Condensates. Cell 182, 799–811. [PubMed: 32822572]

(41). Kim CA, Gingery M, Pilpa RM, and Bowie JU (2002) The SAM domain of polyhomeotic forms a helical polymer. Nat. Struct. Biol 9, 453–7. [PubMed: 11992127]

(42). Afroz T, Hock E-M, Ernst P, Foglieni C, Jambeau M, Gilhespy LAB, Laferriere F, Maniecka Z, Plückthun A, Mittl P, Paganetti P, Allain FHT, and Polymenidou M (2017) Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. Nat. Commun 8, 45. [PubMed: 28663553]

(43). Wang A, Conicella AE, Schmidt HB, Martin EW, Rhoads SN, Reeb AN, Nourse A, Ramirez Montero D, Ryan VH, Rohatgi R, Shewmaker F, Naik MT, Mittag T, Ayala YM, and Fawzi NL (2018) A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. EMBO J. 37, 1–18. [PubMed: 29212815]

(44). Tsoi PS, Choi KJ, Leonard PG, Sizovs A, Moosa MM, MacKenzie KR, Ferreon JC, and Ferreon ACM (2017) The N-Terminal Domain of ALS-Linked TDP-43 Assembles without Misfolding. Angew. Chemie - Int. Ed 56, 12590–12593.

(45). Yu H, Lu S, Gasior K, Singh D, Vazquez-Sanchez S, Tapia O, Toprani D, Beccari MS, Yates JR, Da Cruz S, Newby JM, Lafarga M, Gladfelter AS, Villa E, and Cleveland DW (2021) HSP70 chaperones RNA-free TDP-43 into anisotropic intranuclear liquid spherical shells. Science (80-.). 371.

(46). Seif E, Kang JJ, Sasseville C, Senkovich O, Kaltashov A, Boulier EL, Kapur I, Kim CA, and Francis NJ (2020) Phase separation by the polyhomeotic sterile alpha motif compartmentalizes Polycomb Group proteins and enhances their activity. Nat. Commun 11, 1–19. [PubMed: 31911652]

(47). Conicella AE, Zerze GH, Mittal J, and Fawzi NL (2016) ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure 24, 1537–1549. [PubMed: 27545621]

(48). Conicella AE, Dignon GL, Zerze GH, Schmidt HB, D'Ordine AM, Kim YC, Rohatgi R, Ayala YM, Mittal J, and Fawzi NL (2020) TDP-43 α-helical structure tunes liquid–liquid phase separation and function. Proc. Natl. Acad. Sci 117, 5883–5894. [PubMed: 32132204]

(49). Hallegger M, Chakrabarti AM, Lee FCY, Lee BL, Amalietti AG, Odeh HM, Copley KE, Rubien JD, Portz B, Kuret K, Huppertz I, Rau F, Patani R, Fawzi NL, Shorter J, Luscombe NM, and Ule J (2021) TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. Cell 1–17. [PubMed: 33417857]

(50). Dao TP, Kolaitis RM, Kim HJ, O'Donovan K, Martyniak B, Colicino E, Hehnly H, Taylor JP, and Castañeda CA (2018) Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions. Mol. Cell 69, 965–978.e6. [PubMed: 29526694]

(51). Sanders DW, Kedersha N, Lee DSW, Strom AR, Drake V, Riback JA, Bracha D, Eeftens JM, Iwanicki A, Wang A, Wei M-T, Whitney G, Lyons SM, Anderson P, Jacobs WM, Ivanov P, and Brangwynne CP (2020) Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. Cell 181, 306–324.e28. [PubMed: 32302570]

(52). Harrison AF, and Shorter J (2017) RNA-binding proteins with prion-like domains in health and disease. Biochem. J 474, 1417–1438. [PubMed: 28389532]

(53). Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj AP, Kim HJ, Mittag T, and Taylor JP (2015) Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. Cell 163, 123–133. [PubMed: 26406374]

(54). Murthy AC, Dignon GL, Kan Y, Zerze GH, Parekh SH, Mittal J, and Fawzi NL (2019) Molecular interactions underlying liquid−liquid phase separation of the FUS low-complexity domain. Nat. Struct. Mol. Biol 26, 637–648. [PubMed: 31270472]

(55). Li HR, Chiang WC, Chou PC, Wang WJ, and Huang J. rong. (2018) TAR DNA-binding protein 43 (TDP-43) liquid-liquid phase separation is mediated by just a few aromatic residues. J. Biol. Chem 293, 6090–6098. [PubMed: 29511089]

(56). Martin EW, Holehouse AS, Peran I, Farag M, Incicco JJ, Bremer A, Grace CR, Soranno A, Pappu RV, and Mittag T (2020) Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. Science (80-.). 367, 694–699.

(57). Lin Y, Protter DSW, Rosen MK, and Parker R (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol. Cell 60, 208–219. [PubMed: 26412307]

(58). Krainer G, Welsh TJ, Joseph JA, Espinosa JR, Wittmann S, de Csilléry E, Sridhar A, Toprakcioglu Z, Gudiškytė G, Czekalska MA, Arter WE, Guillén-Boixet J, Franzmann TM, Qamar S, George-Hyslop PS, Hyman AA, Collepardo-Guevara R, Alberti S, and Knowles TPJ (2021) Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. Nat. Commun 12, 1085. [PubMed: 33597515]

(59). Schwartz JC, Cech TR, and Parker RR (2015) Biochemical Properties and Biological Functions of FET Proteins. Annu. Rev. Biochem 84, 355–379. [PubMed: 25494299]

(60). Lukavsky PJ, Daujotyte D, Tollervey JR, Ule J, Stuani C, Buratti E, Baralle FE, Damberger FF, and Allain FHT (2013) Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. Nat. Struct. Mol. Biol 20, 1443–1449. [PubMed: 24240615]

(61). Tang C, Iwahara J, and Clore GM (2006) Visualization of transient encounter complexes in protein–protein association. Nature 444, 383–386. [PubMed: 17051159]

(62). Kim YC, Tang C, Clore GM, and Hummer G (2008) Replica exchange simulations of transient encounter complexes in protein–protein association. Proc. Natl. Acad. Sci 105, 12855–12860. [PubMed: 18728193]

(63). Mohanty P, Agrata R, Habibullah BI, G S A, and Das R (2019) Deamidation disrupts native and transient contacts to weaken the interaction between UBC13 and RING-finger E3 ligases. Elife 8, 1–32.

(64). Pan AC, Jacobson D, Yatsenko K, Sritharan D, Weinreich TM, and Shaw DE (2019) Atomic-level characterization of protein–protein association. Proc. Natl. Acad. Sci 201815431.

(65). Dao TP, and Castañeda CA (2020) Ubiquitin-Modulated Phase Separation of Shuttle Proteins: Does Condensate Formation Promote Protein Degradation? BioEssays 42, 1–13.

(66). Zheng T, Yang Y, and Castañeda CA (2020) Structure, dynamics and functions of UBQLNs: At the crossroads of protein quality control machinery. Biochem. J 477, 3471–3497. [PubMed: 32965492]

(67). Sun D, Wu R, Zheng J, Li P, and Yu L (2018) Polyubiquitin chain-induced p62 phase separation drives autophagic cargo segregation. Cell Res. 28, 405–415. [PubMed: 29507397]

(68). Ciuffa R, Lamark T, Tarafder AK, Guesdon A, Rybina S, Hagen WJH, Johansen T, and Sachse C (2015) The Selective Autophagy Receptor p62 Forms a Flexible Filamentous Helical Scaffold. Cell Rep. 11, 748–758. [PubMed: 25921531]

(69). Yang Y, Willis TL, Button RW, Strang CJ, Fu Y, Wen X, Grayson PRC, Evans T, Sipthorpe RJ, Roberts SL, Hu B, Zhang J, Lu B, and Luo S (2019) Cytoplasmic DAXX drives SQSTM1/p62 phase condensation to activate Nrf2-mediated stress response. Nat. Commun 10, 3759. [PubMed: 31434890]

(70). Yasuda S, Tsuchiya H, Kaiho A, Guo Q, Ikeuchi K, Endo A, Arai N, Ohtake F, Murata S, Inada T, Baumeister W, Fernández-Busnadiego R, Tanaka K, and Saeki Y (2020) Stress- and ubiquitylation-dependent phase separation of the proteasome. Nature 578, 296–300. [PubMed: 32025036]

(71). Bertolaet BL, Clarke DJ, Wolff M, Watson MH, Henze M, Divita G, and Reed SI (2001) UBA domains of DNA damage-inducible proteins interact with ubiquitin. Nat. Struct. Biol 8, 417–422. [PubMed: 11323716]

(72). Li P, Banjade S, Cheng H-C, Kim S, Chen B, Guo L, Llaguno M, Hollingsworth JV, King DS, Banani SF, Russo PS, Jiang Q-X, Nixon BT, and Rosen MK (2012) Phase transitions in the assembly of multivalent signalling proteins. Nature 483, 336–340. [PubMed: 22398450]

(73). Banjade S, Wu Q, Mittal A, Peeples WB, Pappu RV, and Rosen MK (2015) Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. Proc. Natl. Acad. Sci 112, E6426–E6435. [PubMed: 26553976]

(74). Okruta J, Prakasha S, Wuc Q, Kellyd MJS, and Tauntona J (2015) Allosteric N-WASP activation by an inter-SH3 domain linker in Nck. Proc. Natl. Acad. Sci. U. S. A 112, E6436–E6445. [PubMed: 26554011]

(75). Semenov AN, and Rubinstein M (1998) Thermoreversible Gelation in Solutions of Associative Polymers. 1. Statics. Macromolecules 31, 1373–1385.

(76). Schneider JP, Pochan DJ, Ozbas B, Rajagopal K, Pakstis L, and Kretsinger J (2002) Responsive Hydrogels from the Intramolecular Folding and Self-Assembly of a Designed Peptide. J. Am. Chem. Soc 124, 15030–15037. [PubMed: 12475347]

(77). Pochan DJ, Schneider JP, Kretsinger J, Ozbas B, Rajagopal K, and Haines L (2003) Thermally Reversible Hydrogels via Intramolecular Folding and Consequent Self-Assembly of a de Novo Designed Peptide. J. Am. Chem. Soc 125, 11802–11803. [PubMed: 14505386]

(78). Tanaka F (2011) Polymer Physics : Applications to Molecular Association and Thermoreversible Gelation. Cambridge University Press.

(79). Semenov AN, and Rubinstein M (1998) Thermoreversible Gelation in Solutions of Associating Polymers. 2. Linear Dynamics. Macromolecules 31, 1386–1397.

(80). Harmon TS, Holehouse AS, Rosen MK, and Pappu RV (2017) Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. Elife 6, 1–31.

(81). Ranganathan S, and Shakhnovich EI (2020) Dynamic metastable long-living droplets formed by sticker-spacer proteins. Elife 9, 1–25.

(82). Mao YS, Zhang B, and Spector DL (2011) Biogenesis and function of nuclear bodies. Trends Genet. 27, 295–306. [PubMed: 21680045]

(83). Fox AH, Nakagawa S, Hirose T, and Bond CS (2018) Paraspeckles: Where Long Noncoding RNA Meets Phase Separation. Trends Biochem. Sci 43, 124–135. [PubMed: 29289458]

(84). Luo Y, Na Z, and Slavoff SA (2018) P-Bodies: Composition, Properties, and Functions. Biochemistry 57, 2424–2431. [PubMed: 29381060]

(85). Wei MT, Chang YC, Shimobayashi SF, Shin Y, Strom AR, and Brangwynne CP (2020) Nucleated transcriptional condensates amplify gene expression. Nat. Cell Biol 22, 1187–1196. [PubMed: 32929202]

(86). Hudson WH, and Ortlund EA (2014) The structure, function and evolution of proteins that bind DNA and RNA. Nat. Rev. Mol. Cell Biol 15, 749–760. [PubMed: 25269475]

(87). Wang X, Schwartz JC, and Cech TR (2015) Nucleic acid-binding specificity of human FUS protein. Nucleic Acids Res. 43, 7535–7543. [PubMed: 26150427]

(88). Loughlin FE, Lukavsky PJ, Kazeeva T, Reber S, Hock EM, Colombo M, Von Schroetter C, Pauli P, Cléry A, Mühlemann O, Polymenidou M, Ruepp MD, and Allain FHT (2019) The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity. Mol. Cell 73, 490–504.e6. [PubMed: 30581145]

(89). Beusch I, Barraud P, Moursy A, Cléry A, and Allain FH-T (2017) Tandem hnRNP A1 RNA recognition motifs act in concert to repress the splicing of survival motor neuron exon 7. Elife 6, 1–35.

(90). Chong PA, Vernon RM, and Forman-Kay JD (2018) RGG/RG Motif Regions in RNA Binding and Phase Separation. J. Mol. Biol 430, 4650–4665. [PubMed: 29913160]

(91). Stawiski EW, Gregoret LM, and Mandel-Gutfreund Y (2003) Annotating Nucleic Acid-Binding Function Based on Protein Structure. J. Mol. Biol 326, 1065–1079. [PubMed: 12589754]

(92). Vuzman D, and Levy Y (2012) Intrinsically disordered regions as affinity tuners in protein-DNA interactions. Mol. Biosyst 8, 47–57. [PubMed: 21918774]

(93). Roden C, and Gladfelter AS (2021) RNA contributions to the form and function of biomolecular condensates. Nat. Rev. Mol. Cell Biol 22, 183–195. [PubMed: 32632317]

(94). Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovsky A, Poser I, Bickle M, Rizk S, Guillén-Boixet J, Franzmann TM, Jahnel M, Marrone L, Chang Y, Sterneckert J, Tomancak P, Hyman AA, and Alberti S (2018) RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science (80-.). 360, 918–921.

(95). Rengifo-Gonzalez JC, El Hage K, Clément M-J, Steiner E, Joshi V, Craveur P, Durand D, Pastré D, and Bouhss A (2021) The cooperative binding of TDP-43 to GU-rich RNA repeats antagonizes TDP-43 aggregation. Elife 10, 1–30.

(96). Niaki AG, Sarkar J, Cai X, Rhine K, Vidaurre V, Guy B, Hurst M, Lee JC, Koh HR, Guo L, Fare CM, Shorter J, and Myong S (2020) Loss of Dynamic RNA Interaction and Aberrant Phase Separation Induced by Two Distinct Types of ALS/FTD-Linked FUS Mutations. Mol. Cell 77, 82–94.e4. [PubMed: 31630970]

(97). Monahan Z, Ryan VH, Janke AM, Burke KA, Rhoads SN, Zerze GH, O'Meally R, Dignon GL, Conicella AE, Zheng W, Best RB, Cole RN, Mittal J, Shewmaker F, and Fawzi NL (2017) Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. EMBO J. 36, 2951–2967. [PubMed: 28790177]

(98). Ryan VH, Perdikari TM, Naik MT, Saueressig CF, Lins J, Dignon GL, Mittal J, Hart AC, and Fawzi NL (2021) Tyrosine phosphorylation regulates hnRNPA2 granule protein partitioning and reduces neurodegeneration. EMBO J. 40, 1–22.

(99). Cohen TJ, Hwang AW, Restrepo CR, Yuan C-X, Trojanowski JQ, and Lee VMY (2015) An acetylation switch controls TDP-43 function and aggregation propensity. Nat. Commun 6, 5845. [PubMed: 25556531]

(100). Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, and Karpen GH (2017) Phase separation drives heterochromatin domain formation. Nature 547, 241–245. [PubMed: 28636597]

(101). Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, and Narlikar GJ (2017) Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature 547, 236–240. [PubMed: 28636604]

(102). Meehan RR, Kao CF, and Pennings S (2003) HP1 binding to native chromatin in vitro is determined by the hinge region and not by the chromodomain. EMBO J. 22, 3164–3174. [PubMed: 12805230]

(103). Keenen MM, Brown D, Brennan LD, Renger R, Khoo H, Carlson CR, Huang B, Grill SW, Narlikar GJ, and Redding S (2021) HP1 proteins compact dna into mechanically and positionally stable phase separated domains. Elife 10, 1–38.

(104). Nishibuchi G, Machida S, Osakabe A, Murakoshi H, Hiragami-Hamada K, Nakagawa R, Fischle W, Nishimura Y, Kurumizaka H, Tagami H, and Nakayama J (2014) N-terminal

phosphorylation of HP1α increases its nucleosome-binding specificity. Nucleic Acids Res. 42, 12498–12511. [PubMed: 25332400]

(105). Wang L, Gao Y, Zheng X, Liu C, Dong S, Li R, Zhang G, Wei Y, Qu H, Li Y, Allis CD, Li G, Li H, and Li P (2019) Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. Mol. Cell 76, 646–659.e6. [PubMed: 31543422]

(106). Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, and Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410, 120–124. [PubMed: 11242054]

(107). Aranda S, Mas G, and Di Croce L (2015) Regulation of gene transcription by Polycomb proteins. Sci. Adv 1, 1–16.

(108). Robinson AK, Leal BZ, Chadwell LV, Wang R, Ilangovan U, Kaur Y, Junco SE, Schirf V, Osmulski PA, Gaczynska M, Hinck AP, Demeler B, McEwen DG, and Kim CA (2012) The growth-suppressive function of the polycomb group protein polyhomeotic is mediated by polymerization of its sterile alpha motif (SAM) domain. J. Biol. Chem 287, 8702–8713. [PubMed: 22275371]

(109). Lin Y, Currie SL, and Rosen MK (2017) Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. J. Biol. Chem 292, 19110–19120. [PubMed: 28924037]

(110). Schuster BS, Reed EH, Parthasarathy R, Jahnke CN, Caldwell RM, Bermudez JG, Ramage H, Good MC, and Hammer DA (2018) Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. Nat. Commun 9, 1–12. [PubMed: 29317637]

(111). Meyer DE, and Chilkoti A (2004) Quantification of the Effects of Chain Length and Concentration on the Thermal Behavior of Elastin-like Polypeptides. Biomacromolecules 5, 846–851. [PubMed: 15132671]

(112). Simon JR, Eghtesadi SA, Dzuricky M, You L, and Chilkoti A (2019) Engineered Ribonucleoprotein Granules Inhibit Translation in Protocells. Mol. Cell 75, 66–75.e5. [PubMed: 31175012]

(113). Kelley FM, Favetta B, Regy RM, Mittal J, and Schuster BS (2021) Amphiphilic proteins coassemble into multiphasic condensates and act as biomolecular surfactants. Proc. Natl. Acad. Sci 118.

(114). Cui M, Wang X, An B, Zhang C, Gui X, Li K, Li Y, Ge P, Zhang J, Liu C, and Zhong C (2019) Exploiting mammalian low-complexity domains for liquid-liquid phase separation–driven underwater adhesive coatings. Sci. Adv 5, 1–13.

(115). Mohammadi P, Aranko AS, Landowski CP, Ikkala O, Jaudzems K, Wagermaier W, and Linder MB (2019) Biomimetic composites with enhanced toughening using silk-inspired triblock proteins and aligned nanocellulose reinforcements. Sci. Adv 5, 1–12.

(116). Shea J-E, Best RB, and Mittal J (2021) Physics-based computational and theoretical approaches to intrinsically disordered proteins. Curr. Opin. Struct. Biol 67, 219–225. [PubMed: 33545530]

(117). Murthy AC, Tang WS, Jovic N, Janke AM, Seo DH, Perdikari TM, Mittal J, and Fawzi NL (2021) Molecular interactions contributing to FUS SYGQ LC-RGG phase separation and co-partitioning with RNA polymerase II heptads. Nat. Struct. Mol. Biol 28, 923–935. [PubMed: 34759379]

(118). Schuster BS, Dignon GL, Tang WS, Kelley FM, Ranganath AK, Jahnke CN, Simpkins AG, Regy RM, Hammer DA, Good MC, and Mittal J (2020) Identifying sequence perturbations to an intrinsically disordered protein that determine its phase-separation behavior. Proc. Natl. Acad. Sci. U. S. A 117.

(119). Souza PCT, Alessandri R, Barnoud J, Thallmair S, Faustino I, Grünewald F, Patmanidis I, Abdizadeh H, Bruininks BMH, Wassenaar TA, Kroon PC, Melcr J, Nieto V, Corradi V, Khan HM, Domański J, Javanainen M, Martinez-Seara H, Reuter N, Best RB, Vattulainen I, Monticelli L, Periole X, Tieleman DP, de Vries AH, and Marrink SJ (2021) Martini 3: a general purpose force field for coarse-grained molecular dynamics. Nat. Methods 18, 382–388. [PubMed: 33782607]

(120). Zheng W, Dignon GL, Jovic N, Xu X, Regy RM, Fawzi NL, Kim YC, Best RB, and Mittal J (2020) Molecular Details of Protein Condensates Probed by Microsecond Long Atomistic Simulations. J. Phys. Chem. B 124, 11671–11679. [PubMed: 33302617]

(121). Paloni M, Bailly R, Ciandrini L, and Barducci A (2020) Unraveling Molecular Interactions in Liquid–Liquid Phase Separation of Disordered Proteins by Atomistic Simulations. J. Phys. Chem. B 124, 9009–9016. [PubMed: 32936641]

(122). Benayad Z, Von Bülow S, Stelzl LS, and Hummer G (2021) Simulation of FUS Protein Condensates with an Adapted Coarse-Grained Model. J. Chem. Theory Comput 17, 525–537. [PubMed: 33307683]

(123). Dignon GL, Zheng W, Kim YC, Best RB, and Mittal J (2018) Sequence determinants of protein phase behavior from a coarse-grained model. PLOS Comput. Biol (Ofran Y, Ed.) 14, e1005941. [PubMed: 29364893]

(124). Regy RM, Thompson J, Kim YC, and Mittal J (2021) Improved coarse-grained model for studying sequence dependent phase separation of disordered proteins. Protein Sci. 30, 1371–1379. [PubMed: 33934416]

(125). Tesei G, Schulze TK, Crehuet R, and Lindorff-Larsen K (2021) Accurate model of liquid-liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. Proc. Natl. Acad. Sci. U. S. A 118.

(126). Joseph JA, Reinhardt A, Aguirre A, Chew PY, Russell KO, Espinosa JR, Garaizar A, and Collepardo-Guevara R (2021) Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. Nat. Comput. Sci 1, 732–743. [PubMed: 35795820]

(127). Mammen Regy R, Zheng W, and Mittal J (2021) Using a sequence-specific coarse-grained model for studying protein liquid–liquid phase separation. Methods Enzymol. 1st ed. Elsevier Inc.

(128). Perdikari TM, Jovic N, Dignon GL, Kim YC, Fawzi NL, and Mittal J (2021) A predictive coarse-grained model for position-specific effects of post-translational modifications. Biophys. J 120, 1187–1197. [PubMed: 33582133]

(129). Regy RM, Dignon GL, Zheng W, Kim YC, and Mittal J (2020) Sequence dependent phase separation of protein-polynucleotide mixtures elucidated using molecular simulations. Nucleic Acids Res. 48, 12593–12603. [PubMed: 33264400]

(130). Kim YC, and Hummer G (2008) Coarse-grained Models for Simulations of Multiprotein Complexes: Application to Ubiquitin Binding. J. Mol. Biol 375, 1416–1433. [PubMed: 18083189]

(131). Sanyal T, Mittal J, and Shell MS (2019) A hybrid, bottom-up, structurally accurate, Gō-like coarse-grained protein model. J. Chem. Phys 151, 044111. [PubMed: 31370551]

(132). Latham AP, and Zhang B (2021) Consistent Force Field Captures Homologue-Resolved HP1 Phase Separation. Cite This J. Chem. Theory Comput 17, 3134–3144.

(133). Latham AP, and Zhang B (2022) Unifying coarse-grained force fields for folded and disordered proteins. Curr. Opin. Struct. Biol 72, 63–70. [PubMed: 34536913]

(134). Debiec KT, Whitley MJ, Koharudin LMI, Chong LT, and Gronenborn AM (2018) Integrating NMR, SAXS, and Atomistic Simulations: Structure and Dynamics of a Two-Domain Protein. Biophys. J 114, 839–855. [PubMed: 29490245]

(135). Thomasen FE, Pesce F, Roesgaard MA, Tesei G, and Lindorff-Larsen K (2022) Improving Martini 3 for Disordered and Multidomain Proteins. J. Chem. Theory Comput

(136). Babinchak WM, and Surewicz WK (2020) Liquid–Liquid Phase Separation and Its Mechanistic Role in Pathological Protein Aggregation. J. Mol. Biol 432, 1910–1925. [PubMed: 32169484]

(137). Wen J, Hong L, Krainer G, Yao Q-Q, Knowles TPJ, Wu S, and Perrett S (2021) Conformational Expansion of Tau in Condensates Promotes Irreversible Aggregation. J. Am. Chem. Soc 143, 13056–13064. [PubMed: 34374536]

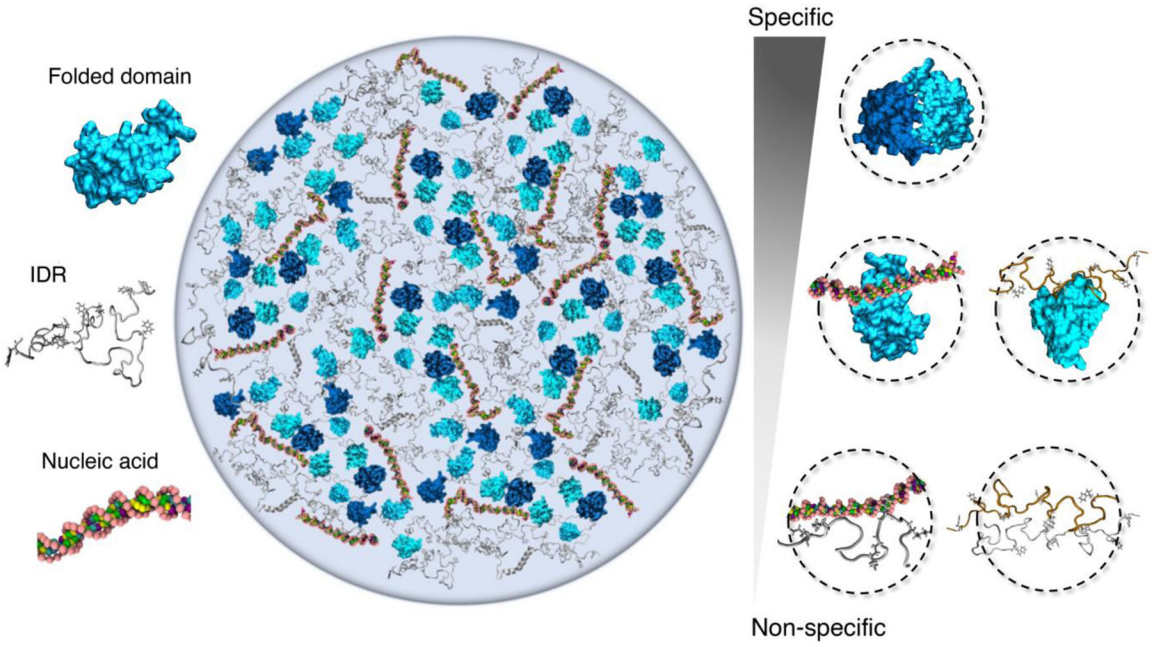

Figure 1. Summary of interactions that can modulate the phase separation of multi-domain proteins.

Depending on the domain architecture, a variety of inter-domain interactions which involve both folded and disordered regions can regulate the phase separation of multi-domain proteins. In terms of specificity, interactions between folded domains and those between folded domains and IDRs are usually stereospecific, i.e. they occur through well-defined surface patches (folded domains) and position-specific repeats, motifs or molecular recognition features (IDRs). Interactions between IDRs are generally non-stereospecific and involve a distributed network of repeats and/or motifs. Nucleic acids can play a critical role in regulating the formation and stability of protein condensates. Interactions between folded domains and nucleic acids can be either specific or non-specific, i.e. between positively-charged sidechains and the negatively-charged sugar-phosphate backbone. IDR-nucleic acid interactions are usually non-specific and involve a distributed network of amino acid repeats and/or motifs and the sugar-phosphate backbone.

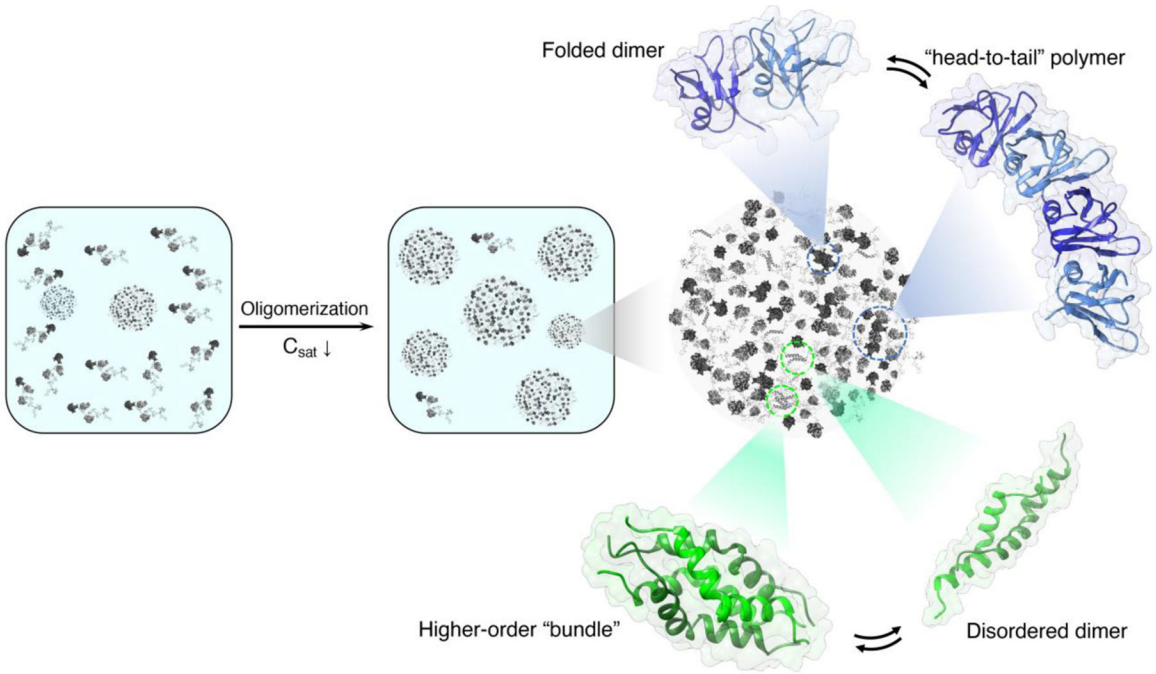

Figure 2. Enhancement of phase separation by homooligomerization through folded and disordered regions.

Homooligomerization can occur through both folded domains and disordered regions with transient secondary structure. TDP-43 N-terminal domain (PDB: 5MDI, aa:2–79) was chosen as an example to illustrate dimerization through a folded domain and TDP-43 C-terminal region (PDB: 2N3X, aa:315–350) was used to illustrate the same for a disordered region. Dimers that form initially can give rise to higher-order structures in a concentration-dependent manner and exert influence on condensate structure over longer length-scales. The ability to form stable (high affinity) and higher-order oligomers (>2 units) lead to a proportionate reduction in c_{sat}.

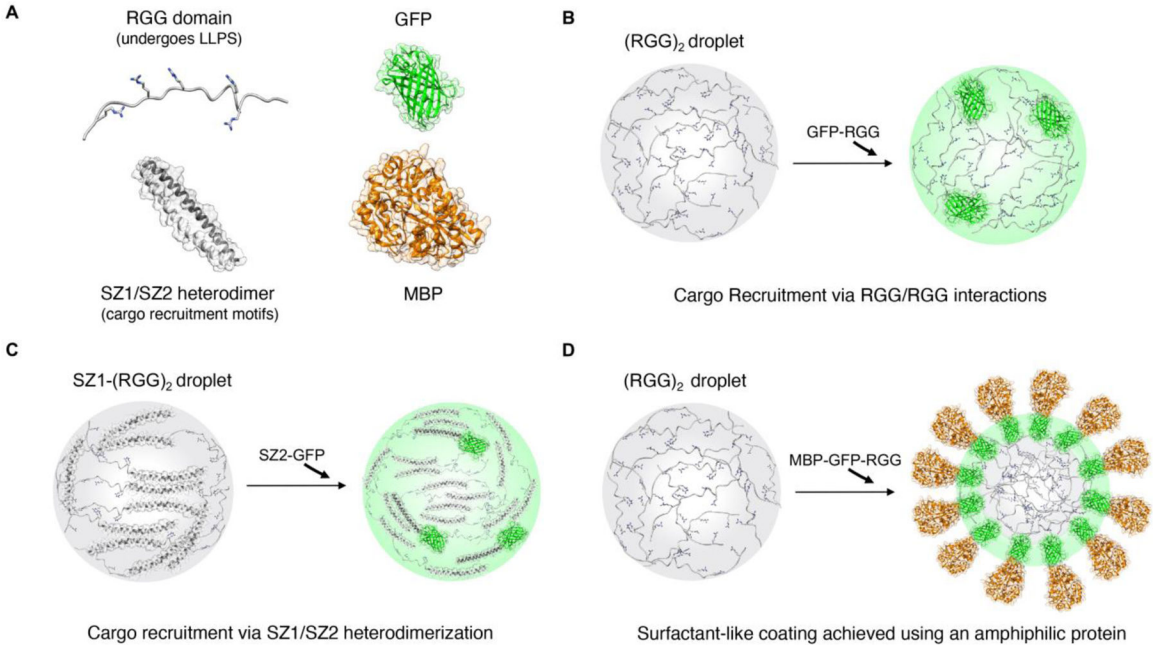

Figure 3. Controlling the composition and structural organization of biomolecular condensates using multi-domain protein constructs.

Phase-separating, disordered domains such as the LAF-1 RGG domain can be utilized as tandem repeats, e.g. $(RGG)_2$ and combined with other folded domains to control the recruitment of cargo proteins and structural organization within the condensed phase. **A.** A representative fragment of LAF1-RGG domain and crystal structures of green fluorescent protein (GFP, PDB: 1GFL), SYNZIP1/SYNZIP2 (SZ1/SZ2) coiled-coil heterodimer (PDB: 3HE5) and Maltose-binding protein (MBP, PDB: 1URD) are shown. B. A cargo protein such as GFP can be recruited to tandem $(RGG)_2$ droplets by either attaching it to a RGG domain that can co-phase separate with $(RGG)_2$ or C., by utilizing the SZ1-SZ2 interaction motifs. **D.** Structural properties of $(RGG)_2$ condensates can be controlled using MBP-based amphiphiles which form enveloped condensates, behave as surfactants and limit droplet size with increasing concentration.

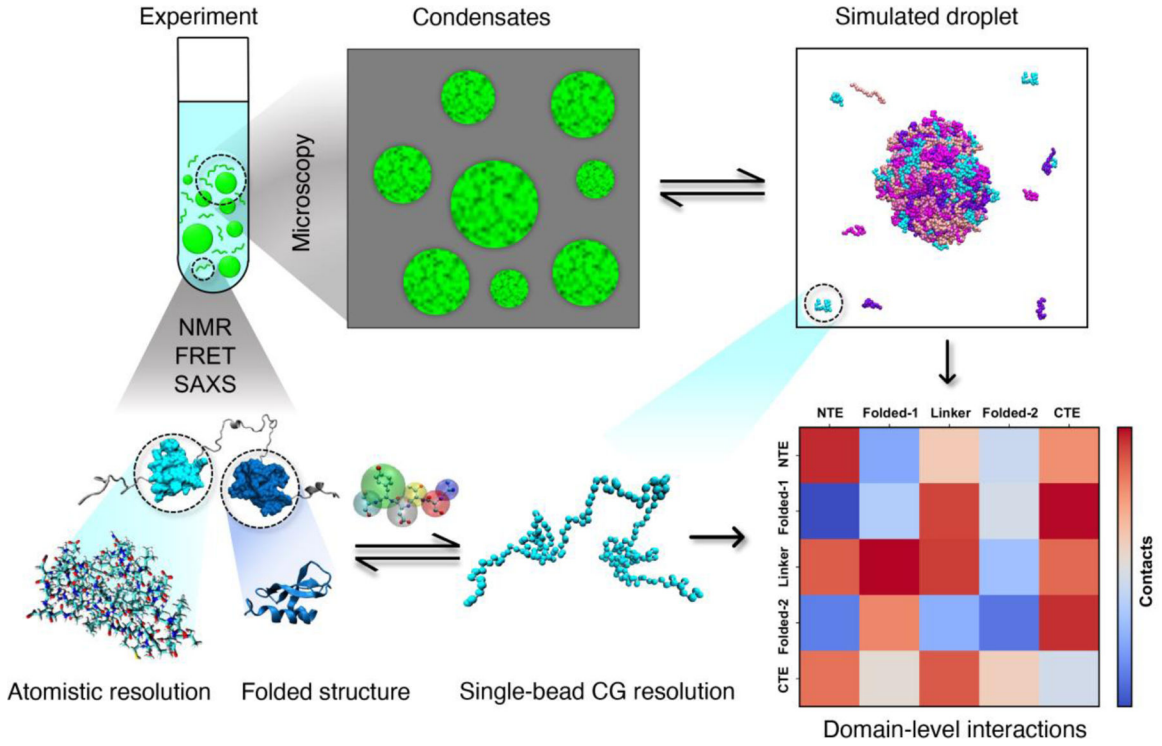

Figure 4. Modeling the phase behavior of multi-domain proteins at various length and time scales using molecular simulations.

Molecular simulations performed at atomistic and CG resolution can provide insights into the influence of domain-level interactions in dilute and droplet phases. Together with biophysical experiments, molecular simulations enhance our understanding of the structure, dynamics and function of multi-domain protein condensates.