Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

IDPs in macromolecular complexes: the roles of multivalent interactions in diverse assemblies

Ho Yee Joyce Fung[1,3], Melissa Birol[1,3] and Elizabeth Rhoades[1,2]

Intrinsically disordered proteins (IDPs) have critical roles in a diverse array of cellular functions. Of relevance here is that they are components of macromolecular complexes, where their conformational flexibility helps mediate interactions with binding partners. IDPs often interact with their binding partners through short sequence motifs, commonly repeated within the disordered regions. As such, multivalent interactions are common for IDPs and their binding partners within macromolecular complexes. Here we discuss the importance of IDP multivalency in three very different macromolecular assemblies: biomolecular condensates, the nuclear pore, and the cytoskeleton.

Addresses
[1] Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, United States
[2] Biochemistry & Molecular Biophysics Graduate Group, University of Pennsylvania, Philadelphia, PA 19104, United States

Corresponding author: Rhoades, Elizabeth (elizabeth.rhoades@sas.upenn.edu)
[3] Co-first authors.

Current Opinion in Structural Biology 2018, **49**:36–43

This review comes from a themed issue on **Macromolecular assemblies**

Edited by **Timm Maier** and **Kira Weissman**

https://doi.org/10.1016/j.sbi.2017.12.007

Introduction

The study of protein biochemistry has long been guided by the paradigm that a protein's function is closely tied to its three-dimensional structure. However, over the past 15 years, there has been a growing appreciation that many proteins do not fit within the structure–function paradigm [1–3]. These proteins, called intrinsically disordered proteins (IDPs), lack tertiary contacts and typically do not exhibit stable secondary structure. Similarly, intrinsically disordered regions (IDRs) are relatively long (>30 residues) stretches of disordered regions which flank globular domains on one or both ends [4]. IDPs and IDRs have been identified as key players in many critical cellular processes [5] as well as being implicated in a diverse array of devastating diseases [6–8]. Of particular relevance here, IDPs are abundant, and play important roles, in the assembly and functions of macromolecular complexes [9,10].

The characteristic features of IDPs make them ubiquitous in macromolecular complexes. IDPs are dynamic and flexible in solution; their extended conformations may permit them to 'search' efficiently for binding partners [11,12]. Their malleability allows them to adopt distinct conformations upon binding to different partners. As a consequence, the same IDP may be a component of several different complexes [13]. Moreover, IDPs are frequent targets for post-translational modifications (PTMs), which can tune their conformational states [14••,15] and interactions with binding partners [16]. Overviews of the types of interactions characteristic of IDPs can be found in several recent reviews [17,18].

IDPs typically interact with binding partners through short sequence motifs; short linear motifs (SLiMs), molecular recognition features (MoRFs), and low-complexity sequences (Box 1) [19–21]. Frequently two or more of these short motifs are found in IDPs or IDRs, making multivalent interactions common [19,21–24]. Multivalency arises when two or more binding sites are present on a protein allowing for simultaneous binding to partner molecules, increasing the overall avidity of the interaction (Box 1). It is important to a broad variety of biological interactions, including protein–carbohydrate interactions [25], clustering of cellular receptors [26], and virus–antibody binding [27]. Here we focus on the role of IDP/IDR multivalency in driving protein complex formation. Using three very different biological assemblies, we illustrate the diverse manifestations of multivalent interactions of disordered protein sequences.

Biomolecular condensates

The role of IDPs in biomolecular condensates (Box 1) is a topic which has garnered significant recent attention. Biomolecular condensates describe the micron-scale protein-rich assemblies resulting from liquid–liquid phase separation (Box 1) that are used by cells as a general mechanism for isolating internal material [28•,29–32]. These assemblies, which include nuclear bodies, Cajal bodies, P-bodies and granules, can play crucial roles in a variety of important biological processes [26,33–35]. They are differentiated from traditional organelles in that they lack a membrane barrier and therefore have the ability to exchange components in response to alterations to their environment. IDPs and IDR-containing proteins are

Box 1 *S*hort *Li*near *M*otifs (SLiMs) — short stretches of sequences (3–10 residues) that direct protein–protein interactions.

*Mo*lecular *R*ecognition *F*eatures (MoRFs) — intrinsically disordered regions of 10–40 residues which undergo a disorder to order transition upon binding to protein partners.

Low-complexity sequences — sequences that contain repeats of single or short amino acid motifs, with less diversity in amino acids than most protein sequences. While many low-complexity sequences are of unknown function, some modulate protein–protein and protein–nucleic acid interactions.

Multivalency — interactions between two or more binding sites on a protein with its binding partners. Achieved through having different types of binding sites interacting with different macromolecules, or repeats of the same binding sites interacting with one or more binding partner.

Liquid–liquid phase separation — demixing of a liquid phase from another, driven by energetically favorable interactions between components of different phases. One familiar example would be the separation of oil and water.

Biomolecular condensates — biomolecular condensates, also termed as pleiomorphic ensembles, liquid droplets or membraneless organelles, are compartments in the cell formed via liquid–liquid phase separation. Their formation is driven by interactions between molecules which are enriched in the phase separated droplets. These compartments are liquid-like as they can fuse with one another, are dynamic and can undergo rapid exchange with the surrounding environment.

enriched in many biomolecular condensates and have been reported to initiate phase separation both *in vitro* and *in cellulo* [36••,37,38]. In many of these cases, phase separation is driven by collective, weak, multivalent interactions between IDPs and/or their protein or nucleic acid partners [39]. PTMs to sequences involved can change the valency and intrinsic solubility of IDPs and thus tune their partitioning into liquid droplets [40•,41].

Phase separation driven by weak multivalent interactions may include heterologous electrostatic interactions between blocks of oppositely charged residues, either between proteins and nucleic acids or between mixtures of oppositely charged proteins (Figure 1, top right panel). The P granule protein LAF-1, thought to drive P-granule assembly *in vivo*, is an extensively studied example. The N-terminal disordered region of LAF-1 is an arginine/glycine (RGG) rich domain containing a mix of positively (arginine) and negatively (aspartic acid) charged residues. Electrostatic interactions between these alternatively charged regions results in its self-assembly into oligomeric structures and is necessary and sufficient for phase separation *in vitro* [36••]. Additionally, the RGG domain of LAF-1 is important for promoting dynamic protein–RNA interactions, shown to be critical in modulating droplet viscosity *in vitro* [36••,42]. Multivalent interactions between LAF-1 and other P granule proteins, including PGL-1, PGL-3, and VBH-1, all of which contain oppositely charged disordered regions, are thought to underlie the maintenance of dynamic but coherent P granule structure [36••].

A closely related example is the germ granule protein DEAD-box helicase 4 (DDX4), a RNA helicase with IDRs at both the N-terminus and C-terminus [37]. DDX4 proteins are essential for the assembly and maintenance of the P-granules in *Caenorhabditis elegan*s, and related assemblies in mammals and Drosophila [37]. Like LAF-1, the N-terminal IDR of DDX4 self-assembles and drives liquid droplet formation *in vitro*. It also contains RGG motifs as well as phenylalanine-glycine (FG) repeats which engage in cation–pi interactions with arginine residues within the RGG motifs, forming both intramolecular and intermolecular contacts [43]. DDX4 also provides an illustrative example as to how PTMs can alter interactions of IDPs; arginine methylation of DDX4 hinders its phase separation, likely due to decreased cation–pi interactions [37].

DDX4 and LAF-1 are examples of simple coacervation and do not require heterotypic interactions with a partner *in vitro*, although multiple additional partners are involved *in vivo*. By contrast, the disordered, negatively charged Nephrin intracellular domain (NICD) requires multiple positively charged ligands/counterions to undergo phase separation [44••]. Although, NICD resembles DDX4 and LAF-1 in that it is highly charged and the charges are clustered into blocks, it has a high net negative charge. When overexpressed in HeLa cells, NICD forms nuclear liquid condensates with positively charged nuclear proteins serving as charge neutralizers [44••].

A variety of other IDRs driving liquid–liquid phase separation *in vitro* through multivalent interactions have been identified, including the amyotrophic lateral sclerosis (ALS) related and stress granule associated proteins, such as hnRNPA1, FUS, and TDP43 [38,45,46]. By contrast to DDX4 and LAF-1, FUS harbors mostly polar and aromatic residues and lacks charged residues. Mutational studies have shown that the aromatic residues are important for driving phase separation of FUS IDR, in particular through pi–pi stacking [45]. Notably all the multivalent interactions discussed here — aromatic, polar, and charge–charge — do not appear to induce stable canonical structure and are short lived, consistent with the dynamic nature of phase-separated droplets in cells [47].

Nuclear pore complex

Segregation of nuclear and cytoplasmic processes by the nuclear membrane is a hallmark of eukaryotic cells. The nuclear pore complex (NPC) is a large megadalton complex in the nuclear membrane that controls the exchange of materials between the two compartments. Small molecules of less than ~5 nm in size can diffuse freely across the NPC while larger complexes are actively transported